



PONTIFICIA UNIVERSIDAD CATÓLICA DE CHILE
ESCUELA DE INGENIERÍA
DEPARTAMENTO DE CIENCIA DE LA COMPUTACIÓN

IIC2523 Sistemas Distribuidos (II/2018) - Sección 1

Profesor: Gabriel Vidal Salazar

Interrogación 1

Fecha: Martes 24 de septiembre de 2018

Parte 1 (10 pts): Preguntas

1. (2 pts) Según la taxonomía de Flynn ¿Cuál es la clasificación de los procesadores modernos? ¿Qué paralelismo explota?
2. (2 pts) Indica a que corresponde la clasificación SPMD y MPMD. Menciona un ejemplo para cada uno de ellos.
3. (2 pts) ¿Por qué es necesario comunicar procesos en un sistema distribuido? ¿Qué rol cumple MPI en este caso?
4. (2 pts) ¿En qué tecnología se basan los proveedores de sistemas de IaaS (como AWS, Google, etc)? ¿Qué características tiene que permite prestar servicio a distintos clientes?
5. (2 pts) ¿Cuál es el problema del mecanismo de coherencia basado en directorio?

Parte 2 (50 pts): Ayudando a un amigo

Uno de tus amigos está abriendo una empresa de 5 personas (en Chile) y quiere centralizar el almacenamiento de los archivos. Un de las cosas que ha detectado es que los archivos que utilizan son pequeños y, en general, siempre es modificado por una persona (por el momento). Él te dice que ha investigado un poco el tema y está convencido que GFS (Google File System) es una buena idea en este caso, ya que le permitiría tener respaldada la información, a bajo costo, y le permitiría crecer. También le menciona que Lustre podría ser otra alternativa (de bajo costo), ya que su empresa va a crecer y necesita rapidez en el acceso a la información.

Para las siguientes preguntas considera relevante las siguientes características: costo, rapidez, concurrencia, consistencia, seguridad, escalabilidad y replicación.

1. (4 pts) ¿Estás de acuerdo con la idea planteada por tu amigo? Justifica tu respuesta para ambos sistemas de archivos distribuidos y en relación a lo planteado por él.
2. (4 pts) Para este caso en particular ¿Cuál sería tu recomendación? ¿Por qué?

3. (4 pts) Si consideráramos un futuro crecimiento de la empresa (a 50 personas) ¿Cambia tu recomendación? ¿Cuáles son las razones?

Después de un tiempo, tu amigo te vuelve a contactar porque su empresa ha crecido bastante y tiene que abrir operaciones en Europa. Actualmente su sistema de archivos distribuido, que tu configuraste, ha funcionado bastante bien pero tu amigo quiere que este sistema esté accesible desde Europa también.

4. (3 pts) ¿Es una buena utilizar una VPN en este caso? Basa tu respuesta en los aspectos de rapidez y seguridad ¿Qué tipo de configuración habría que realizar?
5. (3 pts) ¿Y utilizar la red TOR? Realiza el mismo análisis anterior.

Ahora, ya eres parte de la empresa de tu amigo como asesor técnico y, junto con el equipo de ingeniería, han decidido organizar la aplicación principal como microservicios. Tal como muestra la imagen anterior.

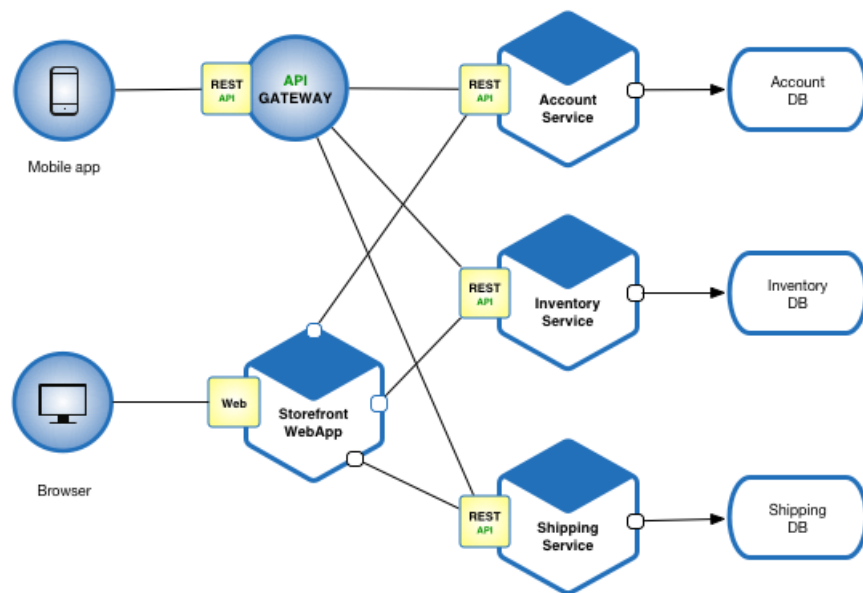


Figure 1: Ejemplo de microservicios (<https://microservices.io/patterns/microservices.html>)

Una vez decidido esto hay varios otros aspectos de infraestructura que hay que analizar.

6. (5 pts) Analiza el cumplimiento del teorema CAP para la arquitectura que se muestra en la imagen. Puedes hacer un supuesto en caso de que lo necesites.
7. (2 pts) Si agregamos replicación de las bases de datos de cada microservicio ¿Cómo se ve afectado el teorema CAP?

Actualmente la empresa cuenta con dos servidores físicos y el equipo de ingeniería se encuentra dividido entre virtualizar o instalar el sistema directamente en los servidores. Finalmente, el objetivo principal es poder tener disponibilidad del servicio.

8. (5 pts) ¿Es posible lograr disponibilidad con ambas opciones (antes descritas)? ¿Cómo?
9. (5 pts) Han decidido que virtualizar es el mejor camino para la empresa. Ahora la discusión se centra en virtualizar basado en contenedores o en hipervisor ¿Cuál es la mayor diferencia entre ambos? ¿Cuál utilizarías si otro de los objetivos es mayor rendimiento? ¿Por qué?

Uno de los grandes problemas que está teniendo la empresa tiene relación con la compra de insumos. Por lo que han determinado que es prioritario poder determinar cuándo se requiere adquirir productos para la venta. Para esto se decidió:

- Estudiar la demanda de los últimos 12 meses para cada producto
- Obtener el promedio de la demanda
- Si en los últimos 3 meses está por sobre el promedio se debe adquirir ese producto.
- Sólo se realiza la compra si más de un cuarto de los productos requiere que se adquieran.

Para esto se entregará un archivo de texto que en la primera columna tendrá el identificador del producto y, en las siguientes, la demanda para los 12 meses anteriores.

10. (10 pts) Escribe (en pseudocódigo) un programa usando MPI para poder realizar lo anterior.
11. (5 pts) ¿Cómo cambia tu programa si la compra se realiza por producto, sólo con el estudio de la demanda?