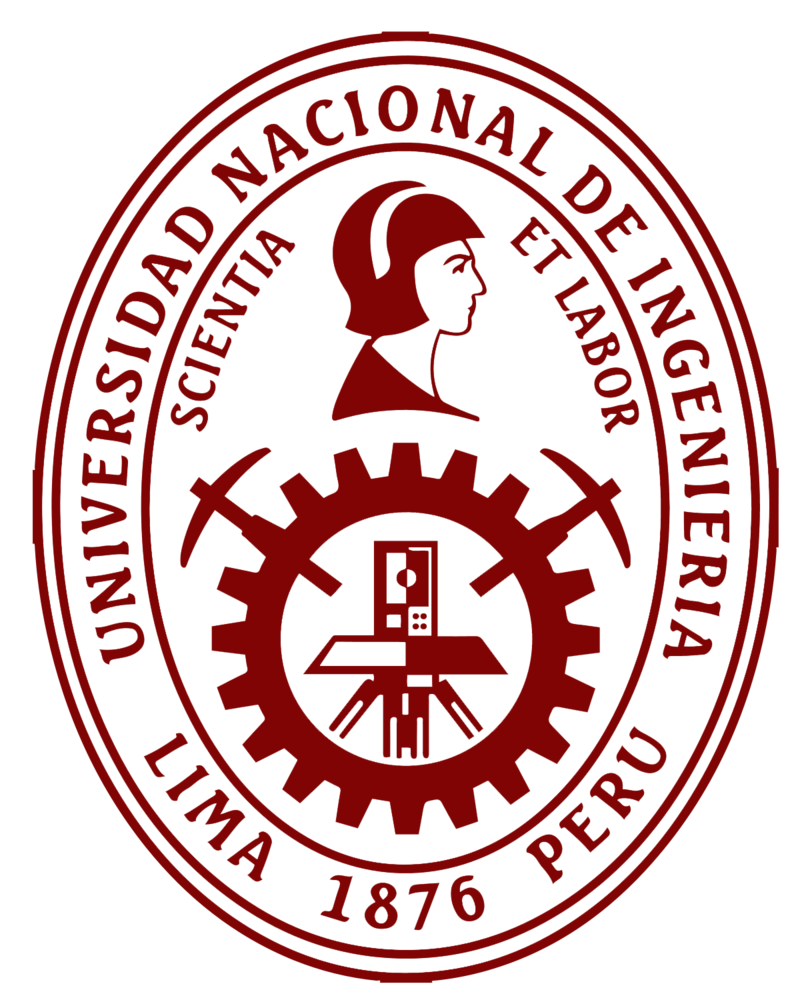
**Universidad Nacional de Ingeniería**

**Facultad de Ingeniería Industrial y de Sistemas**

**Escuela de Ingeniería de Sistemas**

**SI 807 Sistema de Inteligencia de Negocios “U”**

**Practica calificada 1**

Caso de Estudio: “Herramientas de análisis para la SOT en el área de instalación y mantenimiento”

**Integrantes:**

* Callupe Pardo Yoselyn Patricia 20190267H
* Garro Oré, Willian Jesús 20172115E
* Nuñez-Poma-Robert Gianpierro Jesus 20202084E

Lima – Perú

**2025**

1. **ENTORNO DE LA EMPRESA SELECCIONADA**
   1. **Generalidades de la Empresa**
      1. **Empresa**

Claro Perú, que opera legalmente bajo la razón social América Móvil Perú S.A.C., es la filial peruana de la corporación mexicana América Móvil, una de las empresas de telecomunicaciones más grandes del mundo. La compañía ofrece una amplia gama de productos y servicios para clientes individuales y empresariales, incluyendo telefonía móvil, internet fijo y móvil, y televisión por cable.

* + 1. **Giro de la empresa**

El giro de negocio de Claro Perú se define como la provisión integral de servicios de telecomunicaciones, puesto que este modelo de negocio integra telefonía móvil, fija, internet de banda ancha y televisión

* + 1. **Ubicación**

Las oficinas principales se encuentran en la Av. Nicolás Arriola 480, La Victoria, Lima. Además, la empresa cuenta con una extensa red de Centros de Atención al Cliente (CACs) distribuidos en diversos distritos de Lima y a nivel nacional.

* + 1. **Misión**

"Brindamos la mejor experiencia de servicio a través de las más avanzadas soluciones de comunicación, tecnología de información y contenido digital para acelerar el desarrollo de los países donde operamos y promover la igualdad de oportunidades entre la gente."

* + 1. **Visión**

"Consolidarnos como un agente de cambio al proporcionar servicios de conectividad y alta tecnología; preservando nuestro liderazgo en la industria de las telecomunicaciones y reafirmando nuestro compromiso con las personas para hacer un mundo más próspero para todos."

* + 1. **Productos y clientes**

Claro ofrece una amplia gama de productos y servicios a sus clientes, que incluyen tanto a consumidores individuales como a empresas:

* **Servicios móviles:** planes prepago, postpago y equipos móviles.
* **Servicios para el hogar:** planes de internet (incluyendo fibra óptica), televisión por cable y telefonía fija.
* **Servicios empresariales:** soluciones a la medida, como planes móviles para negocios, servicios de internet, soluciones de gestión de ventas y finanzas, y más.
* **Clientes:** El mercado de clientes de Claro es muy amplio, abarcando desde usuarios individuales que buscan servicios de telefonía e internet, hasta grandes y pequeñas empresas que requieren soluciones de conectividad y herramientas tecnológicas.
  + 1. **Organigrama**

Ilustración 01:Organigrama

****

*Fuente:Tesis “Gestión de Atención de Incidencias para el Soporte Comercial en la empresa Claro Perú”*

* + 1. **Cadena de valor**

Ilustración 02:Cadena de valor

****

*Fuente:Tesis “Gestión de Atención de Incidencias para el Soporte Comercial en la empresa Claro Perú”*

* + 1. **Procesos**

Los procesos de negocio de Claro están orientados a la gestión de ventas, distribución, y servicios. Estos incluyen:

