**BỘ MÔN HỆ THỐNG THÔNG TIN – KHOA CÔNG NGHỆ THÔNG TIN**

Logo

Description automatically generated**ĐẠI HỌC KHOA HỌC TỰ NHIÊN THÀNH PHỐ HỒ CHÍ MINH, ĐẠI HỌC QUỐC GIA TP HCM**

Sinh viên thực hiện: 18120327 – Võ Ngọc Đức

GV phụ trách: Hồ Thị Hoàng Vy

  - HTTT PV TTKD

HỌC KỲ I – NĂM HỌC 2021-2022

**Báo cáo bài tập về nhà HW8**

1. **Source -> Stage:**

* **Chuẩn bị**: cài đặt các database đã được nhận từ giáo viên: Jade, Jupiter, WebTower9, NDS, Stage, metadata.
* Thực hiện đổ dữ liệu theo kiểu toàn bảng từ source stage.

Graphical user interface

Description automatically generated

Graphical user interface, text, application

Description automatically generated

* Dữ liệu đã được đổ thành công.
* Tùy chỉnh lại cài đặt thêm LSET và CET để thực hiện đổ dữ liệu các lần sau theo kiểu Incremental extract.
* Cài đặt các thuộc tính LSET, CET ứng với từng bảng trong stage vào bảng Data\_Flow trong metadata.

Graphical user interface, text

Description automatically generated

* Với mỗi luồn data flow đổ từ source -> stage sẽ bắt đầu bằng việc set giá trị CET tương ứng.

Graphical user interface, application

Description automatically generated

* Sau đó thực hiện truncate bảng.

Graphical user interface, application

Description automatically generated

* Tiếp theo thực hiện lấy các giá trị LSET, Cet từ metadata.

Graphical user interface, application

Description automatically generated

* Tiếp theo sẽ đến data flow tast chính, thực hiện đổ nguồn theo query sau:

Graphical user interface, application

Description automatically generated

* Sau khi đổ dữ liệu, thực hiện gắn lại LSET = CET.

Graphical user interface, application

Description automatically generated

Diagram, application

Description automatically generated

*Mô hình luồn đổ dữ liệu từ source -> stage cho bảng Nuoc.*

* Thực hiện tương tự cho các bảng khác.

Graphical user interface

Description automatically generated

* Dữ liệu sẽ chỉ được đổ khi có thêm dữ liệu mới hoặc dữ liệu được cập nhật sau lần cuối đổ dữ liệu.

1. **Stage -> NDS:**

* Với mỗi bảng trong stage ta đổ vào NDS theo mô hình sau:

A screenshot of a computer

Description automatically generated

* Đầu tiên thực hiện load dữ liệu trong bảng ở stage, -> thêm cột SourceID.

Graphical user interface, text

Description automatically generated

* Thực hiện look up trong NDS xem dữ liệu đã tồn tại chưa.

Graphical user interface, application

Description automatically generated

* Tồn tại rồi thì thực hiện thêm cột Update và thực hiện update dữ liệu.

Graphical user interface, text

Description automatically generated

Graphical user interface, application

Description automatically generated

* Nếu chưa có thì thêm cột createdDate, updatedDate, và thêm mới dữ liệu bằng OLE DB Destination.

Graphical user interface, text

Description automatically generated

Graphical user interface, application

Description automatically generated

* Kết quả khi thực hiện với bảng LoaiCuaHang:

Graphical user interface, text, application

Description automatically generated

A screenshot of a computer

Description automatically generated with medium confidence

Graphical user interface, text, application

Description automatically generated

* **Với các bảng có chứa khóa ngoại** thì sẽ được thực hiện đổ dữ liệu sau các bảng khác. Và đồng thời thêm các bước look up để thực hiện thay thế khóa ngoại thành khóa SK trong NDS cho phù hợp.

A screenshot of a computer

Description automatically generated

Mô hình với bảng có chứa khóa ngoại – Bảng CuaHang.

Graphical user interface, text, application

Description automatically generated

Graphical user interface, text

Description automatically generated

*Kết quả thu được.*

1. **NDS -> DDS:**

* Trong metadata tại bảng data\_flow ta có các dòng giá trị để lưu trữ nhãn thời gian LSET và CET cho các bảng trong DDS.
* Bước đầu thực hiện việc đổ dữ liệu từ NDS vào DDS ta cần thực hiện set giá trị CET mới, và load vào giá trị LSET.

Graphical user interface, application

Description automatically generated

Thực hiện set giá trị CET mới vào bảng dataflow.

Graphical user interface, application

Description automatically generated

Lấy giá trị LSET và CET tương ứng.

Graphical user interface, application

Description automatically generated

Các giá trị nhận được sẽ được lưu vào các variable.

* Tiếp theo sử dụng Data flow task để thực hiện đổ dữ liệu.
* Load dữ liệu từ NDS theo nhãn thời gian.

Graphical user interface, text, application

Description automatically generated

* Sau đó thực hiện sử dụng slowly changing dimension để xử lý với các thuộc tính thay đổi chậm trong NDS.
* Với bảng KhachHang thì ở đây các thuộc tính DiaChi, NgheNghiep, SoThich sẽ là thuộc tính historical. Sử dụng loại thay đổi 2.

Graphical user interface, text

Description automatically generated

* Tiếp theo thiết lập cột lưu trạng thái dữ liệu.

Graphical user interface, application

Description automatically generated

* Nhấn next tiếp tục rồi finish, ta được mô hình data flow để xử lý thuộc tính lịch sử.

Diagram

Description automatically generated

* Cuối cùng thực hiện set lại LSET = CET.

Graphical user interface, application

Description automatically generated

* Kết quả thực thi lần đầu:

Diagram

Description automatically generated

Graphical user interface, application, table

Description automatically generated

* Thực hiện test lại, khi thay đổi job của CusNK 008V thành mining trong NDS.

Graphical user interface, text, application, email

Description automatically generated

* Chạy lại flow đến DDS cho bảng KhachHang, ta được kết quả có 1 dòng dữ liệu thay đổi.

A screenshot of a video game

Description automatically generated with medium confidence

* Kiểm tra trong Database ta thấy dòng dữ liệu cũ đã chuyển thành trangthai 0 và có thêm dòng mới với nghề mới được chỉnh sửa.

Graphical user interface, text, application, email

Description automatically generated

* **Thực hiện tương tự với bảng Sản phẩm.**

A screenshot of a video game

Description automatically generated with medium confidence

Diagram

Description automatically generated

* Các thuộc tính thay đổi chậm gồm gia, LoaiSP, TenSP.

Graphical user interface, text

Description automatically generated

Graphical user interface, text, application

Description automatically generated

* **Đối với bảng Cửa hàng:** thực hiện tương tự. Tuy nhiên để trong DDS có tên loại cửa hàng thì ở lúc load dữ liệu, ta cần merge join 2 bảng Cửa hàng và loại cửa hàng trong NDS.

Diagram

Description automatically generated

Graphical user interface

Description automatically generated

* Các thuộc tính lịch sử là nuoc, maloaiCH, tenLoaiCH.

Graphical user interface, text

Description automatically generated

* Kết quả load thành công:

Graphical user interface, text, application

Description automatically generated