

4.3. Arquitectura

Luego del análisis realizado de los requerimientos funcionales y no funcionales es necesario definir la base de herramientas y componentes que compondrán la arquitectura del sistema de BI.

Como se puede identificar en la fig. 4.27, donde se pueden visualizar los distintos componentes que forman parte de cada capa, se presenta una arquitectura en varios niveles, logrando desacoplar las responsabilidades de cada nivel. Fuentes de datos, integración y limpieza, repositorio de datos, herramientas de análisis de datos y metadatos, son algunos de los niveles identificados.

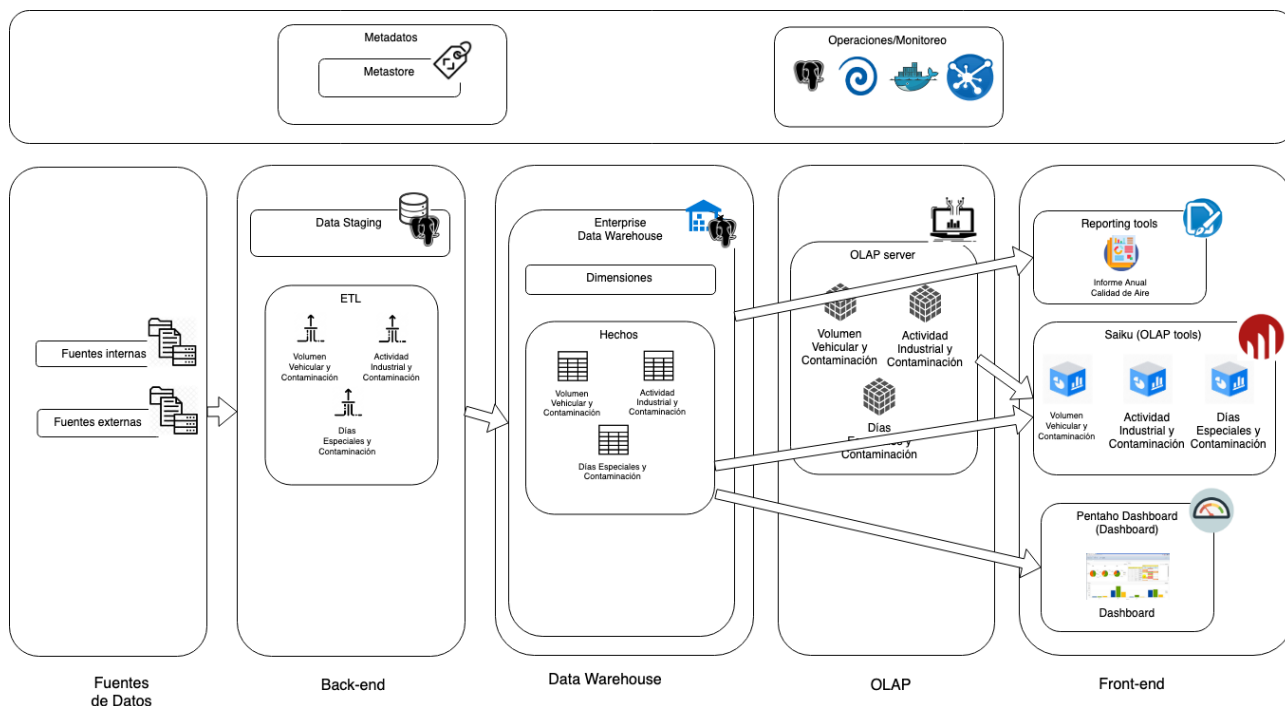


Figura 4.27: Arquitectura sistema de BI

4.3.1. Fuentes de datos

La capa de Fuentes de datos representa las diferentes fuentes de datos que alimentan los datos del DW. La fuente de datos puede estar en cualquier formato: archivo de texto plano, base de datos relacional, otros tipos de base de datos, archivo Excel, etc.

Como se explicitó en la etapa de Análisis, y se muestra en fig. 4.28, se logran identificar fuentes de datos internas del Servicio de Calidad de Aire y externas de Datos Abiertos, ANCAP y ADME.

