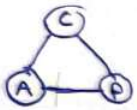
TEMA 6

TEOREMA CAP



En un sistema distribuido NO podemos mantener → SOLO 2 A LA VEZ
- CONSISTENCIA FUERTE (C)
- DISPONIBILIDAD (A)
- PARTICIONADO (P) → Tolerancia al particionado

SOLO 2 A LA VEZ: PA / PC / AC

AC → Sin Particiones: Garantizamos { Consistencia Fuerte, Disponibilidad }
↳ NO VIABLE :(
- No admitimos que la red se particione
- Habrá que garantizar conectividad (replicación)
- Difícil de asegurar → costoso para redes grandes

PC → Sin Disponibilidad: Garantizamos { Consistencia Fuerte, Particionado }
↳ ESCALA NAL :(
- Al particionarse, el grupo con menos nodos deja de funcionar, hay usuarios sin atención para garantizar la consistencia.

PA → Sin Consistencia: Garantizamos { Disponibilidad, Particionado }
↳ ESCALABLE → MEJOR OPCIÓN
- Tras una partición → Siguen trabajando, podrá haber divergencias que deberán resolverse con una consistencia eventual.

REPLICACIÓN MULTI-MÁSTER

Los modelos de replicación Activa o Pasiva son FUERTEMENTE CONSISTENTES → Limita Escalabilidad
↳ Para mejorarlo aparece la Replicación Multi-Máster → Extensión del Modelo Pasivo

→ Cada 'cliente' tiene una réplica principal (Máster), cada copia se encarga de una actividad
Cada nodo es principal para algunas tareas y secundario para otras.

PEREZOSO → Los cambios NO se propagan inmediatamente, se acumulan y se propagan tras un número de cambios. Lo hace + rápido.

VENTAJAS →
- Alta escalabilidad
- Sobrecarga mínima
- Operaciones NO deterministas

INCONVENIENTES →
- Si falla un master, todas sus peticiones pueden perderse → Inconsistencias
- Gestión de fallos igual al modelo pasivo
- NO soporta fallos bizantinos

ALMACENES NoSQL

- Base de datos que NO sigue el modelo relacional → para poder ser altamente escalable y manejar muchas escrituras y lecturas
- SIMPLIFICAR EL ESQUEMA → Utilizan tablas clave-valor →
  - simplifica el sistema
  - reduce el espacio de almacenamiento
  - caben en la memoria ppal.
- ELIMINACIÓN DE TRANSACCIONES → Permiten SOLO operaciones individuales → crea atomicidad en las operaciones y evita bloqueos.

↓

TIPOS

- Almacenes clave-valor
  - Esquema simple compuesto por dos campos: clave, valor
  - La búsqueda se realiza por la clave.
- Almacenes de documentos
  - Esquema compuesto por objetos con nº variable de atributos. (que pueden ser objetos)
  - La búsqueda se puede realizar por atributos.
- Almacenes de registros extensibles
  - Esquema compuesto por tablas con nº variable de columnas
  - Los datos se separan por filas o columnas y estas en diferentes nodos.
  - Sharding → para acceder a diferentes datos buscas en los nodos.

ELASTICIDAD ↑ escalable + adaptable

- Capacidad para adaptarse a variaciones de demanda → adaptarse a cambios de carga
- Asigna a cada aplicación la cantidad justa de recursos de manera autónoma
  - ↳ cubriendo la demanda existente con la mayor precisión posible
- REQUIERE
  - ① Sistema de Monitorización → Supervisa la carga y rendimiento, identifica dónde se necesitan + o - recursos
  - ② Sistema de Actuación → Automatiza la reconfiguración de recursos y servicios
    Ajusta la asignación de recursos en respuesta a los cambios de demanda.
    Pone en marcha lo detectado por el sistema de monitorización

CONTENCIÓN Y CUELLOS DE BOTELLA

↑

Situación en la que un recurso se convierte en un cuello de botella para todo el sistema.

↳ Relentiza todo el proceso

- Causas
  - → Centralización: Todas las tareas pasan por un nodo que realiza todas las tareas pesadas, relentiza todo el sistema.
    SOLUCIÓN: Evitar centralización de tareas pesadas.
  - → Condiciones de carrera: Cuando compaiten recursos
    SOLUCIÓN: Usar herramientas de sincronización
    Optar por asincronía
  - → Tráfico excesivo: SOLUCIÓN: Replicar recursos y mantener consistencia
    Cambiar accesos remotos a locales para reducir latencia