

Sergio Moreno

◀ SSAS 2005 ▶

Desarrollo de
AVA_PIS_Olap 2005

**Cubos de analisis
para AVA PI System**

Autor: Sergio Moreno
Ultima modificación: 20. sep. 2008
Creado el: 20. sep. 2008

AVA_PIS_OLAP 2005

Cubos de análisis para AVA PI System

1. Introducción

El presente trabajo consistió en tomar los cubos desarrollados en la versión de Analysis Services 2000 para actualizarlos a la versión Analysis Services 2005, realizando las mejoras necesarias para su correcto funcionamiento.

1.1. Procesos Sujetos de análisis

Los procesos sujetos de análisis son:

- Balance diario de los tanques de petróleo y agua
- Tickets de movimiento (en tanques y medidores)
- Disposición y pruebas de pozos (composición)
- Alocación de gas y petróleo en cada estructura de pozo y reservorio
- Mediciones y análisis cromatográfico en líquidos y gaseosos

Para cada uno de los anteriores procesos se realizó un levantamiento de información y se proporcionaron hojas de cálculo de ejemplo junto con las estructuras de datos de donde se extraer la información pertinente.

2. Arquitectura

2.1. Base de Datos de Staging

En la base de datos de staging se hacen todas las transformaciones necesarias para poder alimentar los cubos de análisis. Para este proyecto la base se llama AVAPIS_Datamart y está compuesta por 3 tipos de tablas:

- Tablas temporales
- Tablas de dimensión
- Tablas de Hecho
- Tablas de Configuración y registro

2.1.1. Tablas temporales

La funcionalidad de las tablas temporales es replicar la estructura de la base de datos transaccional para hacer las transformaciones a partir de tablas locales y no foráneas.

Estas tablas se limpian completamente cada vez que se realiza una actualización que es realizada a través de paquetes de transformación SSIS.

En los casos de tablas temporales que deben después cargar las tablas de hecho por lo general poseen una fecha de transacción o de instantánea, para procesarlas incrementalmente se creo una tabla de parámetros PAR_ con un campo DaysDelay. Este campo determina la cantidad de registros que serán extraídos de la base de datos transaccional a partir de la fecha actual.

Todas estas tablas tienen el prefijo **TMP**

2.1.2. Tablas de dimensión

Los registros de estas tablas nunca se eliminan sino que se actualizan y aumentan. Cada tabla de dimensión consta de un procedimiento almacenado asociado que la actualiza además de una vista que muestra los datos actuales de la dimensión (según la última actualización realizada) a partir de las tablas temporales.

Para todas las tablas de dimensiones se crearon llaves sustitutas y auto generadas de modo de no crear ninguna dependencia adicional del sistema transaccional.

Todas las tablas de dimensión tienen el prefijo **DIM**, los procedimientos almacenados asociados tienen el prefijo **sp_set_Dim** y las vistas asociadas **VW_Dim**

2.1.3. Tablas de hecho

Estas son las tablas que irán aumentando con el tiempo y su proceso de actualización consiste en eliminar e insertar nuevamente los registros que existen en las tablas temporales.

Todas estas tablas están definidas con llaves foráneas a las tablas de dimensión relacionadas para asegurar la consistencia. además de eso para cada tabla de hecho existe una vista que contiene la información sacada hasta cierta fecha determinada por el **DaysDelay** de las tablas temporales. A la vez existe un procedimiento almacenado que se encarga de eliminar los registros que se encuentren ya en la vista (tomando la llave de la tabla de hechos) y de luego insertar estos registros a la tabla de hechos.

Todas las tablas de hecho tienen el prefijo **Fact**, los procedimientos almacenados asociados tienen el prefijo **sp_set_Dim** y las vistas asociadas **VW_Dim**

2.1.4. Tablas de configuración y registro

Existen dos tablas:

- *DM_PAR_Parameters*
Parámetros para los paquetes SSIS
- *sysdtslog90*
Log de ejecución de los paquetes SSIS

la primera contiene los parámetros de ejecución de la extracción de información del sistema transaccional, la otra contiene los Logs de ejecución de los paquetes SSIS

2.2. Servicios de Integración

Los servicios de integración están dentro del proyecto de integración AVAPIS_SIS incluido en el proyecto.

