# Kiến trúc hệ thống

Mô hình triển khai:

Diagram

Description automatically generated

Nguồn là dữ liệu

* Hệ thống bán hàng, lưu trữ ở schema **ORDERS**, chạy trên nền tảng oracle database:

|  |  |
| --- | --- |
| **Tên bảng** | **Ý nghĩa bảng** |
| SRC\_REGION | Bảng lưu thông tin vùng miền |
| SRC\_CITY | Bảng lưu thông tin Tỉnh/ thành phố |
| SRC\_CUSTOMER | Bảng lưu thông tin khách hàng |
| SRC\_PRODUCT | Bảng lưu thông tin về sản phẩm |
| SRC\_ORDERS | Bảng lưu thông tin chung của đơn hàng |
| SRC\_ORDER\_LINES | Bảng lưu thông tin chi tiết các mặt hàng trong đơn hàng |
| SRC\_SALE\_PERSON | Bảng lưu thông tin của nhân viên bán hàng |

Vùng Tạm (Staging), Schema trong database: **STAGING** chạy trên Oracle Database (user/pass: STAGING/STAGING)

Data Warehouse (DWH): Schema **ODI\_SALES\_PROD** chạy trên Oracle Database (pass: oracle123)

ETL sử dụng Oracle ODI

BI thì dùng PowerBI

# Xây dựng DWH cơ bản (chưa dùng SCD)

## 1. Source to Staging

Diagram

Description automatically generated

1. Trong Project trên ODI, tạo 1 **folder** đặt tên là: **SRC - STG** để lưu trữ toàn bộ job từ Source vào Staging
2. Tạo các bảng trên vùng Staging có tên: **<Họ tên>\_STG\_<tên bảng nguồn>**
3. Tạo ra các MAPPING đặt tên: **<họ tên>-SRC-STG-<tên bảng nguồn>** để thực hiện tải dữ liệu từ Nguồn về vùng tạm.**Load 1:1 tức là KHÔNG làm biến đổi dữ liệu, bảng ở source và bảng ở Staging giống hệt nhau** *(lưu ý: trong thực tế khi ETL bảng giao dịch sẽ chặn dữ liệu theo ngày, ví dụ: chỉ ETL dữ liệu bán hàng ngày hôm nay)*

## 2. Staging to DWH

Diagram

Description automatically generated

1. Trong Project trên ODI, Tạo 1 folder đặt tên là: **STG - DWH** để lưu trữ toàn bộ job từ Staging vào DWH
2. Tạo data model sau trong database (dùng Schema **<Họ tên>\_SALES\_PROD)**

Bảng danh mục (Dimension):

