

Sistemas de almacenamiento y procesamiento distribuido

Presentación

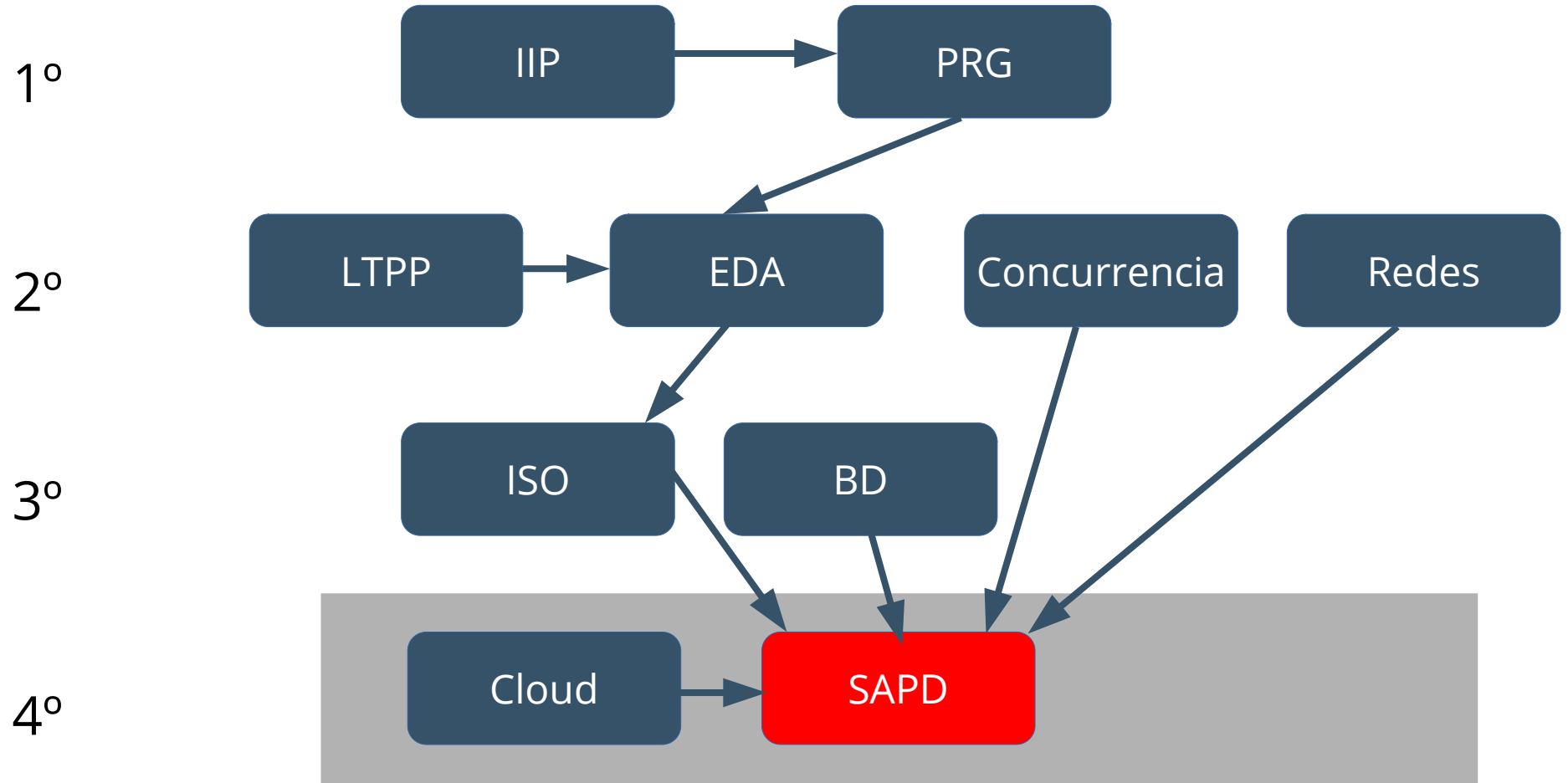
Profesores

- Javier Esparza Peidro – jesparza@dsic.upv.es
- Profesor responsable
- Teoría y laboratorio
- Despacho F4D8
- Tutorías: bajo demanda

Detalles de la asignatura

- Asignatura de 4A
- Horario
 - Teoría y laboratorio
 - (C3AI2): Miércoles 17:15-19:15
 - (C3L1-C3AI2): Viernes 10:15-12:15 (4,6,9,11,13,17,20,21)

Relación con otras asignaturas



¿Por qué?

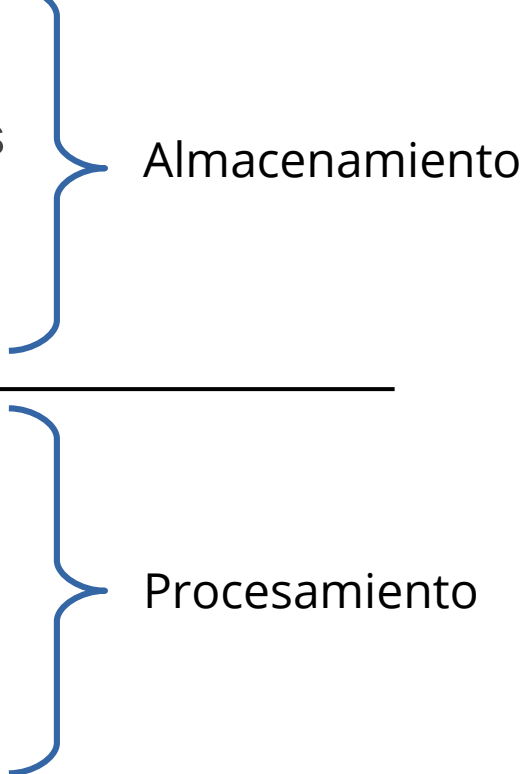
- En la actualidad cada usuario genera al menos 1,7 Mb por segundo de información
- Las técnicas de inteligencia artificial actuales requieren ingentes cantidades de datos
- La gestión de datos es uno de los mayores desafíos de la informática

Objetivos

- Analizar el funcionamiento de los sistemas de datos actuales
- Aplicar técnicas de procesamiento en batch y en streaming
- Diseñar e integrar sistemas de datos

Teoría

Contenido

- Tema 1: Introducción
 - Datos, sistemas de datos, propiedades, arquitecturas
 - Tema 2: Almacenamiento de datos
 - Agregados, indexación, replicación, consistencia
-
- Tema 3: Procesamiento en Batch
 - MapReduce, Hadoop, Spark
 - Tema 4: Procesamiento en Streaming
 - Spark, Kafka
- 

Teoría

Metodología

- Exposición de la teoría mediante traspas, elaboración de ejemplos y propuesta de ejercicios
- Los alumnos irán elaborando un portafolio a partir de los ejercicios propuestos, se depositarán en git
- Materiales: apuntes, traspas, documentación oficial

Laboratorio

Contenido

- Un laboratorio sobre cada bloque (2x)

Metodología

- En cada laboratorio se trabaja un bloque
- La solución se depositará en git y PoliformaT
- Materiales: documentación sobre los laboratorios, material de teoría, máquinas virtuales DSIC

Evaluación

Evaluación continua

- Nota final = 30% teoría + 20% portafolio + 50% labs
- Teoría (2x): examen tipo test tras cada bloque (10q/20')
- Portafolio (2x): trabajo diario entregado antes del examen
- Labs (2x): laboratorios entregados a lo largo del curso

Residual

- El examen diferenciará Teoría (50%) + Labs (50%)

¿Empezamos?