

# Clúster EMR

- **Nodo principal:** Gestiona el clúster
  - Monitorea el estado de las tareas y la salud del clúster
  - Instancia EC2 única (puede ser incluso un clúster de un solo nodo)
  - También conocido como "nodo líder"
- **Nodo central:** Aloja los datos en HDFS y ejecuta tareas
  - Puede ser escalado hacia arriba y hacia abajo
- **Nodo de tareas:** Ejecuta tareas, no aloja datos (*\*Opcional*)
  - No hay riesgo de pérdida de datos al removerlo
  - Buen uso de instancias de spot

