**Planificación y Cronograma**

**1. Planificación y Diseño**

**Definir Requisitos**

* **Inicio:** 1 de agosto de 2024
* **Finalización:** 7 de agosto de 2024
* **Acciones:**
  + Reuniones con stakeholders
  + Documentación de requisitos
  + Creación del documento de especificaciones técnicas

**Elegir Herramientas y Tecnologías**

* **Inicio:** 8 de agosto de 2024
* **Finalización:** 14 de agosto de 2024
* **Acciones:**
  + Evaluación de herramientas
  + Pruebas de concepto
  + Documentación de elección de herramientas

**Diseño de la Arquitectura**

* **Inicio:** 15 de agosto de 2024
* **Finalización:** 21 de agosto de 2024
* **Acciones:**
  + Creación de diagramas de arquitectura
  + Definición de flujos de datos
  + Documentación del diseño de la arquitectura

**2. Infraestructura**

**Hardware**

* **Inicio:** 22 de agosto de 2024
* **Finalización:** 28 de agosto de 2024
* **Acciones:**
  + Evaluación de necesidades
  + Selección de hardware

**Servicios en la Nube**

* **Inicio:** 29 de agosto de 2024
* **Finalización:** 4 de septiembre de 2024
* **Acciones:**
  + Selección de proveedor de nube
  + Configuración de cuentas y permisos

**Contenedores y Orquestación**

* **Inicio:** 5 de septiembre de 2024
* **Finalización:** 11 de septiembre de 2024
* **Acciones:**
  + Configuración de Docker
  + Configuración de Kubernetes

**3. Ingesta de Datos**

**Apache NiFi**

* **Inicio:** 12 de septiembre de 2024
* **Finalización:** 18 de septiembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Diseño de flujos de datos

**Apache Sqoop**

* **Inicio:** 19 de septiembre de 2024
* **Finalización:** 25 de septiembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de trabajos de Sqoop

**Apache Kafka**

* **Inicio:** 26 de septiembre de 2024
* **Finalización:** 2 de octubre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de topics y producers/consumers

**4. Almacenamiento**

**HDFS**

* **Inicio:** 3 de octubre de 2024
* **Finalización:** 9 de octubre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de NameNode y DataNodes

**Apache Cassandra**

* **Inicio:** 10 de octubre de 2024
* **Finalización:** 16 de octubre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de clúster y replicación de datos

**Amazon S3**

* **Inicio:** 17 de octubre de 2024
* **Finalización:** 23 de octubre de 2024
* **Acciones:**
  + Configuración de buckets
  + Configuración de permisos y políticas de acceso

**5. Procesamiento de Datos**

**Apache Spark**

* **Inicio:** 24 de octubre de 2024
* **Finalización:** 30 de octubre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de clúster
  + Desarrollo y despliegue de aplicaciones PySpark

**Apache Flink**

* **Inicio:** 31 de octubre de 2024
* **Finalización:** 6 de noviembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Desarrollo y despliegue de aplicaciones de procesamiento en tiempo real

**Apache Beam**

* **Inicio:** 7 de noviembre de 2024
* **Finalización:** 13 de noviembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Desarrollo y despliegue de aplicaciones de procesamiento por lotes y en tiempo real

**PySpark**

* **Inicio:** 14 de noviembre de 2024
* **Finalización:** 20 de noviembre de 2024
* **Acciones:**
  + Configuración de entorno de desarrollo
  + Desarrollo y despliegue de scripts de procesamiento de datos

**6. Gobernanza y Seguridad**

**Apache Atlas**

* **Inicio:** 21 de noviembre de 2024
* **Finalización:** 27 de noviembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de metadatos y linaje de datos

**Apache Ranger**

* **Inicio:** 28 de noviembre de 2024
* **Finalización:** 4 de diciembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Configuración de políticas de acceso y seguridad

**Cifrado**

* **Inicio:** 5 de diciembre de 2024
* **Finalización:** 11 de diciembre de 2024
* **Acciones:**
  + Implementación de cifrado de datos
  + Configuración de claves de cifrado y gestión de claves

