

Arquitectura de servicio de Power BI

Eleonora Rodríguez

La arquitectura de servicio de Power BI

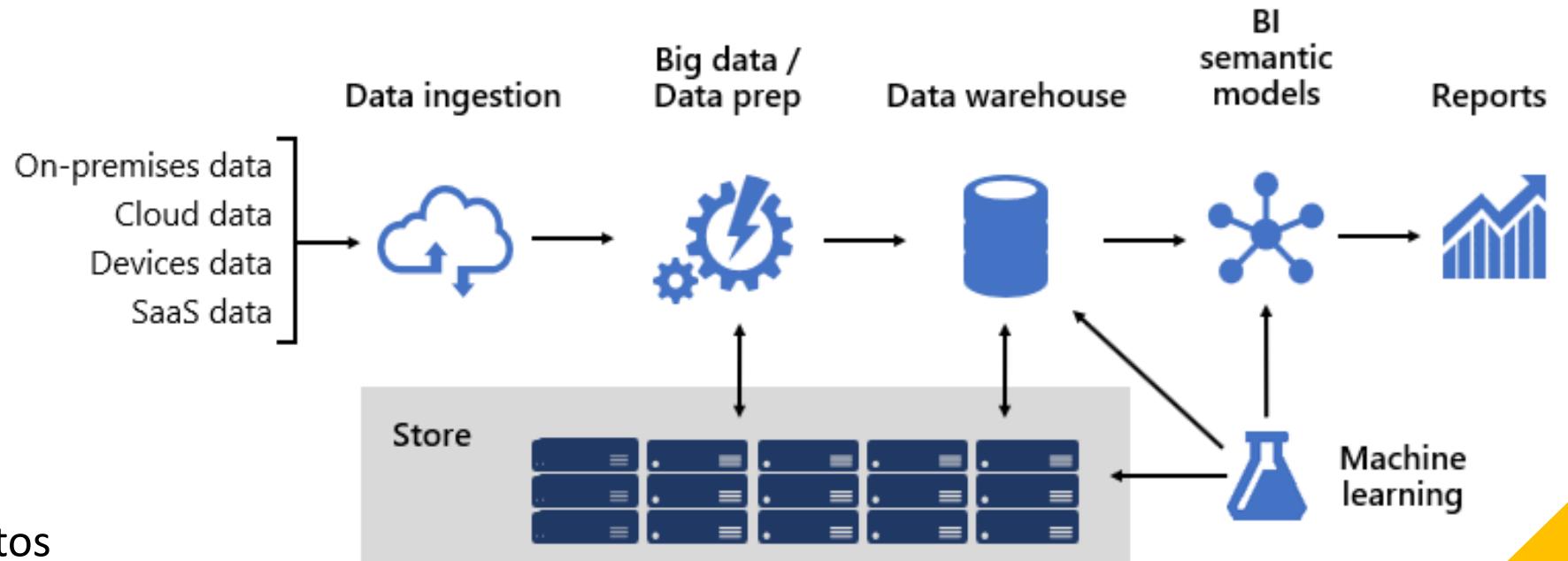
Se basa en dos conjuntos o clústeres.

- El primero es el WFE o front-end web.
- El segundo es el back-end .
- El primero de ellos, el WFE es responsable de la conexión inicial y la autenticación en el servicio de Power BI. El back-end controla todas las interacciones de usuario que se producen después de haber realizado la autenticación.

- Power BI es una herramienta que engloba diferentes servicios y aplicaciones que son capaces de Identificar la información que se quiere obtener, identificar los orígenes de datos, crear un modelo de datos, crear la infraestructura de base de datos, el proceso de carga, implementar el modelo de datos, el sistema de reporting y cuadros de mando.
- Algunas de esas aplicaciones se encuentran en Power BI Desktop, que dan la posibilidad de recopilar datos y generar informes visuales.
- Esta posibilidad se debe a la completa arquitectura de Power BI que al encontrarse ubicado en la nube y gracias a Sharepoint permite abrir el conocimiento al resto de actores involucrados en la toma de decisiones.