|  |  |  |  |
| --- | --- | --- | --- |
| **Bảng trên DWH** | **Cột** | **Ý nghĩa** | **Logic Mapping** |
| DIM\_DATE (học viên tự tìm trên Google SQL bảng Date) | DATE\_ID | Định dạng kiểu chữ, format là YYYYMMDD |  |
|  | DATE\_KEY | Định dạng là Date |  |
|  | MONTH | Tháng |  |
|  | Quarter | Quý |  |
|  | ... |  |  |
|  | …. |  |  |
|  | ... |  |  |
| DIM\_CUSTOMER | CUSTOMER\_ID | Mã Khách hàng | STG\_CUSTOMER.CUSTID |
| DIM\_CUSTOMER | DEAR |  | STG \_CUSTOMER.DEAR |
| DIM\_CUSTOMER | LAST\_NAME | Họ Khách hàng | STG \_CUSTOMER.LAST\_NAME |
| DIM\_CUSTOMER | FIRST\_NAME | Tên Khách hàng | STG \_CUSTOMER.FIRST\_NAME |
| DIM\_CUSTOMER | ADDRESS | Địa chỉ Khách hàng | STG \_CUSTOMER.ADDRESS |
| DIM\_CUSTOMER | PHONE | Số điện thoại Khách hàng | STG\_CUSTOMER.PHONE |
| DIM\_CUSTOMER | AGE | Tuổi Khách hàng | STG \_CUSTOMER.AGE |
| DIM\_PRODUCT | PRODUCT\_ID | Mã sản phẩm | STG \_PRODUCT.PRODUCT\_ID |
| DIM\_PRODUCT | PRODUCT\_NAME | Tên sản phẩm | STG \_PRODUCT.PRODUCT |
| DIM\_PRODUCT | PRICE | Giá tiền sản phẩm | STG \_PRODUCT.PRICE |
| DIM\_PRODUCT | CATEGORY | Loại sản phẩm | STG \_PRODUCT.FAMILY\_NAME |
| DIM\_SALES\_PERSON | SALES\_PERSON\_ID | Mã người bán hàng | STG\_SALES\_PERSON.SALES\_PERS\_ID |
| DIM\_SALES\_PERSON | FIRST\_NAME | Họ người bán hàng | SRC\_SALES\_PERSON.FIRST\_NAME |
| DIM\_SALES\_PERSON | LAST\_NAME | Tên người bán hàng | STG\_SALES\_PERSON.LAST\_NAME |
| DIM\_SALES\_PERSON | HIRE\_DATE | Ngày người bán hàng được thuê | STG\_SALES\_PERSON.HIRE\_DATE |
| DIM\_CITY | CITY\_ID | Mã thành phố | STG\_CITY.CITY\_ID |
| DIM\_CITY | CITY\_NAME | Tên thành phố | STG\_CITY.CITY |
| DIM\_CITY | POPULATION | Số dân | STG\_CITY.POPULATION |
| DIM\_CITY | COUNTRY\_NAME | Tên quốc gia | STG\_REGION.COUNTRY |
| DIM\_CITY | REGION\_NAME | Tên khu vực | STG\_REGION.REGION |
|  | <học viên tự điền tiếp ra file Excel> |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |
|  |  |  |  |

* Bảng tính toán (fact):

|  |  |  |  |
| --- | --- | --- | --- |
| **Bảng** | **Cột** | **Ý nghĩa** | **Công thức** |
| Fact\_ORDERS | ORDER\_DATE | Ngày phát sinh đơn hàng | = STG\_ORDERS.ORDER\_DATE |
| Fact\_ORDERS | CITY\_ID | Mã thành phố của đơn hàng | = STG\_CITY.CITY\_ID |
| Fact\_ORDERS | CUST\_ID | Mã khách hàng | = STG\_ORDERS.CUST\_ID |
| Fact\_ORDERS | ORDER\_ID | Mã đơn hàng | = STG\_ORDERS.ORDER\_ID |
| Fact\_ORDERS | SALE\_PERSON\_ID | Mã nhân viên bán hàng | = STG\_CUSTOMER.SALES\_PERS\_ID |
| Fact\_ORDERS | PRODUCT\_ID | Mã sản phẩm trong đơn hàng | = STG\_ORDERS\_LINES.PRODUCT\_ID |
| Fact\_ORDERS | QUANTITY | Số lượng sản phẩm được mua | = STG\_ORDER\_LINES.QTY |
| Fact\_ORDERS | UNIT\_PRICE | Đơn giá sản phẩm được mua | = STG\_PRODUCT.PRICE |
| Fact\_ORDERS | GROSS\_SALES | Doanh số bán được | = STG\_ORDER\_LINES.QTY \* STG \_PRODUCT.PRICE |

1. Xây các job đưa dữ liệu từ STG vào các bảng DWH, đặt tên **<học viên>-STG-DWH-<Tên bảng đích>**

## DWH to DataMart

Diagram

Description automatically generated

* Trong Project trên ODI, Tạo 1 folder đặt tên là: **DWH – DATA MART** để lưu trữ toàn bộ job từ DWH vào Data mart
* Tạo ra 1 Data Mart, đặt tên “RPT\_SALE\_MONTH”: Báo cáo tháng cho từng sản phẩm

|  |  |  |  |
| --- | --- | --- | --- |
| **Bảng** | **Cột** | **Ý nghĩa** | **Công thức** |
| RPT\_SALE\_MONTH | PRODUCT\_ID | Mã sản phẩm (INT) | = FACT\_ORDERS.PRODUCT\_ID |
| MONTH | Tên tháng (YYYYMM) | = TO\_CHAR( FACT\_ORDERS.DATE , ’YYYYMM’ ) |
| QUANTITY | Số lượng đã bán (FLOAT) | = SUM (FACT\_ORDERS.QUANTITY) |
| GROSS\_SALES | Doanh thu (FLOAT) | = SUM (FACT\_ORDERS.GROSS\_SALES) |