4.3.2. Back-end

La capa de Back-end, ver fig. 4.29, contiene tres secciones importantes: Capa de extracción de datos, Área de pruebas y procesos ETL.

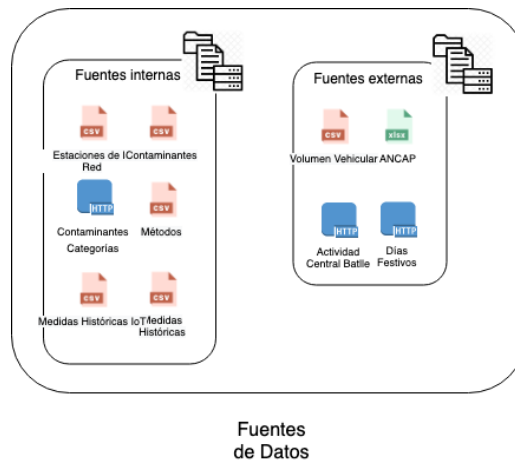


Figura 4.28: Fuentes de Datos

Dentro del Área de pruebas (Data Staging) podemos visualizar las distintas tablas basadas en las fuentes de datos que se identificaron, tanto internas como externas.

En los procesos de ETL se define uno por cada requerimiento identificado.

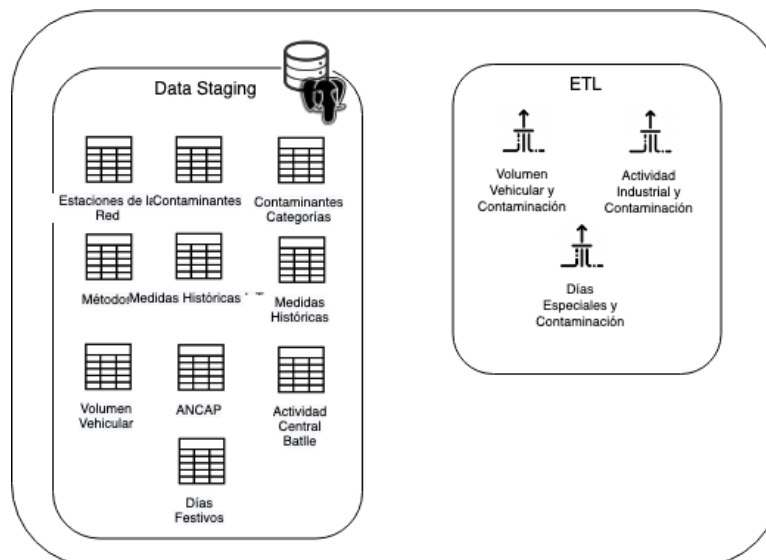
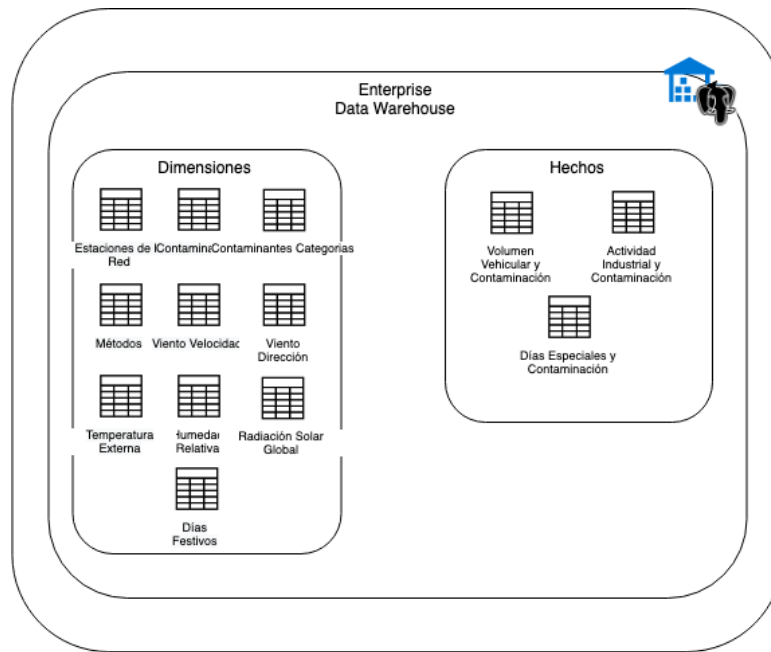


Figura 4.29: Back-end

4.3.3. DW

En la capa de almacenamiento de datos se tiene una base de datos relacional que oficia de repositorio de datos.