Arquitectura de la solución de BI en el centro de excelencia

- El diseño de una plataforma de BI sólida es algo parecido a crear un puente; un puente que conecta los datos de origen transformados y enriquecidos con los consumidores de datos. El diseño de esta estructura compleja requiere una mentalidad de ingeniería, aunque puede ser una de las arquitecturas de TI más creativas y gratificantes que pueda diseñar. En una organización de gran tamaño, una arquitectura de la solución de BI puede constar de lo siguiente:



1. Orígenes de datos
2. Ingesta de datos
3. Preparación de datos o macrodatos
4. Almacenamiento de datos
5. Modelos semánticos de BI
6. Informes

1. Orígenes de datos

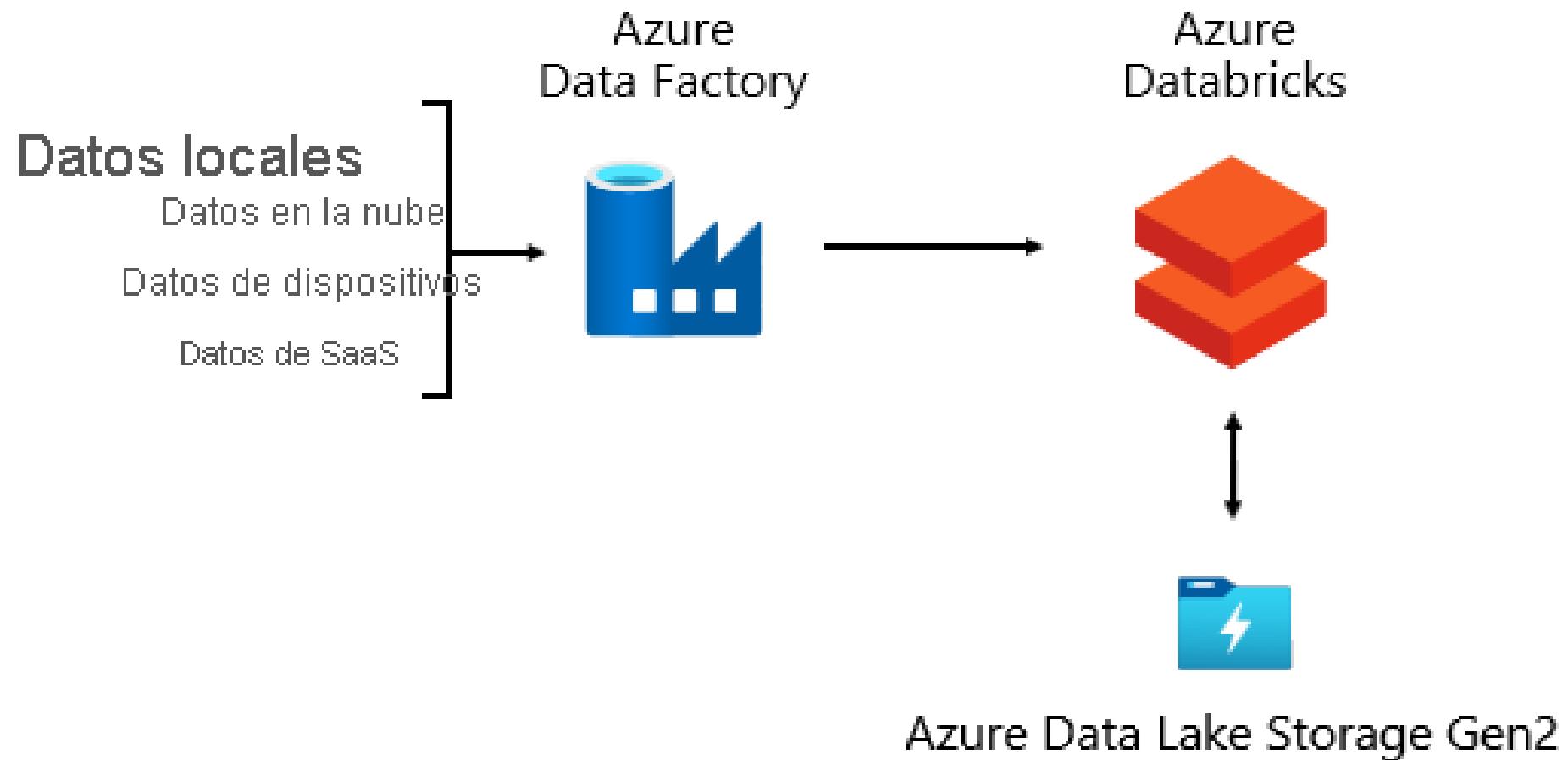
- Un almacenamiento de datos puede consolidar datos de prácticamente cualquier origen de datos. Se basa principalmente en orígenes de datos de aplicación de línea de negocio, que suelen ser bases de datos relacionales que almacenan datos específicos de cada tema para ventas, marketing, finanzas, etc. Estas bases de datos pueden estar hospedadas en la nube o residir en el entorno local.
- Otros orígenes de datos pueden estar basados en archivos, especialmente registros web o datos de IOT (Internet de las Cosas) procedentes de dispositivos. Además, los datos pueden tener su origen en proveedores de software como servicio (SaaS).

