

Coordinación de Actividades

- **Coordinación.**
 1. División/Despacho/Difusión.
 2. Ejecución.
 3. Consolidación.
- **Replicación.**
 1. Difusión.
 2. Ejecución (todos lo mismo).
 3. Consolidación.
- **Acceso a Recursos Compartidos.**
 1. Serialización de requests.
 2. Ejecución serial.

Open MPI

- Transmisión y recepción de mensajes.
- **Ejecución transparente** de 1 a N nodos.
- Librería con **abstracciones** de uso general.
 - Foco en **cómputo distribuido**.
- Implementa middleware de comunicación de grupos.
 - MPI_Send, MPI_Recv, MPI_Bcast.
 - MPI_Scatter, MPI_Gather.
 - MPI_Allgather, MPI_Reduce.

Apache Flink

- Plataforma p/ **procesamiento distribuido** de datos.
- Motor de ejecución de **pipelines de transformación**.
- Framework p/ crear pipelines.
- **Ventanas** para streaming.
- Casos de uso: ETL, Data Pipelines.

Dataflow

DAG de operaciones sobre un flujo de datos.

- **Streams:** flujo de información que eventualmente no finaliza.
- **Batches:** dataset de tamaño conocido.

Bloques de pipeline

- **Source:** inyecta datos.
- **Transformation (operador):** modifica o filtra datos.
- **Sink:** almacenamiento final

Apache Beam

Modelo de **definición de pipelines** de procesamiento de datos con **portabilidad de lenguajes y motores de ejecución**.

- **Runners:**
 - Ejecución directa.
 - Motores de cluster.
 - Plataformas cloud.

Bloques de un pipeline

- **Input y Output** (*source y sink*).
- **PCollection**: colecciones paralelizables de elementos (*streams*).
- **Transformations**: modificaciones elemento a elemento (*operators*).