* **Controlar inventario de activos en almacenes:** Proceso responsable de controlar el ingreso y salida de inventario en los almacenes de la empresa, con el fin de mantener stock disponible para garantizar las ventas.
* **Gestionar almacenes e infraestructura de red:** Proceso encargado de adquirir nuevos almacenes o locales que cumplan esta función para abastecimiento de exceso de inventario, así como, identificar lugares estratégicos para la colocación o instalación de antenas de red.
* **Gestionar el abastecimiento de infraestructura tecnológica:** Proceso que es responsable de identificar y proponer innovación tecnológica para las distintas plataformas de la empresa.
* **Desarrollar proyectos de inversión:** Proceso que se encarga de la definición e implementación de proyectos de inversión que generen valor a la empresa y mayor retorno a mediano y largo plazo.
* **Desarrollar proyectos de innovación tecnológica:** Proceso que se encarga de la definición e implementación de proyectos de innovación tecnológica que generen valor a la empresa y lo coloquen a la vanguardia del mercado.
* **Gestionar canales de distribución:** Proceso responsable de la correcta y puntual entrega de dispositivos telefónicos, como equipos celulares, módems, accesorios telefónicos, entre otros, a los diversos canales de distribución, con el fin de cumplir con la demanda del mercado.
* **Desarrollar publicidad de productos comerciales:** Proceso encargado del desarrollo de la estrategia de marketing para la publicidad de los diversos productos comerciales que entrega la empresa, con el fin de que llegue a la mayor cantidad de clientes potenciales.
* **Gestionar construcción continua de la marca:** Proceso responsable de desarrollar la estrategia para la identificación y significado de la marca, con el fin de lograr una relación de lealtad activa e intensa entre los clientes y nuestra marca.
* **Gestionar acercamiento al cliente objetivo:** Proceso que se encarga de la definición e implementación de servicios y campañas bajo las necesidades del mercado con el fin de llegar al cliente objetivo.
* **Gestionar servicios adicionales post venta:** Proceso responsable de la implementación de servicios adicionales complementarios a las campañas y planes existentes que se ofrecen a los clientes.
* **Gestionar mantenimientos de plataformas y redes:** Proceso responsable de brindar mantenimiento oportuno a las plataformas de negocio y redes, con el fin de mantener la continuidad de servicios.
  1. **Identificación de Problemas del Negocio**

Los problemas del negocio identificados son los siguientes:

* **Alta tasa de cancelaciones de SOT por demoras en la instalación:** Un porcentaje significativo de órdenes de trabajo (SOT) son canceladas por los clientes debido a los largos tiempos de espera, lo que afecta directamente los ingresos y la satisfacción del cliente.
* **Falta de visibilidad sobre los motivos de rechazo de las SOT:** No se tiene un registro claro y centralizado de las razones más comunes por las que una SOT es rechazada o no se puede completar en el primer intento.
* **Altos niveles de rechazo en las SOT:** Se presentan rechazos recurrentes en las órdenes de trabajo (SOT) debido a factores como tipo de servicio, ubicación o condiciones en campo. Esto genera reprocesos, incrementa los tiempos de atención y afecta la eficiencia operativa.
* **Tiempos de espera prolongados entre la generación de la SOT y la instalación:** Existe un "cuello de botella" en el proceso, lo que alarga el ciclo de vida de la SOT y aumenta la insatisfacción del cliente.
* **Ausencia de métricas de conversión de SOT:** No se cuenta con indicadores que midan el porcentaje de solicitudes que se convierten en SOT, lo que dificulta evaluar la eficiencia del proceso de gestión y detectar oportunidades de mejora en la asignación o seguimiento de órdenes.
  1. **Necesidades de información y decisiones críticas.**

| **Nivel** | **Tipo de decisión** |  | **Necesidad de información** |
| --- | --- | --- | --- |
| Estratégico | Evaluar la viabilidad de nuevos servicios de instalación a domicilio y reestructurar los procesos operativos. |  | Analizar el costo de las demoras y cancelaciones por tipo de servicio y región. Determinar el impacto financiero de las fallas en el proceso. |
| Táctico | Mejorar la programación de técnicos y la asignación de SOT. Implementar protocolos de comunicación con el cliente en tiempo real. |  | Conocer la dilación promedio por cada tipo de atención y por cada equipo de trabajo. Identificar los rechazos más comunes y su ubicación geográfica. |
| Operativo | Priorizar las SOT críticas. Reasignar técnicos en caso de retrasos inesperados. |  | Consultar en tiempo real el estado actual de las SOT y la disponibilidad de los técnicos. Acceder a un listado de los motivos de cancelación o rechazo más frecuentes. |

* 1. **KPI’s Iniciales**

| **Nombre del KPI** | **Descripción** | **Fórmula** | **Unidad de medida** | **Frecuencia** | **Nivel de decisión** |
| --- | --- | --- | --- | --- | --- |
| Tasa de conversión de SOT | Mide el porcentaje de órdenes de trabajo completadas exitosamente. | (SOT instaladas / SOT totales) \* 100 | Porcentaje (%) | Diaria | Operativo / Táctico |
| Tiempo Promedio de Instalación | Mide el tiempo promedio que toma instalar una SOT, desde su creación hasta su cierre. | Suma del tiempo de instalación de todas las SOT / Número de SOT instaladas | Horas / Días | Diaria / Semanal | Táctico |
| Tasa de Rechazo de SOT | Mide el porcentaje de órdenes de trabajo que son rechazadas por el cliente o no pueden completarse. | (SOT rechazadas / SOT totales) \* 100 | Porcentaje (%) | Semanal | Táctico |
| Tiempo Promedio de Dilación por creación | Mide el tiempo adicional que toma una SOT en completarse por un evento inesperado (ej. no había nadie en casa). | (Tiempo real - Tiempo planificado) / Número de SOT con dilación | Horas | Diaria | Operativo |
| Tiempo Promedio de Dilación por factibilidad | Mide el tiempo adicional que toma una SOT en completarse por un evento inesperado (ej. no había nadie en casa). | (Tiempo real - Tiempo planificado) / Número de SOT con dilación | Horas | Diaria | Operativo |
| Tasa de sot rechazados por Claro | Mide el porcentaje de órdenes de trabajo que son rechazadas por Claro | (SOT rechazados por Claro / Sot rechazadas) \* 100 | Porcentaje (%) | Semanal | Táctico |
| Tasa de sot rechazados por Cliente | Mide el porcentaje de órdenes de trabajo que son rechazadas por Cliente | (SOT rechazados por Cliente / Sot rechazadas) \* 100 | Porcentaje (%) | Semanal | Táctico |
| Ratio de Conversión de la SOT respecto al mes anterior | Cálculo de la tasa de conversión del mes presente con el mes anterior | (Cantidad SOT mes presente) / (Cantidad mes anterior) | Porcentaje (%) | Mensual | Estratégico |
| Tasa de SOT Duplicadas/Fraude/ | Cálculo de la tasa respecto a las SOT totales | SOT duplicadas/fraude / SOT totales | Porcentaje (%) | Mensual | Estratégico |
| Tasa de SOT pendiente | Cálculo de la tasa respecto a las SOT totales | SOT pendientes / Sot Totales | Porcentaje (%) | Mensual | Estratégico |
| Efectividad promedio de la contrata | Cálculo de la SOT convertida por la contrata | Sot Atendida por contrata x/sot generadas asignado | Porcentaje (%) | Mensual | Estratégico |