La estructura de los paquetes es la siguiente:

cinco paquetes de extracción:

- E_COMMON (Temporales comunes usados en varios cubos)
- E_ALLOC (Temporales de alocacion)
- E_TANK (Temporales de análisis de tanques)
- E_METER (Temporales de análisis de medidores)
- E_WELL (Temporales de análisis de pozos)

Un paquete ETL_MAIN que llama a los anteriores 5, ejecuta las transformaciones a través de los procedimientos almacenados y luego procesa la base de datos multidimensional

(E por extracción , T por transformación y L por cargado - loading)

En el paquete ETL_MAIN se ha tratado de no seguir una secuencia lineal de los paquetes de modo de tratar de aprovechar al máximo la capacidad de procesamiento simultáneo del servidor

Todos los paquetes SIS tienen una conexión con la base de staging y además con la transaccional si se trata de procesos de extracción. en el caso de que se tuviera que mover la base de datos de staging y el servidor de análisis a otro servidor, bastaría con cambiar las conexión compartida a la base transaccional y volver a generar los paquetes de extracción para reemplazarlos en el servidor de integración

Todos los paquetes de extracción siguen una lógica sencilla:

1. Efectuar un TRUNCATE de la tabla(s) temporal(es) destino
2. Consultar los parámetros de tiempo para atrás que se tomara para las tablas de hecho
3. Realizar el traspaso de la información

2.3. Servicios de Análisis

Está compuesta de los siguientes cubos de análisis:

2.3.1. Cubos de tanques

1. TANK_OilTankDataDay

Para el balance diario de tanques de petróleo

Tablas de Hecho: DM_Fact_OilTankDataday

Dimensiones: DIM Date y DIM OilTankStructure

Procedimiento almacenado de carga: sp_set_Fact_OilTankDataDay

En este cubo todas las medidas se agregan como sumas con excepción de dos : “Oil Initial Stock” y “Oil Final Stock” que toman el primero de los hijos con valor y el ultimo hijo respectivamente

2. *TANK_WaterTankDataDay*

Para el balance diario de tanques de agua

Tablas de Hecho: *DM_Fact_WaterTankDataday*

Dimensiones: DIM Date y DIM WaterTankStructure

Procedimiento almacenado de carga: *sp_set_Fact_WaterTankDataDay*

En este cubo todas las medidas se agregan como sumas con excepción de dos : “Water Initial Stock” y “Water Final Stock” que toman el primero de los hijos con valor y el ultimo hijo respectivamente

3. *TANK_OilTankTicketData*

Para transferencias entre distintos elementos de la planta por medio de tickets medidos en los tanques

Tablas de Hecho: *DM_Fact_OilTankTicketData*

Dimensiones: DIM Date, DIM OilTankStructure, DIM TicketElements (2 veces), DIM Oil Ticket, DIM Ticket Abstract

Procedimiento almacenado de carga: *sp_set_Fact_WaterTankDataDay*

En este cubo todas las medidas se agregan como sumas.

4. *TANK_OilMeterTicketData*

Para transferencias entre distintos elementos de la planta por medio de tickets medidos en los medidores

Tablas de Hecho: *DM_Fact_OilMeterTicketData*

Dimensiones: DIM Date, DIM TankOilMeterStructure, DIM TicketElements (2 veces), DIM OilMeterOilTicket, DIM Ticket Abstract

Procedimiento almacenado de carga: *sp_set_Fact_WaterTankDataDay*

En este cubo todas las medidas se agregan como sumas.

2.3.2. Cubos de Medidores

1. *METER_Liquid*

Contiene información en turnos diarios de los medidores de líquidos

Tablas de Hecho: *DM_Fact_MeterLiquidDataDay*

Dimensiones: DIM Date, DIM MeterLiquidStructure, DIM MeterOperation

Procedimiento almacenado de carga: *sp_set_Fact_MeterLiquidDataDay*

En este cubo todas las medidas se agregan como sumas con excepcion de “Correct API” que se agrega como un promedio ponderado

2. *METER_GASCromatographyDay*

Contiene información diaria de medidores de gases además de sus medidas de cromatografía

Tablas de Hecho: *DM_Fact_CromatographyEnergyDay*

Dimensiones: DIM Date, DIM MeterGasStructure, DIM MeterTestTtype

Procedimiento almacenado de carga: *sp_set_Fact_CromatographyEnergyDay*

En este cubo todas las medidas se agregan como sumas

3. *METER_GASDataDayControl*

Contiene información en turnos diarios de los medidores de gases

Tablas de Hecho: *DM_Fact_CromatographyEnergyDay*,
DM_Fact_GasMeterValveDataDay, *DM_Fact_GasMeterValveDataDay*

Dimensiones: DIM Date, DIM MeterGasStructure, DIM MeterOperation

Procedimiento almacenado de carga: *sp_set_Fact_CromatographyEnergyDay*,
sp_set_Fact_MeterGasDataDay, *sp_set_Fact_MeterGasValveDataDay*

En este cubo todas las medidas se agregan como sumas excepto : Gravity, CO2, Dew Point y H2O Content que se calculan como promedio ponderado

2.3.3. Cubos de Pozos

1. *WELL_WellDisposition*

Contiente la medicion diaria de los pozos hasta nivel de reservorios

Tablas de Hecho: *DM_Fact_WellDisposition*

Dimensiones: DIM Date, DIM WellStructure, DIM AllocationReference, DIM WellOperation, DIM WHCK

Procedimiento almacenado de carga: *sp_set_Fact_WellDisposition*

En este cubo todas las medidas se agregan como sumas

2. ***WELL_WellTestCromatography***

Contiene las mediciones tomadas en las pruebas de pozo así como las medidas de cromatografía obtenidas.