**7. Análisis y Machine Learning**

**Apache Mahout**

* **Inicio:** 12 de diciembre de 2024
* **Finalización:** 18 de diciembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Desarrollo y despliegue de algoritmos de aprendizaje automático

**TensorFlow**

* **Inicio:** 19 de diciembre de 2024
* **Finalización:** 25 de diciembre de 2024
* **Acciones:**
  + Instalación y configuración
  + Desarrollo y entrenamiento de modelos de aprendizaje profundo

**MLlib**

* **Inicio:** 26 de diciembre de 2024
* **Finalización:** 1 de enero de 2025
* **Acciones:**
  + Desarrollo y despliegue de modelos de aprendizaje automático

**8. Visualización y BI**

**Grafana**

* **Inicio:** 2 de enero de 2025
* **Finalización:** 8 de enero de 2025
* **Acciones:**
  + Instalación y configuración
  + Creación de dashboards para monitorización en tiempo real

**Metabase**

* **Inicio:** 9 de enero de 2025
* **Finalización:** 15 de enero de 2025
* **Acciones:**
  + Instalación y configuración
  + Creación de dashboards interactivos

**Apache Superset**

* **Inicio:** 16 de enero de 2025
* **Finalización:** 22 de enero de 2025
* **Acciones:**
  + Instalación y configuración
  + Creación de visualizaciones y dashboards

**9. Orquestación de Tareas**

**Apache Airflow**

* **Inicio:** 23 de enero de 2025
* **Finalización:** 29 de enero de 2025
* **Acciones:**
  + Instalación y configuración
  + Desarrollo de DAGs para orquestación de flujos de trabajo

**10. Monitorización y Logging**

**ELK Stack**

* **Inicio:** 30 de enero de 2025
* **Finalización:** 5 de febrero de 2025
* **Acciones:**
  + Instalación y configuración de Elasticsearch, Logstash y Kibana
  + Configuración de ingesta, búsqueda y visualización de logs

**Prometheus**

* **Inicio:** 6 de febrero de 2025
* **Finalización:** 12 de febrero de 2025
* **Acciones:**
  + Instalación y configuración
  + Configuración de monitorización y alertas

**11. Automatización y DevOps**

**CI/CD Pipelines**

* **Inicio:** 13 de febrero de 2025
* **Finalización:** 19 de febrero de 2025
* **Acciones:**
  + Configuración de Jenkins para automatización de despliegues
  + Creación de pipelines de CI/CD

**Infrastructure as Code (IaC)**

* **Inicio:** 20 de febrero de 2025
* **Finalización:** 26 de febrero de 2025
* **Acciones:**
  + Configuración de Terraform o Ansible para gestión de infraestructura
  + Desarrollo de scripts de IaC

**12. Despliegue y Configuración**

**Configuración de Entornos**

* **Inicio:** 27 de febrero de 2025
* **Finalización:** 5 de marzo de 2025
* **Acciones:**
  + Configuración de entornos de desarrollo, prueba y producción
  + Establecimiento de políticas de gestión de configuraciones

**Gestión de Configuraciones**

* **Inicio:** 6 de marzo de 2025
* **Finalización:** 12 de marzo de 2025
* **Acciones:**
  + Configuración de Ansible o Chef
  + Desarrollo de scripts de configuración

**Pruebas**

* **Inicio:** 13 de marzo de 2025
* **Finalización:** 19 de marzo de 2025
* **Acciones:**
  + Realización de pruebas unitarias, de integración y de rendimiento
  + Documentación de resultados y ajustes necesarios

**13. Documentación y Capacitación**

**Documentación Técnica**

* **Inicio:** 20 de marzo de 2025
* **Finalización:** 26 de marzo de 2025
* **Acciones:**
  + Creación de guías de usuario, diagramas y especificaciones técnicas
  + Documentación de procesos y configuraciones

**Capacitación**

* **Inicio:** 27 de marzo de 2025
* **Finalización:** 2 de abril de 2025
* **Acciones:**
  + Realización de formación para el equipo en herramientas y tecnologías utilizadas
  + Organización de workshops y sesiones de capacitación