- En Microsoft, algunos de sus sistemas internos envían datos operativos directos a ADLS Gen2 con formatos de archivo sin procesar. Además de su lago de datos, otros sistemas de origen incluyen aplicaciones de línea de negocio relacionales, libros de Excel, otros orígenes basados en archivos y repositorios de datos personalizados y de administración de datos maestros (MDM).
- Los repositorios de MDM nos permiten administrar los datos maestros para garantizar versiones de datos autoritativas, estandarizadas y validadas.

2. Ingesta de datos

- La ingestión de datos se refiere a la extracción, carga y transformación de datos (ETL). La diferencia estriba en el lugar donde se produce la transformación. Las transformaciones se aplican para limpiar, ajustar, integrar y estandarizar los datos.
- En última instancia, el objetivo es cargar los datos correctos en el modelo empresarial con la mayor rapidez y eficacia posibles.

Ejemplo:



3. Almacenamiento de datos

- Los datos refinados se almacenan en una base de datos relacional para ofrecer un almacén de datos de alto rendimiento y altamente escalable para los modelos empresariales, con seguridad, gobernanza y capacidad de administración.

Modelos

El modelado de datos es el proceso de analizar y definir todos los diferentes datos que una organización recopila y produce, así como las relaciones entre esos bits de datos.

- Sólo profesionales de marcado perfil técnico en bases de datos y de la herramienta de BI eran capaces de transformar grandes volúmenes de datos en informes rápidos de ejecutar y entender.
- Pero una cosa es crear un conjunto de datos pequeño de uso individual y otra muy distinta es la de crear uno con cierta amplitud, complejidad y volumen que tengan que utilizar otros. La capa semántica, al margen de su aspecto más técnico viene a facilitar es uso compartido con normas bastante simples.

Modelos semánticos de BI

- Representan una capa semántica en los modelos empresariales, es decir, es la que comunica los datos con el usuario y sirve exactamente para asegurarse que los usuarios comprenderán fácilmente los datos que maneja encontrando sin ambigüedades aquellas medidas y campos que necesitan.
- Los desarrolladores de BI y los usuarios empresariales los compilan y mantienen. Los desarrolladores crean modelos semánticos de BI principales, los cuales obtienen datos de los modelos empresariales.
- Los usuarios empresariales pueden crear modelos independientes de menor escala, o bien pueden ampliar los modelos semánticos de BI principales con orígenes externos o de departamento. Los modelos semánticos de BI normalmente se centran en una sola área temática y a menudo son ampliamente compartidos.

- Las capacidades empresariales no solo se habilitan por datos, sino también por modelos semánticos de BI que describen conceptos, relaciones, reglas y estándares. De este modo, representan estructuras intuitivas y fáciles de entender que definen relaciones de datos y encapsulan reglas de negocios como cálculos. También pueden aplicar permisos de datos específicos, que garanticen que las personas adecuadas tengan acceso a los datos correctos. Lo importante es que aceleran el rendimiento de las consultas y proporcionan un análisis interactivo con gran capacidad de respuesta, incluso con varios terabytes de datos.

- En una plataforma de BI en la nube, los desarrolladores de BI pueden implementar modelos semánticos de BI en Azure Analysis Services o capacidades de Power BI Premium. Se recomienda su implementación en Power BI cuando se usa como capa de informes y análisis. Estos productos admiten diferentes modos de almacenamiento, lo que permite que las tablas del modelo de datos almacenen en memoria caché sus datos o utilicen DirectQuery, que es una tecnología que pasa las consultas a través del origen de datos subyacente. DirectQuery es un modo de almacenamiento ideal cuando las tablas del modelo representan grandes volúmenes de datos o es necesario ofrecer resultados casi en tiempo real. Los dos modos de almacenamiento se pueden combinar: Los modelos compuestos combinan tablas que usan distintos modos de almacenamiento en un único modelo.