En el mismo se pueden diferenciar dos grandes grupos: las tablas que sirven como Dimensiones y las que sirven como Hechos, como se muestra en fig. 4.30.



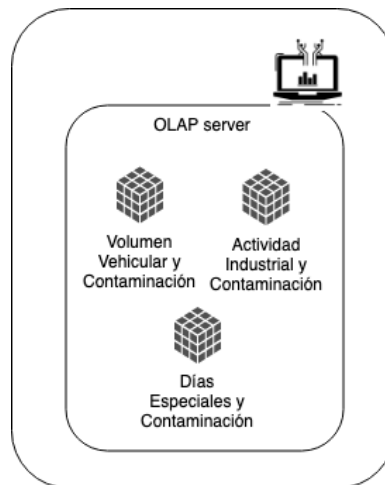
Data Warehouse

Figura 4.30: DW

4.3.4. OLAP

En la capa OLAP se mantienen las reglas de negocio que se aplican a la información almacenada en el Enterprise DW para ser presentada de forma intuitiva por las herramientas de análisis de datos.

En el mismo, ver fig. 4.31, se definen tres cubos, uno por cada requerimiento identificado.



OLAP

Figura 4.31: OLAP

4.3.5. Front-end

Finalmente, en la capa de análisis de datos, se ofrecen al usuario herramientas con los siguientes objetivos:

1. Mecanismos de acceso a la información eficaces.
2. Tener conexión eficaz al DW.

Primero se utiliza Saiku, una herramienta de la suite Pentaho, para que los especialistas puedan realizar las consultas de forma dinámica logrando la mayor flexibilidad posible para analizar los datos.

A su vez, se disponibiliza la capacidad de generar un informe, similar al publicado anualmente, para automatizar la generación del mismo.

Finalmente, se genera un dashboard a medida para poder visualizar de forma rápida y directa los indicadores preestablecidos.

En la siguiente fig. 4.32 se pueden diferenciar las distintas herramientas de análisis.

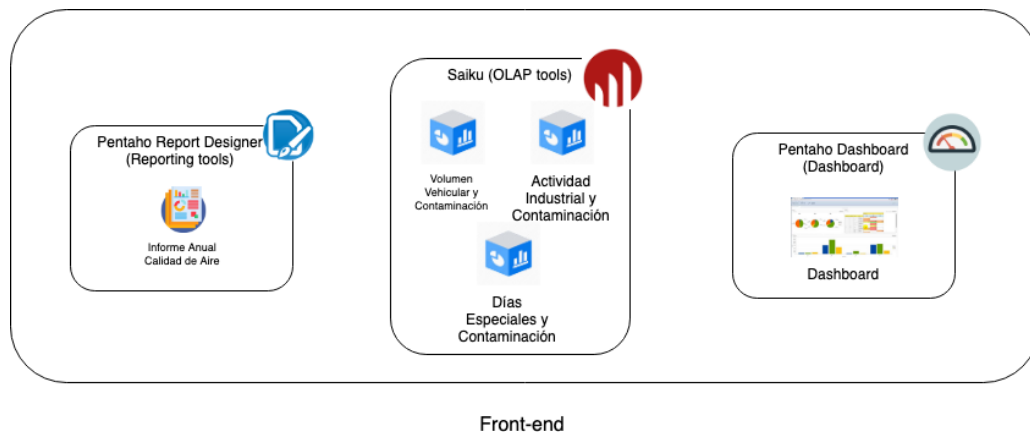


Figura 4.32: Front-end

4.3.6. Metadatos

Horizontalmente, ver fig. 4.33, se puede visualizar la capa de metadatos donde la información sobre los datos almacenados en el sistema de BI es almacenada.