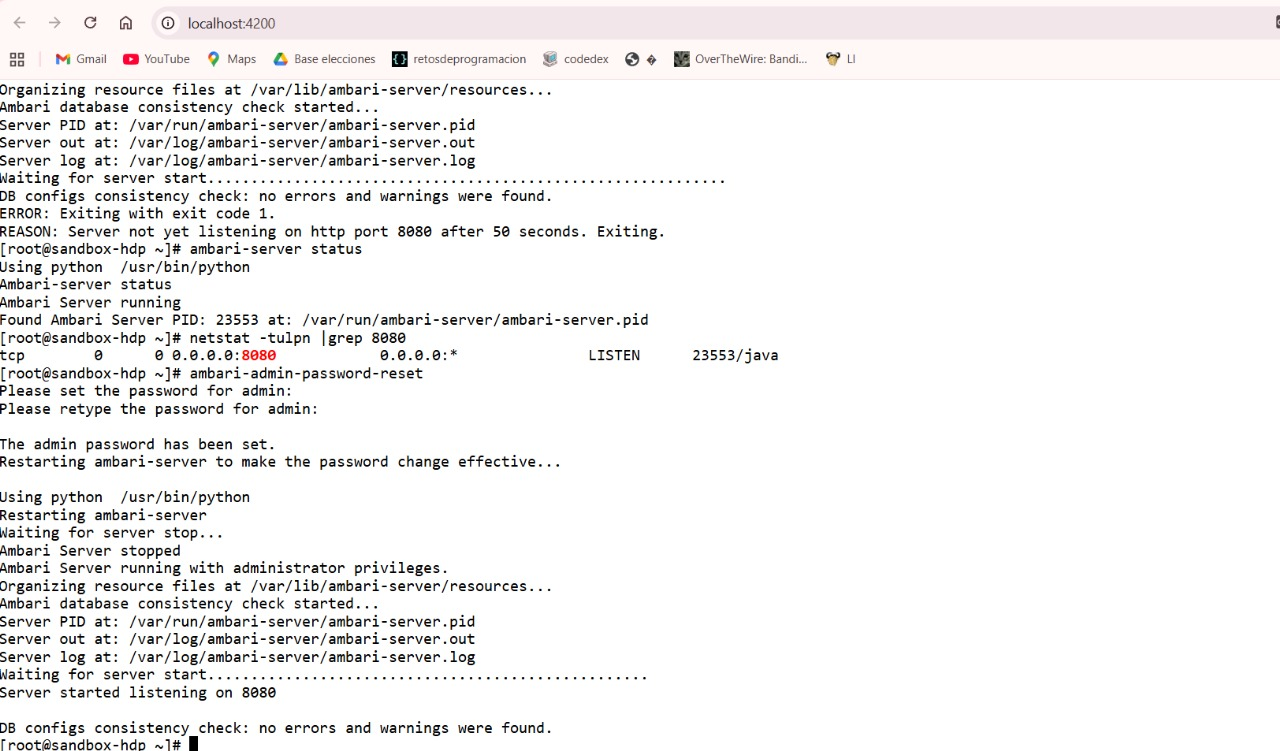
1. **EVIDENCIA TÉCNICA**
   1. **Acceder a Ambari**

Antes de acceder a Ambari, se cambiará la contraseña por defecto para mejorar la seguridad del entorno. Para hacer esto, primero debemos acceder a la terminal de la máquina virtual utilizando SSH.

Se accede a la terminal SSH, lo que te permitirá cambiar la contraseña de acceso al sistema. Se ejecuta el siguiente comando para cambiar la contraseña de Ambari:

ambari-admin-password-reset

Después de ejecutar el comando, el sistema te pide una nueva contraseña para el usuario administrador de Ambari, luego de esto se reinicia.

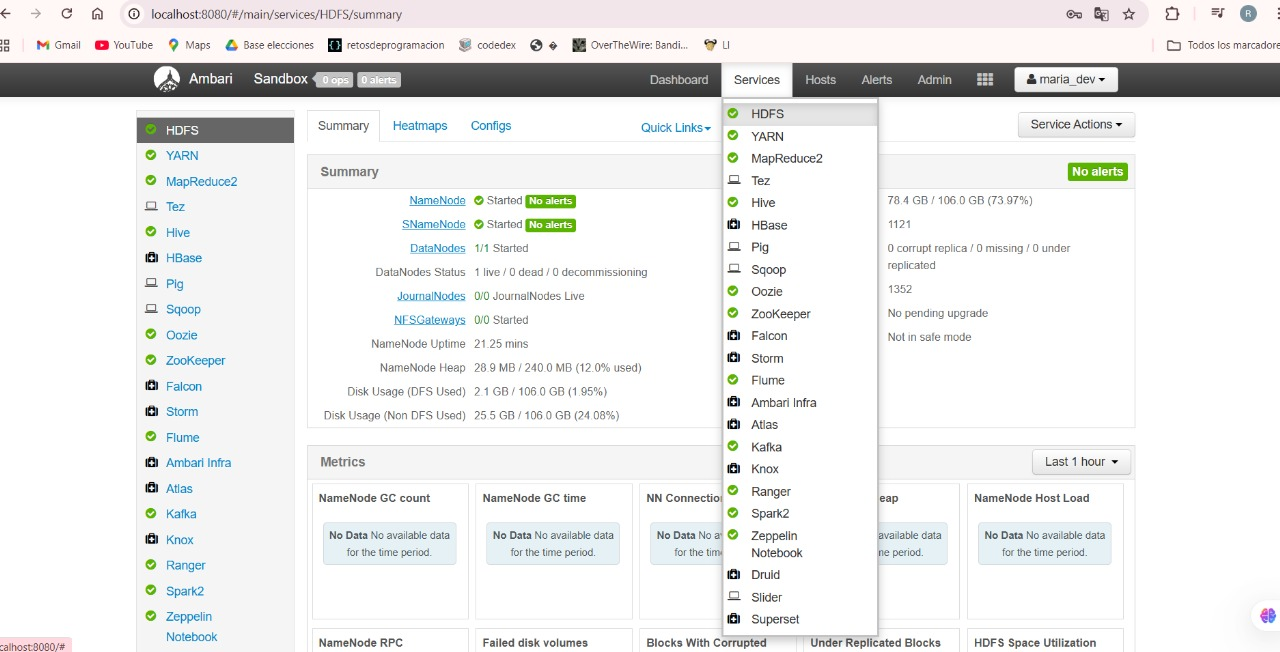
****

Visualización de la interfaz que muestra los diferentes servicios del clúster. En la imagen se muestra que los servicios ya se encuentran iniciados y se puede acceder correctamente a ellos.