Diseñando modelos de datos

- Cuando se parte de una tabla de hechos (por ejemplo, la facturación de una empresa) se tiende a pensar en el plano técnico, de eficiencia, aunque por el camino del diseño del modelo de datos es posible que surjan dudas, y la solución debe ser, aunque sea la más compleja técnicamente, la que mayor aceptación se crea que vaya a tener en el usuario final.
- Podemos tener escenarios del tipo: necesitamos la facturación trimestral. La decisión técnica puede ser variada. Podemos tener una medida llamada “Facturación trimestral” o bien obligar al usuario a escoger la medida “Facturación” y el periodo “Trimestre” para obtener el mismo resultado. No hay una respuesta válida general, dos conjuntos de datos pueden tener una decisión válida distinta. Dependerá de varios factores: la cultura de la empresa, el tipo de usuario final o la extensión del modelo de datos.

Reuniendo la información necesaria

- Un datawarehouse es aquel almacén de datos que reúne toda la información de la empresa en un único lugar y la expone a herramientas como Power BI.
- Pero muchos proyectos no parten de un datawarehouse, simplemente recogen datos directamente de las fuentes transaccionales (CRM, ERP, programas contables, de Recursos Humanos, etc) y con ellos se crean conjuntos de datos que contienen una visión de la organización con datos departamentales (contables, comerciales, etc.). Este es uno de los pasos más importantes a estudiar pues hay que asegurarse que el usuario final podrá disponer de toda la información necesaria.

Reuniendo la información necesaria

- Si se está considerando un conjunto de informes destinado al área comercial, debemos pensar en su información importante y entender de donde la debemos recuperar. En ocasiones esto implica importar datos del CRM para disponer de la actividad sobre clientes, del ERP para contabilizar ventas y puede que de uno o varios archivos de Excel para obtener los objetivos de venta y otras informaciones.

Estandarizando el nombre de las cosas

- Lo habitual en el modelo transaccional es encontrar campos del tipo ID, COD_Prod y otras abreviaciones. En el modelo que se construye en Power BI se debe seguir un patrón en la denominación de las tablas y campos, deben ser lo más descriptivas posibles, de esta manera si es necesario que el usuario maneje el código de producto, se debe denominar de forma natural “Código de producto”.
- También lo más lógico es que si se debe mostrar otros códigos, estos se denominen “Código de ...”. En ocasiones existen dudas y se esté tentado a escribir “Producto código” para favorecer la ordenación de los objetos, esto no suele funcionar. Mejor hacerlo con lenguaje natural aunque sea necesario a veces eliminar conjunciones.

Agrupando medidas

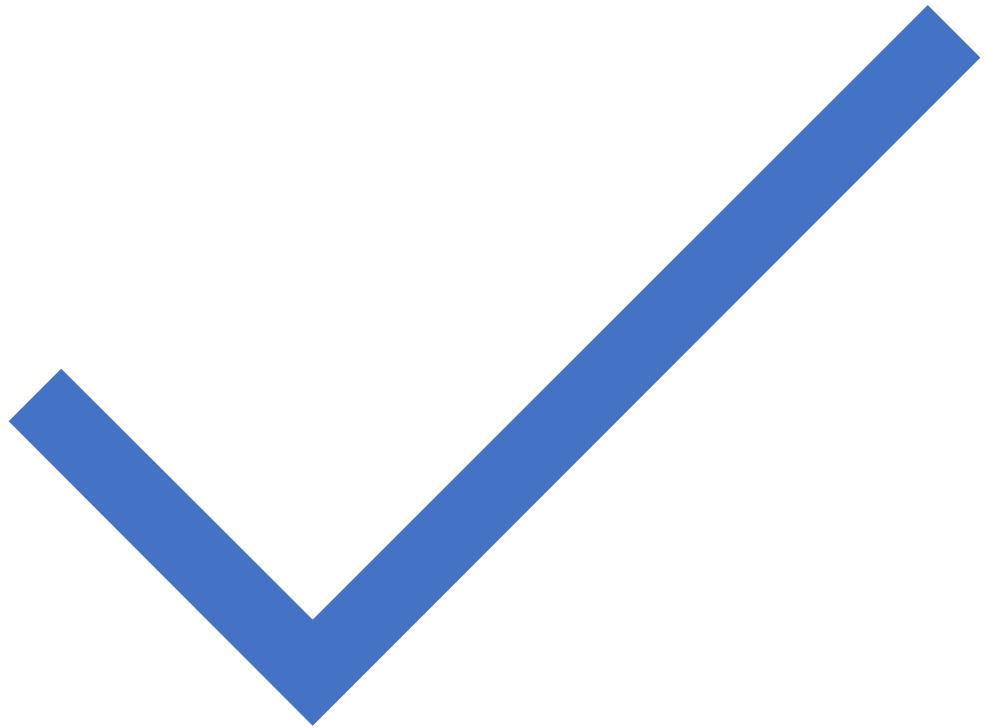
- Dependiendo de las circunstancias es posible que sea mejor para el usuario final disponer de todas las medidas en una misma carpeta.