Tablas de Hecho: *DM_Fact_WellProductionTestCromatography*,
DM_Fact_WellProductionTest

Dimensiones: DIM WellTestDate, DIM WellStructure, DIM WHCK, DIM WellTest, Dim WellSeparatorData

Procedimiento almacenado de carga: *sp_set_Fact_WellProductionTest*

En este cubo todas las medidas se agregan como sumas

2.3.4. **Cubos de Alocación**

1. ***ALLOCATION_DailyReservoir***

Contiene la alocación por línea y reservorio

Tablas de Hecho: *DM_Fact_AllocationReservoirDaily*,
DM_Fact_AllocationReservoirMonth

Dimensiones: DIM Date, DIM WellStructure, DIM WHCK, DIM Reservoir, Dim WellOperation

Procedimientos almacenados de carga: *sp_set_Fact_AllocationReservoirDaily*,
sp_set_Fact_AllocationReservoirMonth

En este cubo todas las medidas toman y suman desde la data diaria si es que no se está agregando a nivel de meses, en ese caso se toma la data de la segunda tabla de hechos

2. ***ALLOCATION_DailyComposition***

Contiene la alocación por Completion Zone

Tablas de Hecho: *DM_Fact_AllocationCompositionDaily*,
DM_Fact_AllocationCompositionMonth

Dimensiones: DIM Date, DIM WellStructure, DIM WHCK, Dim WellOperation

Procedimientos almacenados de carga: *sp_set_Fact_AllocationCompositionDaily*,
sp_set_Fact_AllocationCompositionMonth

En este cubo todas las medidas toman y suman desde la data diaria si es que no se está agregando a nivel de meses, en ese caso se toma la data de la segunda tabla de hechos

3. ***ALLOCATION_DailyReservoirCumm***

Contiene la alocacion por reservorio calculada de manera mensual y acumulada

Tablas de Hecho: *DM_Fact_AllocationReservoirCumm*,

Dimensiones: DIM Date, DIM WellStructure, DIM Reservoir, Dim WellOperation

Procedimiento almacenado de carga: *sp_set_Fact_AllocationReservoirCumm*

En este cubo todas las medidas se suman

4. *ALLOCATION_DailyCompositionCumm*

Contiene la alocacion por Completion Zone calculada de manera mensual y acumulada

Tablas de Hecho: *DM_Fact_AllocationCompositionCumm*,

Dimensiones: DIM Date, DIM WellStructure, DIM Reservoir, Dim WellOperation

Procedimiento almacenado de carga: *sp_set_Fact_AllocationCompositionCumm*

En este cubo todas las medidas se suman

2.3.5. Dimensiones

Todas las dimensiones tienen el prefijo **DIM** y tienen dentro de la base de datos de staging una vista y un procedimiento almacenado de carga (a menos que se cargue directamente de la tabla de hechos)

las vistas comienzan con el prefijo VW_Dim y los procedimientos almacenados con *sp_set_Dim*.

Las dimensiones son:

1. *TicketElements*

Representan todos los orígenes y destinos de transacciones de los cubos 3 y 4 (ticketData)

se encuentra agrupada por unidad de producción, compañía y elemento.

Tabla de Dimensión: *DM_Dim_TicketElements*

Vista asociada: *VW_Dim_TicketElements*

Procedimiento almacenado de carga: *sp_set_Dim_TicketElements*

2. *MeterGasStructure*

Representan todos los medidores de gas y están agrupados por unidad de producción, compañía, tipo de medidor y medidor. Los tipos de medidores son Gas, Logic y Valve

Tabla de Dimensión: *DM_Dim_MeterStructure*

Vista asociada: *VW_Dim_MeterStructure*

Procedimiento almacenado de carga: *sp_Set_Dim_MeterStructure*

3. TankOilMeterOilTicket

Representa el código de ticket asignado a una transacción registrada en un medidor, esta es una dimension simple

Esta dimensión es sacada directamente de la tabla de hechos por tratarse de una codificación

4. TankOilMeterStructure

Representa todos los medidores de tanque y está agrupada por unidad de producción, compañía y medidor