* Tạo ra 1 Data Mart, đặt tên “RPT\_REGION\_MONTH”: Báo cáo tháng cho từng khu vực

|  |  |  |  |
| --- | --- | --- | --- |
| **Bảng** | **Cột** | **Ý nghĩa** | **Công thức** |
| RPT\_REGION\_MONTH | REGION\_ID | Mã khu vực (INT) | = DIM\_REGION.REGION\_ID |
| MONTH | Tên tháng (YYYYMM) | = TO\_CHAR( FACT\_ORDERS.DATE , ’YYYMM’ ) |
| QUANTITY | Số lượng đã bán (FLOAT) | = SUM (FACT\_ORDERS.QUANTITY) |
| GROSS\_SALES | Doanh thu (FLOAT) | = SUM (FACT\_ORDERS.GROSS\_SALES) |

* Tạo các mapping để kéo dữ liệu từ DWH vào các data mart, đặt tên **<học viên>-DWH-DM-<Tên bảng đích>**

## 4. Luồng tổng (Load plan)

* Tạo Load Plan để thực thi toàn bộ các job
* Lập lịch chạy bằng Agent

## 5. BI Report

Kết nối theo tài liệu: https://inda.vn/powerbi/ket-noi-power-bi-desktop-voi-oracle-database/

Diagram

Description automatically generated

* Kết nối tới **RPT\_SALE\_MONTH** và **RPT\_REGION\_MONTH** và kéo ra báo cáo
* Dùng Power BI tạo ra filter Năm, Tháng

# Xây DWH Nâng cao (dùng SCD type 2)

Ở phục II.2. Staging to DWH, bạn hãy:

* Bạn hãy tự thiết kế bảng **DIM\_CUSTOMER\_SCD** để lưu lại lịch sử thay đổi của khách hàng (tham khảo bài SCD)
* Tạo ra 1 bảng **“FACT\_ORDERS\_NEW”,** học viên sẽ thêm Surrogate key của bảng DIM\_CUSTOMER\_SCD vào bảng Fact **“Fact\_ORDERS\_NEW”**

|  |  |  |  |
| --- | --- | --- | --- |
| **Bảng** | **Cột** | **Ý nghĩa** | **Công thức** |
| **FACT\_ORDERS\_NEW** | ORDER\_DATE | Ngày phát sinh đơn hàng | =STG\_ORDERS.ORDER\_DATE |
| **FACT\_ORDERS\_NEW** | CITY\_ID | Mã thành phố của đơn hàng | = STG\_CITY.CITY\_ID |
| **FACT\_ORDERS\_NEW** | CUST\_SCD | Mã khách hàng theo SCD (Surrogate key) | = **DIM\_CUSTOMER\_SCD**. CUST\_SCD |
| **FACT\_ORDERS\_NEW** | ORDER\_ID | Mã đơn hàng | = STG\_ORDERS.ORDER\_ID |
| **FACT\_ORDERS\_NEW** | SALE\_PERSON\_ID | Mã nhân viên bán hàng | = STG\_CUSTOMER.SALES\_PERS\_ID |
| **FACT\_ORDERS\_NEW** | PRODUCT\_ID | Mã sản phẩm trong đơn hàng | = STG\_ORDERS\_LINES.PRODUCT\_ID |
| **FACT\_ORDERS\_NEW** | QUANTITY | Số lượng sản phẩm được mua | = STG\_ORDER\_LINES.QTY |
| **FACT\_ORDERS\_NEW** | UNIT\_PRICE | Đơn giá sản phẩm được mua | = STG\_PRODUCT.PRICE |
| Fact\_ORDERS | GROSS\_SALES | Doanh số bán được | = STG\_ORDER\_LINES.QTY \* STG \_PRODUCT.PRICE |