Informes (consumo de datos)

- En la capa de informes, los servicios de negocio consumen datos empresariales procedentes del almacenamiento de datos. También tienen acceso a los datos directamente en el lago de datos para realizar tareas de análisis ad hoc o de ciencia de datos.
- Los permisos específicos se aplican en todas las capas: en el lago de datos, los modelos empresariales y los modelos semánticos de BI. Los permisos garantizan que los consumidores de datos solo pueden ver los datos para los que tienen derechos de acceso.
- En Microsoft, se utilizan informes y paneles de Power BI e informes paginados de Power BI. Algunos informes y análisis ad hoc se realizan en Excel, especialmente en el caso de los informes financieros.

- Se publican diccionarios de datos, que proporcionan información de referencia a los modelos de datos. Se han puesto a disposición de los usuarios a fin de que puedan descubrir información sobre la plataforma de BI. Los diccionarios documentan los diseños de modelos y proporcionan descripciones sobre entidades, formatos, estructura, linaje de datos, relaciones y cálculos.
- Normalmente, los patrones de consumo de datos difieren en función del rol:
- Los analistas de datos se conectan directamente a los modelos semánticos de BI principales. Cuando los modelos semánticos de BI principales contienen todos los datos y la lógica que necesitan, usan conexiones dinámicas para crear informes y paneles de Power BI. Cuando tienen que ampliar los modelos con datos departamentales, crean modelos compuestos de Power BI. Si se necesitan informes de tipo de hoja de cálculo, usan Excel para generar informes basados en modelos semánticos de BI principales o modelos semánticos de BI de departamento.

- Los desarrolladores de BI y los autores de informes operativos se relacionan directamente con los modelos empresariales. Usan Power BI Desktop para crear informes de análisis de conexión dinámica. También pueden crear informes de BI de tipo operativo como informes paginados de Power BI si escriben consultas SQL nativas para acceder a los datos de los modelos empresariales de Azure Synapse Analytics mediante T-SQL, o los modelos semánticos de Power BI mediante DAX o MDX.
- Los científicos de datos se conectan directamente a los datos del lago de datos. Usan cuadernos de Azure Databricks y Python para desarrollar modelos de ML, que a menudo son experimentales y requieren aptitudes especializadas para su uso en producción.



**Gracias por
su atención**

Referencias Bibliográficas

- <https://www.makesoft.es/arquitectura-de-power-bi/#:~:text=La%20arquitectura%20de%20servicio%20de,segundo%20es%20el%20back%2Dend%20>
- <https://www.arimetrics.com/glosario-digital/power-bi>
- <https://learn.microsoft.com/es-es/power-bi/guidance/center-of-excellence-business-intelligence-solution-architecture>
- https://www.google.com/search?q=que+son+los+modelos+sem%C3%A1nticos+en+Power+BI%3F&rlz=1C1ALOY_esSV1047SV1047&oq=que+son+los+modelos+sem%C3%A1nticos+en+Power+BI%3F&aqs=chrome..69i57j33i160l3.9920j1j7&sourceid=chrome&ie=UTF-8
- <http://blog.blauservices.com/?p=2008#:~:text=Una%20capa%20sem%C3%A1ntica%20es%20la,medidas%20y%20campos%20que%20necesitan>