* 1. **Exploración Ambari**

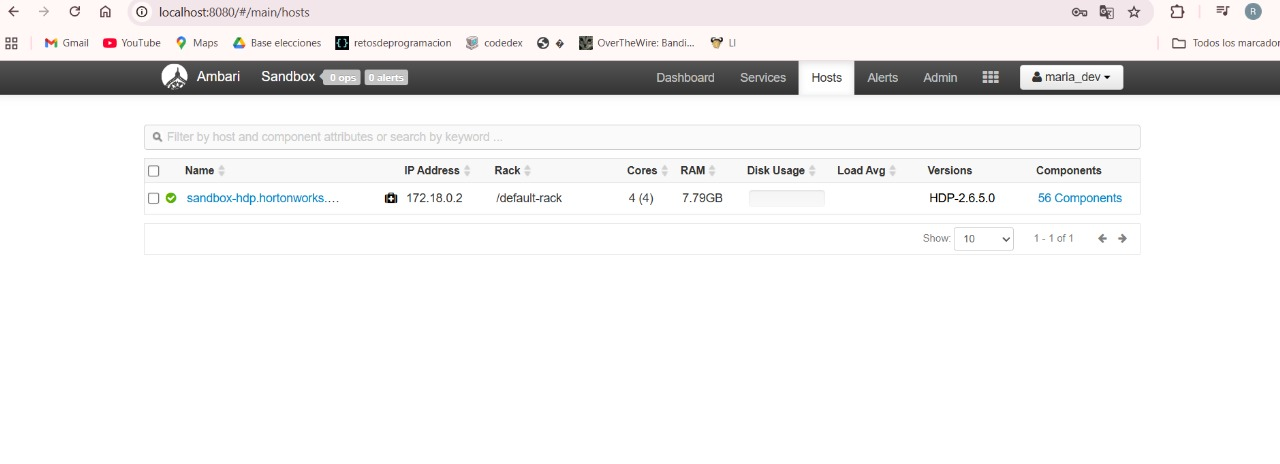
Visualización de la interfaz que muestra los diferentes servicios del clúster. En la imagen se muestra que los servicios ya se encuentran iniciados y se puede acceder correctamente a ellos.

**Servicios en la plataforma HDP:**

****

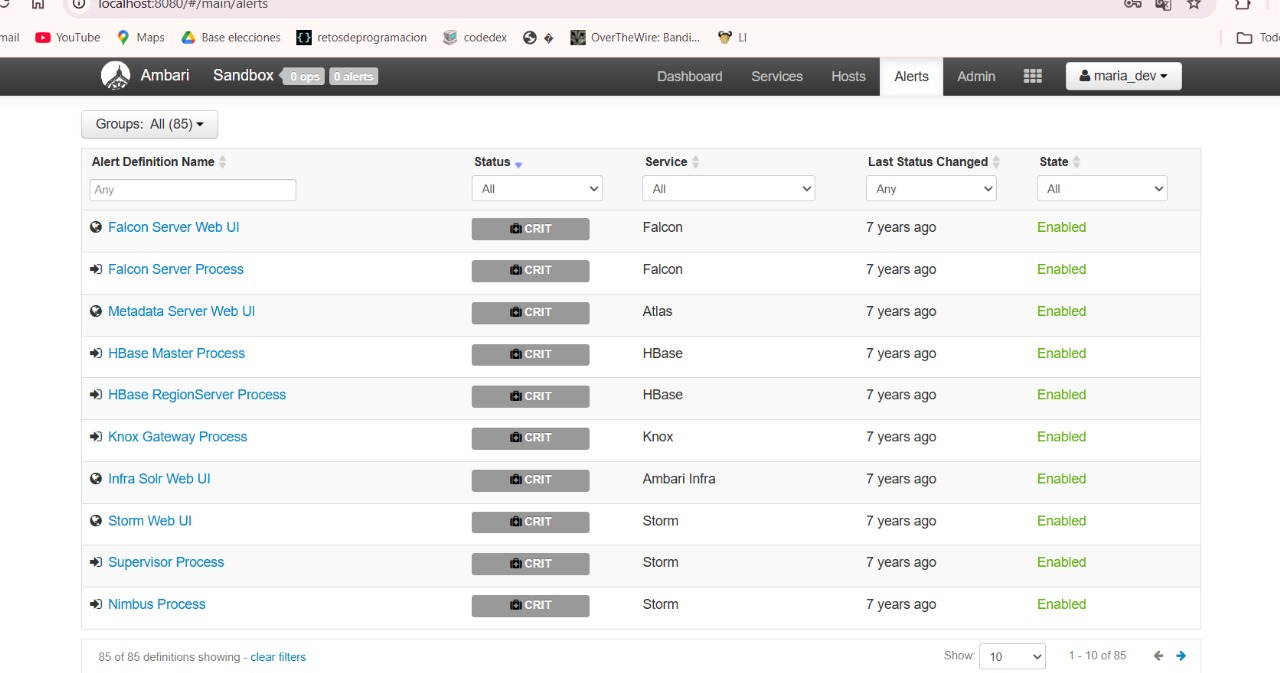
**Hosts:**

Esta parte muestra la lista de todos los nodos (hosts) en el clúster. Aquí podemos ver información detallada sobre cada nodo, como su estado, recursos disponibles y el rendimiento. También puedes gestionar los nodos desde aquí, como iniciar, detener o reiniciar servicios en un nodo específico.

