**1. Planificación y Diseño**

**1.1 Definir Requisitos**

**Objetivos:**

* Establecer metas claras y medibles para la arquitectura de Big Data.
* Determinar los casos de uso específicos que la arquitectura debe soportar.

**Acciones Detalladas:**

1. **Reuniones con Stakeholders (1 de agosto de 2024 - 4 de agosto de 2024)**
   * **Paso 1:** Identificar a los stakeholders clave, incluyendo gerentes de producto, directores de TI, analistas de datos y usuarios finales.
   * **Paso 2:** Programar reuniones iniciales para discutir objetivos generales y expectativas.
   * **Paso 3:** Recolectar información sobre las necesidades específicas y los casos de uso deseados.
   * **Paso 4:** Documentar las discusiones y obtener la aprobación de los stakeholders sobre los objetivos preliminares.
2. **Documentación de Requisitos (5 de agosto de 2024 - 6 de agosto de 2024)**
   * **Paso 1:** Crear un documento de requisitos funcionales y no funcionales basado en la información recolectada.
     + **Requisitos Funcionales:** Casos de uso, tipos de datos, volúmenes de datos, frecuencia de procesamiento, etc.
     + **Requisitos No Funcionales:** Rendimiento, escalabilidad, seguridad, disponibilidad, etc.
   * **Paso 2:** Revisar el documento con los stakeholders para asegurar que todos los requisitos están correctamente capturados.
   * **Paso 3:** Obtener la aprobación final del documento de requisitos.
3. **Crear un Documento de Especificaciones Técnicas (7 de agosto de 2024)**
   * **Paso 1:** Convertir los requisitos en especificaciones técnicas detalladas.
     + **Especificaciones Técnicas:** Capacidades de procesamiento, almacenamiento necesario, herramientas de análisis requeridas, etc.
   * **Paso 2:** Detallar los KPI (Indicadores Clave de Desempeño) para medir el éxito del proyecto.
   * **Paso 3:** Revisar y obtener la aprobación de los stakeholders.

**1.2 Elegir Herramientas y Tecnologías**

**Objetivos:**

* Evaluar y seleccionar herramientas basadas en los requisitos definidos.

**Acciones Detalladas:**

1. **Evaluación de Herramientas (8 de agosto de 2024 - 11 de agosto de 2024)**
   * **Paso 1:** Identificar un conjunto inicial de herramientas y tecnologías que pueden cumplir con los requisitos.
     + Herramientas de Ingesta: Apache NiFi, Apache Sqoop, Apache Kafka.
     + Herramientas de Almacenamiento: HDFS, Apache Cassandra, Amazon S3.
     + Herramientas de Procesamiento: Apache Spark, Apache Flink, Apache Beam, PySpark.
     + Gobernanza y Seguridad: Apache Atlas, Apache Ranger.
     + Análisis y Machine Learning: Apache Mahout, TensorFlow, MLlib.
     + Visualización y BI: Grafana, Metabase, Apache Superset.
     + Orquestación: Apache Airflow.
     + Monitorización y Logging: ELK Stack (Elasticsearch, Logstash, Kibana), Prometheus.
     + Automatización y DevOps: Jenkins, Terraform, Ansible.
   * **Paso 2:** Crear una tabla comparativa de las características y capacidades de cada herramienta.
   * **Paso 3:** Realizar una evaluación preliminar para filtrar las herramientas más adecuadas.
2. **Pruebas de Concepto (12 de agosto de 2024 - 14 de agosto de 2024)**
   * **Paso 1:** Configurar entornos de prueba para cada herramienta seleccionada.
   * **Paso 2:** Ejecutar casos de uso simples para evaluar el rendimiento, facilidad de uso, integración y escalabilidad.
   * **Paso 3:** Documentar los resultados de las pruebas de concepto.
   * **Paso 4:** Revisar los resultados con el equipo y los stakeholders para tomar decisiones informadas.
3. **Documentación y Justificación (14 de agosto de 2024)**
   * **Paso 1:** Crear un documento que justifique la elección de cada herramienta, detallando las razones detrás de la selección.
   * **Paso 2:** Incluir en el documento cualquier limitación o consideración especial que deba tenerse en cuenta.
   * **Paso 3:** Obtener la aprobación final del documento de selección de herramientas.

**1.3 Diseño de la Arquitectura**

**Objetivos:**

* Crear un diagrama detallado de la arquitectura que muestre todos los componentes y sus interacciones.

**Acciones Detalladas:**

1. **Creación de Diagramas de Arquitectura (15 de agosto de 2024 - 18 de agosto de 2024)**
   * **Paso 1:** Utilizar herramientas de diagramación (e.g., Lucidchart, Draw.io) para crear diagramas iniciales.
   * **Paso 2:** Incluir todos los componentes seleccionados y sus interacciones.
   * **Paso 3:** Mostrar flujos de datos detallados entre los componentes.
     + **Ingesta:** Cómo los datos se mueven desde las fuentes hasta el sistema de almacenamiento.
     + **Almacenamiento:** Cómo se distribuyen y almacenan los datos.
     + **Procesamiento:** Cómo se procesan los datos para cumplir con los casos de uso.
     + **Análisis y BI:** Cómo se visualizan y analizan los datos procesados.
     + **Gobernanza y Seguridad:** Cómo se gestionan y protegen los datos.
2. **Revisión y Optimización de Diagramas (19 de agosto de 2024 - 20 de agosto de 2024)**
   * **Paso 1:** Revisar los diagramas iniciales con el equipo técnico y los stakeholders.
   * **Paso 2:** Obtener feedback y realizar ajustes necesarios para optimizar el diseño.
   * **Paso 3:** Asegurarse de que el diseño cumple con todos los requisitos funcionales y no funcionales.
3. **Documentación del Diseño de la Arquitectura (21 de agosto de 2024)**
   * **Paso 1:** Crear un documento que detalle el diseño de la arquitectura, incluyendo diagramas, flujos de datos y justificaciones para cada decisión de diseño.
   * **Paso 2:** Incluir en el documento planes de escalabilidad y resiliencia para asegurar que la arquitectura pueda manejar aumentos en la carga de trabajo y recuperar de fallos.
   * **Paso 3:** Obtener la aprobación final del diseño de la arquitectura por parte de los stakeholders.

**Cronograma**

| **Fase** | **Inicio** | **Finalización** |
| --- | --- | --- |
| Definir Requisitos | 1 de agosto de 2024 | 7 de agosto de 2024 |
| Elegir Herramientas y Tecnologías | 8 de agosto de 2024 | 14 de agosto de 2024 |
| Diseño de la Arquitectura | 15 de agosto de 2024 | 21 de agosto de 2024 |