



Sistemas Distribuidos I (75.74)

Flights Optimizer

TP Escalabilidad: Middleware y Coordinación de Procesos

Docentes

- Pablo D. Roca
- Guido Albarello

- Gabriel Robles
- Franco Barreneche

- Tomás Nocetti
- Nicolás Zulaica



Requerimientos Funcionales

- Se solicita un sistema distribuido que analice 6 meses de registros de precios de vuelos de avión p/proponer mejoras en la oferta a clientes.
- Los registros poseen trayectos (aeropuertos origen-destino), tarifa total, distancia total, duración, cada segmento con escalas y aerolíneas.
- Los registros se ingresan progresivamente al ser escaneados de internet.
- Se debe obtener:
 - ID, trayecto, precio y escalas de vuelos de 3 escalas o más.
 - ID, trayecto y distancia total de vuelos cuya distancia total sea mayor a cuatro veces la distancia directa entre puntos origen-destino.
 - ID, trayecto, escalas y duración de los 2 vuelos más rápidos para cada trayecto entre todos los vuelos de 3 escalas o más.
 - El precio avg y max por trayecto de los vuelos con precio mayor a la media general de precios.



Requerimientos No Funcionales

- Para construir una simulación realista se define la serie de datos:
 - <https://kaggle.com/datasets/dilwong/flightprices>
 - <https://kaggle.com/datasets/pabloroca/airports-opendatahub>
 - <https://kaggle.com/datasets/pabloroca/flight-prices-2M> (extract)
- Usar la librería *geopy.distance* para calcular distancias.
- El sistema debe estar optimizado para entornos multicomputadoras
- Se debe soportar el incremento de los elementos de cómputo para escalar los volúmenes de información a procesar
- Se requiere del desarrollo de un Middleware para abstraer la comunicación basada en grupos.
- Se debe soportar una única ejecución del procesamiento y proveer *graceful quit* frente a señales SIGTERM.



Se espera del alumno:

- Empleo del tiempo de consultas en clase para resolver dudas y clarificar el negocio del sistema a construir previo a su diseño
- Exposición y verificación en clase de la arquitectura propuesta antes de iniciar su implementación
- Empleo del grupo de correos para realizar consultas que no pudieran ser resueltas en clase
- Consideración de prácticas distribuidas según lo estudiado en clase para elaborar una arquitectura flexible, escalable y robusta
- Aprobación del cuerpo docente para el uso de cualquier librería.
- Demo del sistema en funcionamiento previamente ensayada



Entrega - Parte 1

- Fecha de entrega: 26/09/2023
- Formato de entrega:
 - Entrega presencial.
 - Documento de arquitectura 4+1 Views o C4Model incluyendo al menos: diagrama de robustez, despliegue y actividades.
 - Listado de tareas a ejecutar y división entre integrantes.
 - Se discutirá y validará el diseño planteado por el alumno para que pueda continuar con el TP, **no se evaluará código.**



Entrega - Parte 2

- Fecha de entrega: 12/10/2023
- Fecha de re-entrega: 26/10/2023
- Formato de entrega:
 - Entrega digital mediante correo personal.
 - Demostración del sistema en funcionamiento.
 - Documento de arquitectura 4+1 Views o C4Model incluyendo al menos: diagramas de robustez, despliegue, actividades, paquetes, secuencia y DAG.