Tabla de Dimensión: *DM_Dim_MeterStructure*

Vista asociada: *VW_Dim_MeterStructure*

Procedimiento almacenado de carga: *sp_Set_Dim_MeterStructure*

5. TankWaterStructure

Representa todos los tanques de agua y está agrupada por unidad de producción, compañía y tanque

Tabla de Dimensión: *DM_Dim_WaterTankStructure*

Vista asociada: *VW_Dim_WaterTankStructure*

Procedimiento almacenado de carga: *sp_Set_Dim_WaterTankStructure*

6. TicketAbstract

Representa una mini dimension de los tickets que aglomera el tipo de operacion (venta, transferencia...), el Estado del Ticket y el tipo de transporte usado. Aqui no existe ninguna jerarquía

Tabla de Dimensión: *DM_Dim_TicketAbstract*

Vista asociada: *VW_Dim_TicketAbstract*

Procedimiento almacenado de carga: *sp_Set_Dim_TicketAbstract*

7. TankOilTicket

Representa el código de ticket asignado a una transacción registrada en un tanque, esta es una dimension simple

Esta dimensión es sacada directamente de la tabla de hechos por tratarse de una codificación

8. MeterTestType

Es el tipo de test cromatográfico que aplican los medidores al hacer la cromatografía

Esta dimensión es sacada directamente de la tabla de hechos
DM_Fact_CromatographyEnergyDay

9. MeterLiquidStructure

Representan todos los medidores de líquidos y están agrupados por unidad de producción, compañía, tipo de medidor y medidor. Los tipos de medidores son NGL, Logic, Water y Oil

Tabla de Dimensión: *DM_Dim_MeterStructure*

Vista asociada: *VW_Dim_MeterStructure*

Procedimiento almacenado de carga: *sp_Set_Dim_MeterStructure*

10. WellStructure

Representa la estructura de 6 niveles de un pozo, que incluye : Unidad de producción, Bloque, Campo, Pozo, Linea, y Zona de Reservorio

Tabla de Dimensión: *DM_Dim_WellStructure*

Vista asociada: *VW_Dim_WellStructure*

Procedimiento almacenado de carga: *sp_Set_Dim_WellStructure*

11. WellAllocationReference

Dentro de al análisis de disposición del pozo se refiere a los distintos códigos de alocación asociados

Esta dimensión es sacada directamente de la tabla de hechos

DM_Fact_WellDisposition

12. WellTestDate

Son las distintas fechas donde se hicieron las pruebas de pozo

Tabla de Dimensión: *DM_Dim_Date (filtrada)*

Procedimiento almacenado de carga: *sp_Set_Dim_Date*

13. WellTest

Representa el codigo asignado a una prueba de pozo

Esta dimensión es sacada directamente de la tabla de hechos por tratarse de una codificación

14. Reservoir

Es una dimensión complementaria de la disposición de los pozos y las pruebas que dice de que se reservorio se produce

Tabla de Dimensión: *DM_Dim_Reservoir*

Vista asociada: *VW_Dim_Reservoir*

Procedimiento almacenado de carga: *sp_Set_Dim_Reservoir*

15. WHCK

Choque

Esta dimensión es sacada directamente de las tablas de hecho que contienen la medida WHCK

16. Date

Es la dimensión básica de tiempo dividida en año, mes y día. Se han generado las fechas desde 1990 hasta 2020

Tabla de Dimensión: *DM_Dim_Date*

Procedimiento almacenado de carga : *sp_Set_Dim_Date*

17. AllocOperation

Es una minidimension que enumera los distintos operation status y los distintos data items(turnos) que se dan en las alocaiones de pozo

Tabla de Dimensión: *DM_Dim_MeterOperation*

Vista asociada: *VW_Dim_MeterOperation*

Procedimiento almacenado de carga: *sp_Set_Dim_MeterOperation*

18. MeterOperation

Es una minidimension que enumera los distintos operation status y los distintos data items (turnos) que se dan en los medidores

Tabla de Dimensión: *DM_Dim_MeterOperation*

Vista asociada: *VW_Dim_MeterOperation*

Procedimiento almacenado de carga: *sp_Set_Dim_MeterOperation*

19. WellOperation

Es una minidimension que enumera los distintos operation status y los distintos data items (turnos) que se dan en los pozos

Tabla de Dimensión: *DM_Dim_MeterOperation*

Vista asociada: *VW_Dim_MeterOperation*

Procedimiento almacenado de carga: *sp_Set_Dim_MeterOperation*

20. WellSeparatorData

Dentro de las pruebas de pozo divide las que fueron hechas con High Presure or Low Presure

Esta dimensión es sacada directamente de la tabla de hechos

DM_Fact_WellProductionTestCromatography