

# EJERCICIO 4

TUNNING Y PATRONES DE DISEÑO



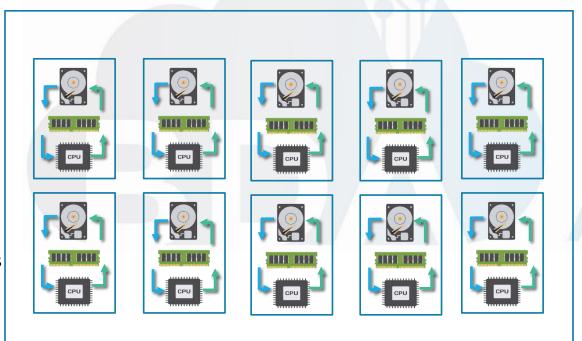
## Clúster de ejemplo

Lo primero que debemos saber es la potencia del clúster y qué significa el 100% de la potencia

1000 TB DISCO

1000 GB RAM

400 NÚCLEOS DE CPU



#### **POTENCIA TOTAL**

RAM: 1000 GB CPU: 400 CPU



### EJERCICIO 4

# IMPLEMENTAR EL SIGUIENTE PROCESO



### EJERCICIO 4

#### **CONSIDERAR:**

- 1. CUALQUIER ARCHIVO QUE VENGA DE DISCO DURO: REPARTITION Y CACHÉ
- 2. CUALQUIER DATAFRAME QUE SE ALMACENARÁ EN DISCO DURO: **CACHÉ**
- 3. USAR CHECKPOINT PARA LIBERAR DE LA RAM **CADENAS DE PROCESOS QUE YA NO SE USARÁN**

[RESOLVERLO USANDO EL **10% DE LA POTENCIA DEL CLÚSTER**] Ejercicio 4 [CONSIDERAR QUE CUALQUIER ARCHIVO QUE VENGA DE DISCO DURO: REPARTITION Y CACHÉ] **LECTURA MODELAMIENTO REGLAS DE CALIDAD** PREPARACIÓN DE TABLONES **PROCESAMIENTO** ALMACENAMIENTO <<T\_P>> CHECKPOINT, REPARTITION Y CACHÉ REPARTITION Y CACHÉ IfTransaccionLim pib [ • dfTransaccion transacciones\_balcarias.json <<T\_P\_E>> dfPersonaLimp dfPersona REPORTE 1.csv dfReporte1 <<T\_P\_E\_R>> (O) dfTransacciones dfEmpresaLimpio dfEmpresa dfTablon dfReporte2 REPORTE\_2.csv **O**.0 df3 RIESGO\_CREDITICIO.csv <<RIESGO PONDERADO>> dfReporte3 REPORTE\_3.csv CACHÉ dfRiesgoLimpio dfRiesgo CHECKPOINT, REPARTITION Y CACHÉ