****

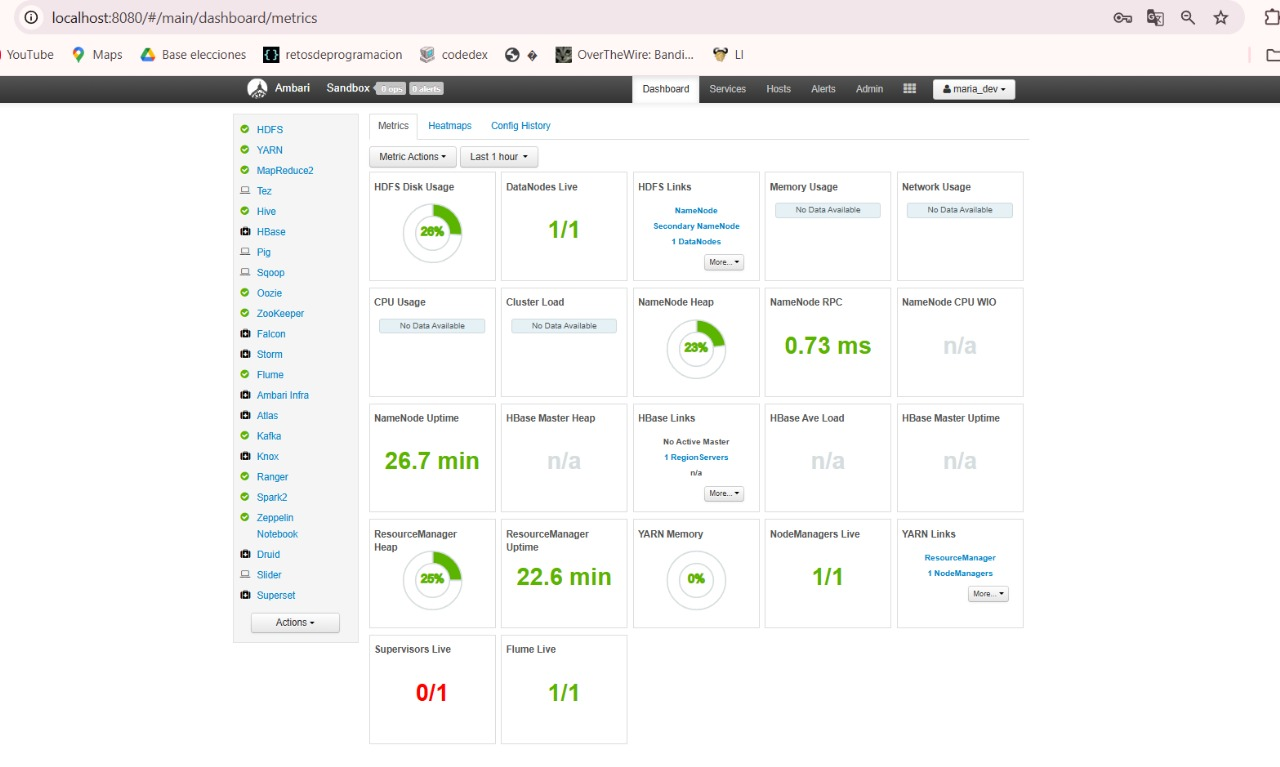
**Alertas:**

Esta sección muestra las alertas generadas por los servicios en el clúster. Aquí podemos ver si hay problemas de rendimiento o errores que deben ser atendidos. Las alertas pueden clasificarse en diferentes niveles (críticos, advertencias, etc.) y se pueden configurar notificaciones para eventos específicos.

****

**Métricas:**

En esta sección, se puede ver un resumen general del estado de los servicios y el clúster. Aquí se muestran métricas clave, como la utilización de recursos y el estado general de los servicios. Esto incluye gráficos que muestran la utilización de recursos, el rendimiento de las aplicaciones, y otra información crítica para el monitoreo del clúster.

****

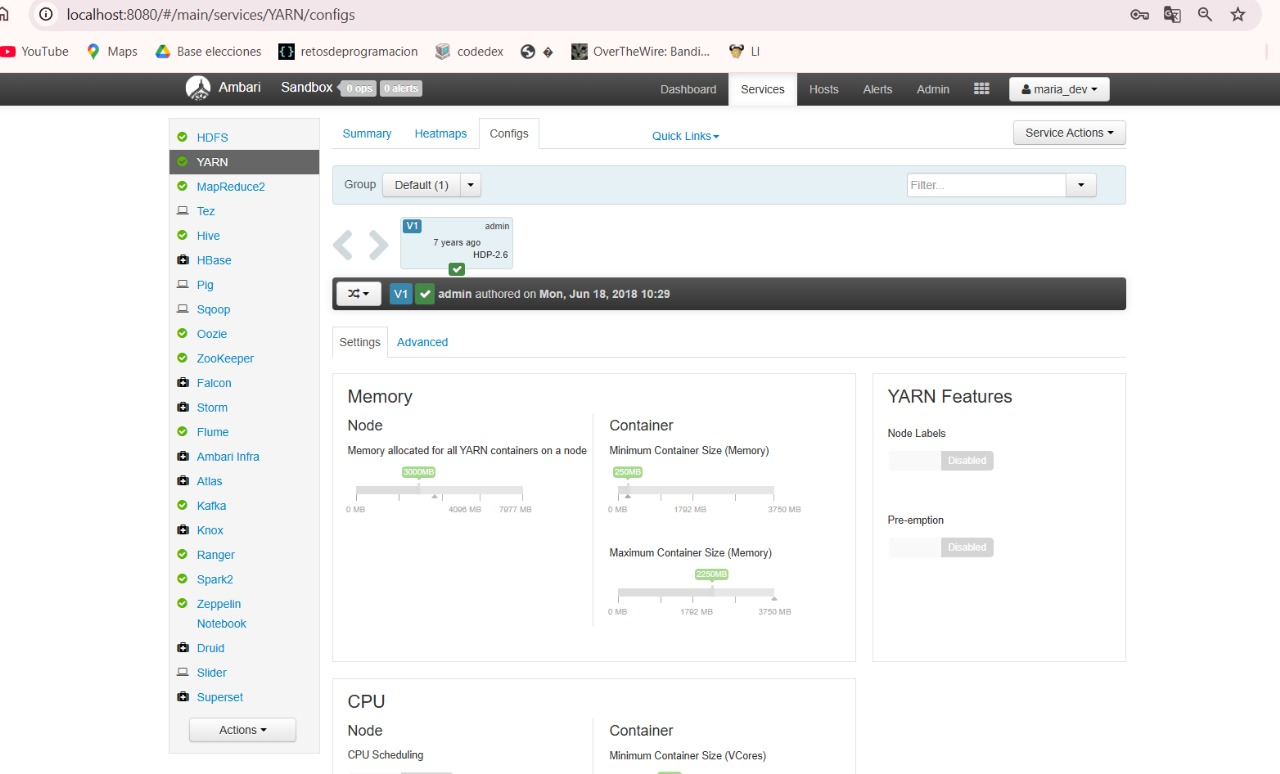
* 1. **Gestionar Cluster en HDP**

Cuando se instaló esta versión de Hadoop en un entorno sandbox (un entorno de pruebas), por defecto solo tenemos un clúster ya configurado por defecto. Esto es suficiente para ambientes de prueba y aprendizaje. En caso estemos ante un entorno de producción real, se debe optar por múltiples clústeres dependiendo de la arquitectura y necesidades de la organización.

Como nuestro caso es el primero, la configuración deberia desarrollarse de la siguiente manera:

En Dashboard visualiza YARN entre sus servicios:

Para visualizar la configuración del clúster, selecciona la pestaña "Configurar" en la parte superior derecha dentro de la página de YARN.