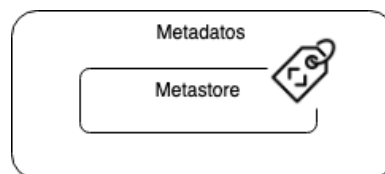


Figura 4.33: Metadatos

4.3.7. Operaciones del sistema

Esta capa incluye información sobre cómo está funcionando el sistema de BI, cuál es el estado de trabajo ETL, cuál es el rendimiento del sistema y el historial de acceso de los usuarios.

Como se puede visualizar en fig. 4.34, en esta capa se pueden visualizar múltiples elementos de operaciones que sirven para el monitoreo:

1. Monitoreo de solución en Docker Host mediante Grafana [31] y Prometheus [41].
2. Status de PBI.
3. Status de PDI (Carte).
4. Aplicación web desarrollada a medida.



Figura 4.34: Operaciones

4.4. Diseño

Dado que el área de Servicio de Calidad de Aire se enfoca especialmente en el monitoreo y evaluación de la calidad de aire, un área de negocio con objetivos específicos, se utiliza la metodología bottom-up.

A continuación se especifican cada una de las etapas de diseño realizadas.

4.4.1. Diseño Conceptual

La etapa de Diseño Conceptual se realiza utilizando un enfoque combinado, haciendo uso paralelamente de los requerimientos identificados y el modelo conceptual de las fuentes de datos generado, ambas salidas generadas en la etapa de Análisis. Finalmente se integran los resultados de ambos enfoques.

Adicionalmente, se incluye la representación gráfica de los distintos componentes del diseño conceptual utilizando el modelo CMDM.

4.4.1.1. Dimensiones y jerarquías

Para modelar la realidad del problema es necesario construir las siguientes dimensiones que sirven como criterios de análisis de los datos como se pueden ver en los distintos grupos de dimensiones en las figuras fig. 4.35, fig. 4.36, fig. 4.37, y fig. 4.38.

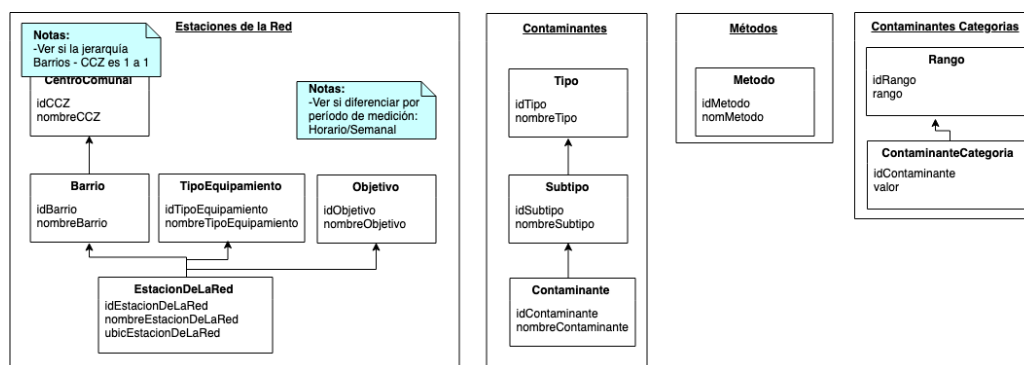


Figura 4.35: Dimensiones de Red de Monitoreo de Calidad de Aire de Montevideo

A continuación se describen las distintas dimensiones en detalle:

Estaciones de la Red

La dimensión Estaciones de la Red representa las estaciones que permiten evaluar la concentración de material particulado en diversas fracciones y de gases que habitualmente se encuentran en ambientes urbanos (dióxido de azufre y dióxido de nitrógeno).

Niveles

EstacionDeLaRed: Compuesto por las estaciones de la red de monitoreo.

Barrio: Compuesto por los barrios de Montevideo.