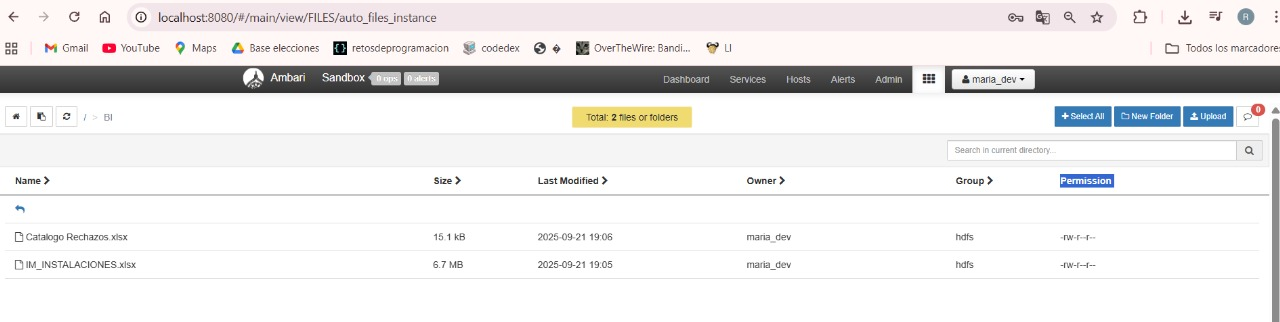
****

El entorno predeterminado de un sandbox está diseñado para manejar pequeñas cargas de trabajo, que son suficientes para tareas de aprendizaje y desarrollo.

La configuración por defecto tiene límites que permiten que el sandbox funcione correctamente con recursos limitados, y cambiar estos parámetros puede desbalancear el sistema y hacer que el sandbox deje de funcionar adecuadamente.

* 1. **Explorando HDFS**

HDFS (Hadoop Distributed File System) es el sistema de archivos distribuido de Hadoop, diseñado para almacenar grandes volúmenes de datos de forma fiable y eficiente en clústeres de nodos. Lo que lo diferencia de otros sistemas de almacenamiento tradicionales es su capacidad para dividir archivos grandes en bloques, los cuales se distribuyen a través de múltiples nodos en el clúster. Estos bloques se replican para garantizar la alta disponibilidad y la tolerancia a fallos, lo que permite a HDFS seguir funcionando incluso si uno o más nodos fallan.

****

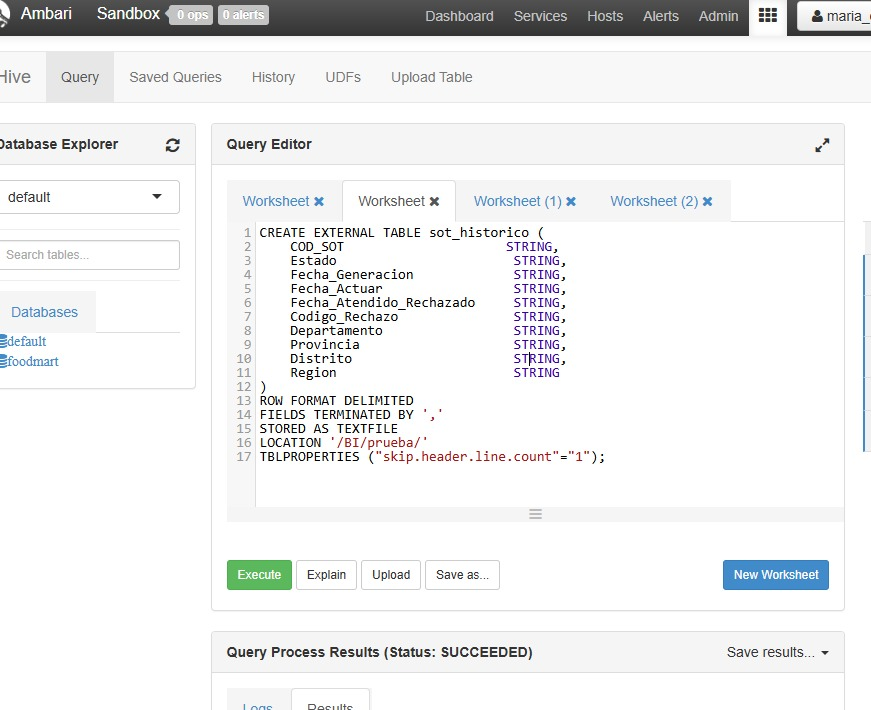
* 1. **Apache Hive**

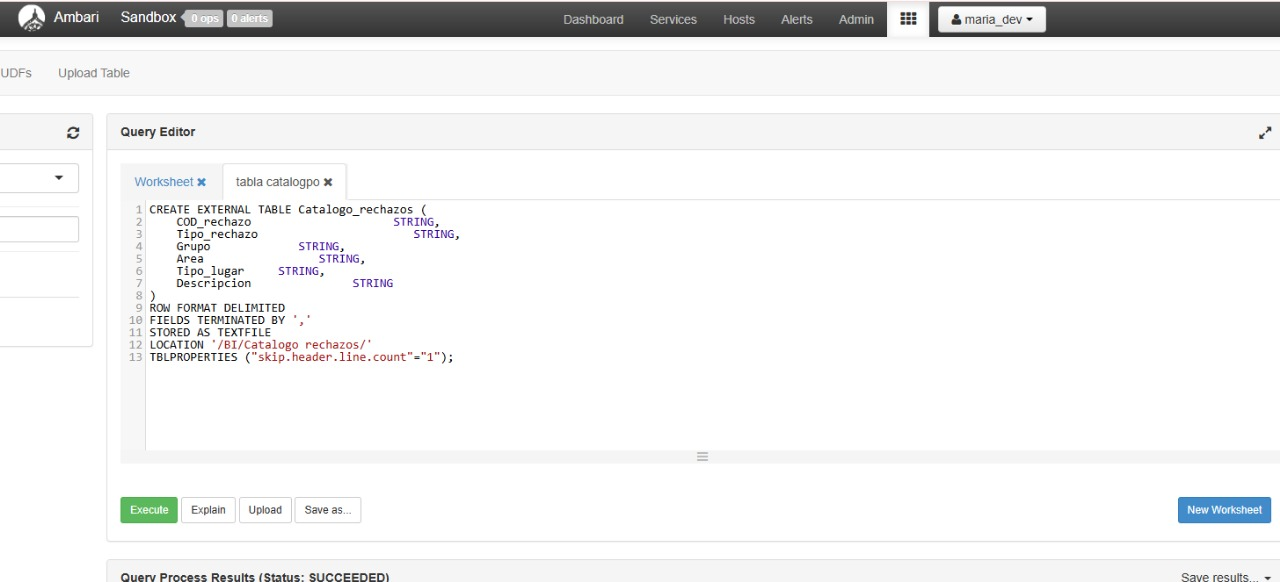
Apache Hive es una herramienta de data warehousing diseñada para trabajar con grandes volúmenes de datos almacenados en sistemas distribuidos como HDFS (Hadoop Distributed File System). Proporciona una interfaz similar a SQL llamada HiveQL para realizar consultas y análisis sobre estos datos. En lugar de escribir código complejo en MapReduce, Hive permite a los usuarios ejecutar consultas SQL de manera sencilla. Entre las funciones basicas de Hive tenemos:

* Consultas SQL: Hive permite realizar consultas sobre grandes conjuntos de datos almacenados en Hadoop, similar a cómo se interactúa con bases de datos relacionales.
* Definición de tablas: Puedes crear tablas para organizar los datos en HDFS y hacer más fácil su manejo.
* ETL: Se puede usar Hive para transformar, limpiar y cargar datos.
* Integración: Hive se puede integrar con otras herramientas de análisis y procesamiento, como Spark y Zeppelin, para análisis avanzados.
  1. **Conexión Hive con HDFS:**

Cuando se crea una tabla en Hive, hay dos tipos de tablas que puedes crear: tablas internas y tablas externas. Las primeras ocurren cuando generamos consultas dentro de Hive sin depender de otros archivos ni rutas y se almacenan en el mismo Hive. Mientras que las tablas externas son las que nos permiten hacer conectar Hive con HDFS.

Una tabla externa en Hive se utiliza cuando se desea que los datos permanezcan fuera del control directo de Hive. Si eliminas una tabla externa, los datos subyacentes en HDFS no se eliminan. Por lo tanto, creamos una tabla externa:





Usaremos un SELECT para que nos muestre los datos. Una vez que se ha creado la tabla en Hive, nos conectamos a ella desde Zeppelin usando PySpark y ejecutar consultas sobre los datos.

* 1. **Apache Zepellin:**

Apache Zeppelin es una herramienta de análisis de datos que permite crear y compartir documentos interactivos llamados "notebooks". Zeppelin proporciona una interfaz que facilita el análisis de datos, la creación de gráficos y la integración con varios motores de procesamiento de datos, como Apache Spark, Hive entre otros.

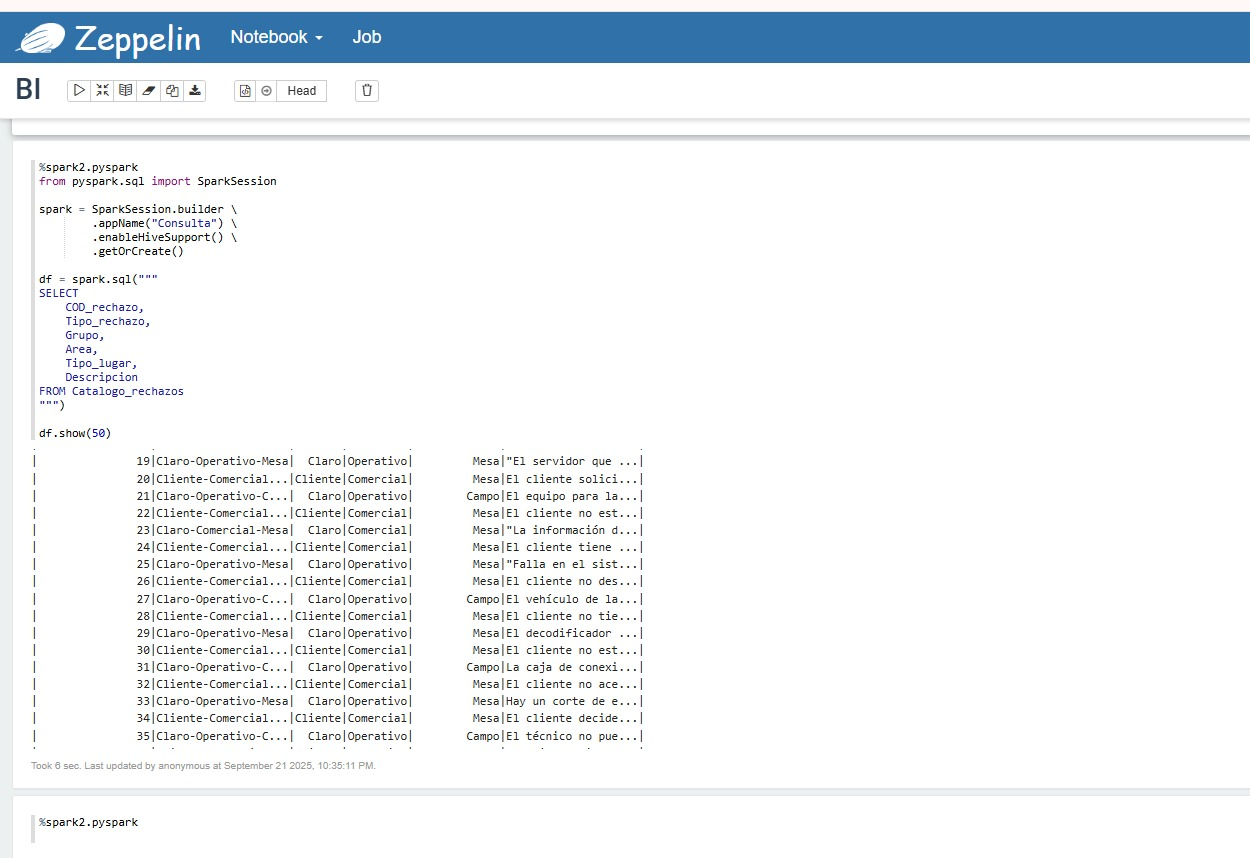
Es especialmente útil para realizar análisis interactivos, visualizar datos en proyectos de análisis de Big Data. Una de sus características más destacadas es su capacidad para ejecutar código en múltiples lenguajes, como Python, Scala, SQL, entre otros, y combinar estos lenguajes en un solo documento.

**Crear documentos:**

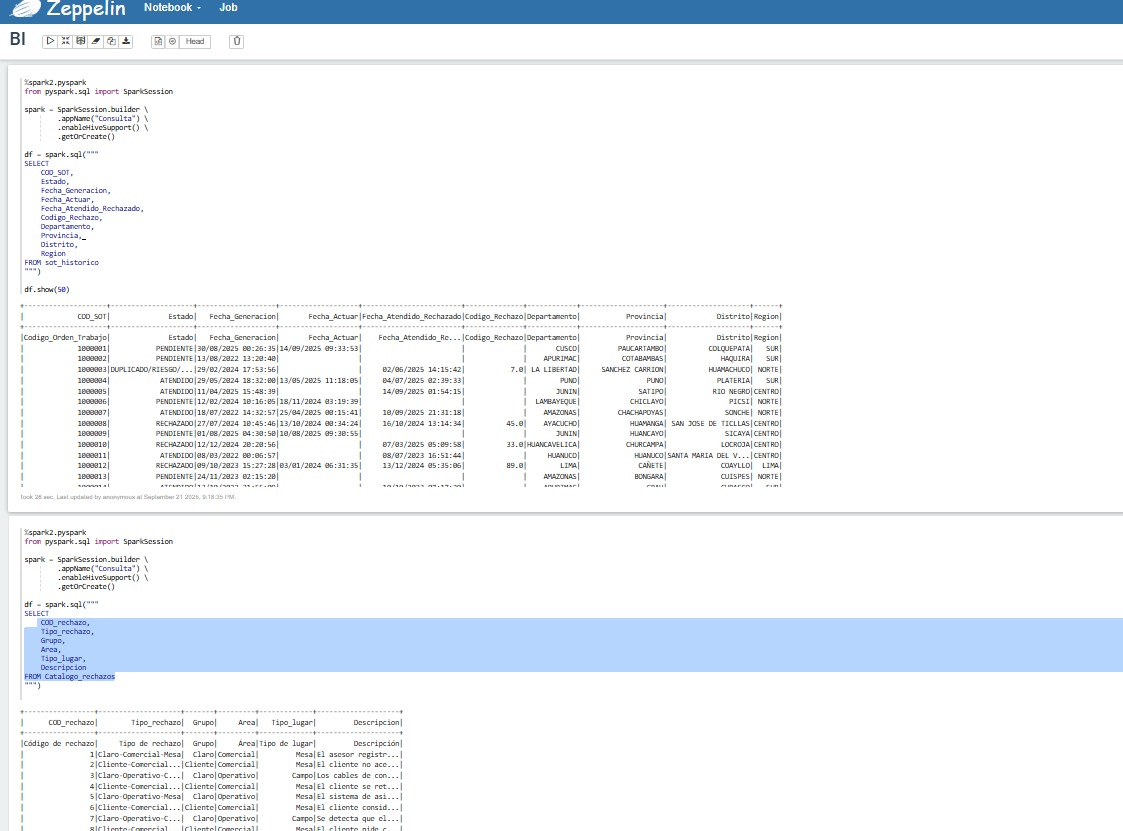
Uno de los principales usos de Zeppelin es trabajar con Apache Spark, una plataforma de procesamiento distribuido que permite el análisis de grandes volúmenes de datos. Zeppelin puede ejecutar comandos de PySpark (Python sobre Spark) o Scala para interactuar con Spark.

**Conectar a Datos en Hive:**

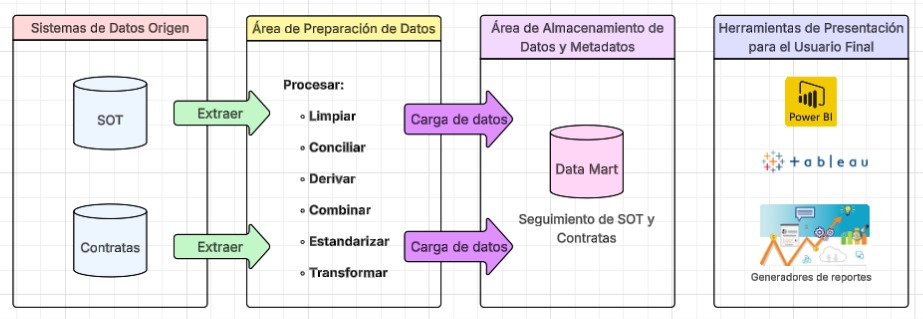
También podemos conectarnos con Hive gracias a una tabla externa. Para ello utilizamos este script.



El código proporcionado puede ser ejecutado directamente desde Apache Zeppelin usando el intérprete de PySpark (%spark2.pyspark) para interactuar con los datos en Hive.



**Diagrama de arquitectura inicial**

****

**Diseño del Data Mart para Claro Perú: Mercado Residencial Fijo**

**1. Sistemas de Datos Origen (Source Data Systems)**

Estos son los sistemas de origen de donde se extraen los datos.

La información que viene desde los ERP que manejan las contratas donde la información recae en una base de datos centralizada de la fija donde un BOT TOA la recoge y la deposita en las BD de I&M Instalación.

* SOT (Service Order Tracking): Base de datos que contiene todas las órdenes de trabajo para servicios de instalación y mantenimiento.
* Contratas (Contractor): Base de datos con información detallada sobre los contratistas externos que realizan los servicios.

**2. Área de Preparación de Datos**

En esta etapa, se limpian, combinan y transforman los datos de los sistemas de origen.

* Extraer: Se extraen los datos de ambas bases de datos (SOT y Contratas).
* Procesar:
* Limpiar: Eliminar datos incorrectos o inconsistentes (ej. fechas inválidas, códigos de SOT duplicados).
* Conciliar: Asegurar que los datos de ambas fuentes coincidan (ej. el contratista asignado a una SOT existe en la base de datos de Contratas).
* Derivar: Crear nuevas métricas o atributos (ej. tiempo promedio para completar una SOT, antigüedad del contratista).
* Combinar: Unir los datos de SOT y Contratas usando una clave común, como el ID del contratista o el número de SOT.
* Estandarizar: Formatear los datos de manera uniforme (ej. todos los nombres de contratistas en mayúsculas, fechas en un formato estándar).
* Transformar: Organizar los datos en un formato adecuado para el análisis dimensional.

**3. Área de Almacenamiento de Datos y Metadatos**

Esta es el área principal donde se almacenan los datos transformados, listos para ser consultados. Aquí se crea el Data Mart.

Data Mart (Seguimiento de SOT y Contratas): Una única base de datos optimizada para el análisis de los procesos de seguimiento.

**4. Herramientas de Presentación para el Usuario Final**

Una vez que los datos están en el Data Mart, los usuarios pueden acceder a ellos con diferentes herramientas para generar reportes, análisis y visualizaciones.

* Herramientas de Consulta Ad Hoc: Permiten a los usuarios crear consultas personalizadas (ej. SQL, Power BI, Tableau).
* Generadores de Reportes (Report writers): Herramientas para crear reportes estructurados sobre el desempeño de los contratistas o el estado de las SOT (ej. número de SOTs completadas por contratista, tiempo promedio por distrito).
* Herramientas de Modelado y Minería (Modeling/mining tools): Para análisis más avanzados (ej. pronosticar el rendimiento futuro de los contratistas, identificar cuellos de botella en el proceso).
* Herramientas de Visualización (Visualization tools): Para crear gráficos, dashboards y cuadros de mando interactivos que ayuden a entender la información de un vistazo.