

Guía de Ejercicios

Selección de Bases de Datos NoSQL

Gestión y Arquitectura de Datos, Universidad de San Andrés

Si encuentran algún error en el documento o hay alguna duda, mandenme un mail a rodriguezr@udesa.edu.ar y lo revisamos.

1. Ejercicios

A continuación se presentan 10 casos de uso reales donde se debe decidir qué tipo(s) de base(s) de datos NoSQL utilizar. Para cada caso:

- Analice los requerimientos del sistema
- Identifique las características clave que influyen en la elección
- Seleccione la(s) tecnología(s) más apropiada(s)
- Justifique su elección

Intenten resolver los casos por su cuenta. Las respuestas se encuentran al final del documento.

1.1. Caso 1: E-commerce Básico

Una pequeña tienda online está comenzando su negocio. Necesitan almacenar un catálogo de productos con descripciones y categorías y tener carritos de compra temporales. También necesitan saber el historial de pedidos por usuario. La tienda espera tener unos 1,000 productos y 5,000 usuarios en su primer año.

1.2. Caso 2: Red Social de Fotografías

Una startup está desarrollando una red social centrada en compartir fotos. Necesitan almacenar fotos y sus metadatos, manejar relaciones de “seguir” entre usuarios, gestionar likes y comentarios y tener un feed personalizado para cada usuario. La red social espera tener unos 10,000 usuarios y 100,000 fotos en su primer año.

1.3. Caso 3: Sistema de Monitoreo Industrial

Una fábrica necesita monitorear sus máquinas en tiempo real. Los sensores envían datos cada segundo y se necesita almacenar históricos por 6 meses. La fábrica tiene 100 máquinas, cada una con 5 sensores diferentes.

1.4. Caso 4: Plataforma de Streaming de Videos

Un servicio de streaming necesita un catálogo de videos con metadatos, un sistema de recomendaciones, un historial de visualización por usuario, preferencias y configuraciones de usuario y caché de videos populares. El servicio espera manejar millones de usuarios concurrentes.

1.5. Caso 5: Sistema de Reservas de Hotel

Una cadena hotelera internacional necesita un sistema para gestionar disponibilidad en tiempo real, manejar reservas y cancelaciones, perfiles de clientes con preferencias, historial de estadías, sistema de fidelización y reportes de ocupación. La cadena tiene 50 hoteles en 20 países diferentes.

1.6. Caso 6: Sistema de Logística

Una empresa de logística necesita un sistema para rastrear ubicación de paquetes en tiempo real, optimizar rutas de entrega, gestionar flota de vehículos, historial de entregas y análisis de eficiencia. La empresa maneja 1,000 entregas diarias con 100 vehículos en 5 ciudades.

1.7. Caso 7: Plataforma de Juegos Online

Un estudio de videojuegos necesita una plataforma para perfiles de jugadores y estadísticas, rankings y logros, inventarios de items in-game, chat en tiempo real, matchmaking y transacciones de items entre jugadores. Esperan picos de 100,000 jugadores simultáneos.

1.8. Caso 8: Sistema de Salud

Un hospital necesita modernizar su sistema para historias clínicas electrónicas, gestión de citas y turnos, resultados de estudios (incluyendo imágenes),

prescripciones médicas, seguimiento de tratamientos e integración con otros centros médicos. El hospital atiende 2,000 pacientes diarios y necesita acceso 24/7 a los datos.

1.9. Caso 9: Plataforma IoT Smart City

Una ciudad inteligente necesita gestionar datos de sensores de tráfico, monitoreo de calidad del aire, control de semáforos, sistema de parking inteligente, consumo energético y análisis predictivo para servicios públicos. La ciudad tiene 1 millón de habitantes y miles de sensores distribuidos.

1.10. Caso 10: Sistema Financiero

Un banco digital necesita manejar transacciones en tiempo real, detección de fraude, historial de operaciones, perfiles de clientes, análisis de riesgo crediticio, reportes regulatorios e integración con otros bancos. El banco tiene 500,000 clientes y procesa millones de transacciones diarias.

2. Anexo: Respuestas

2.1. Respuesta Caso 1: E-commerce Básico

MongoDB (Base de Datos Documental) como base principal:

- Ideal para el catálogo de productos por su esquema flexible
- Permite almacenar productos con diferentes atributos y categorías
- Bueno para el historial de pedidos con documentos anidados

Redis (Base de Datos Clave-Valor) para:

- Carritos de compra temporales
- Caché de productos populares
- Sesiones de usuario

Justificación: Para un e-commerce pequeño, esta combinación ofrece:

- Flexibilidad para el crecimiento del catálogo
- Alta velocidad para operaciones frecuentes
- Simplicidad de implementación
- Buena relación costo-beneficio

2.2. Respuesta Caso 2: Red Social de Fotografías

MongoDB (Base de Datos Documental) para:

- Perfiles de usuario
- Metadatos de fotos
- Comentarios y likes

Neo4j (Base de Datos de Grafos) para:

- Relaciones de "seguir"

- Recomendaciones de usuarios
- Análisis de redes sociales

Redis (Base de Datos Clave-Valor) para:

- Caché de feeds
- Contadores de likes
- Sesiones activas

Justificación: Esta combinación permite:

- Escalabilidad para el crecimiento rápido
- Optimización de consultas sociales
- Alta velocidad en operaciones frecuentes

2.3. Respuesta Caso 3: Sistema de Monitoreo Industrial

Cassandra (Base de Datos Columnar) para:

- Almacenamiento de datos de sensores
- Series temporales
- Alta disponibilidad

Redis (Base de Datos Clave-Valor) para:

- Análisis en tiempo real
- Caché de últimas lecturas
- Alertas temporales

Justificación:

- Cassandra es excelente para series temporales y escrituras intensivas
- Permite escalabilidad horizontal para históricos
- Redis complementa con análisis en tiempo real

2.4. Respuesta Caso 4: Plataforma de Streaming

MongoDB (Base de Datos Documental) para:

- Catálogo de videos y metadatos
- Perfiles de usuario
- Historial de visualización

Redis (Base de Datos Clave-Valor) para:

- Caché de videos populares
- Sesiones de usuario
- Estado de reproducción

Neo4j (Base de Datos de Grafos) para:

- Sistema de recomendaciones
- Análisis de preferencias

Justificación: Esta arquitectura permite:

- Alta disponibilidad
- Escalabilidad horizontal
- Recomendaciones personalizadas eficientes

2.5. Respuesta Caso 5: Sistema de Reservas de Hotel

MongoDB (Base de Datos Documental) para:

- Perfiles de clientes
- Detalles de hoteles
- Historial de estadias

Redis (Base de Datos Clave-Valor) para:

- Disponibilidad en tiempo real

- Bloqueo temporal de habitaciones
- Caché de búsquedas frecuentes

Justificación:

- Permite consistencia eventual entre regiones
- Alta disponibilidad para reservas
- Buen manejo de datos distribuidos

2.6. Respuesta Caso 6: Sistema de Logística

MongoDB (Base de Datos Documental) para:

- Información de paquetes
- Detalles de vehículos
- Historial de entregas

Redis (Base de Datos Clave-Valor) para:

- Ubicación en tiempo real
- Estado de entregas
- Caché de rutas

Neo4j (Base de Datos de Grafos) para:

- Optimización de rutas
- Análisis de red logística

Justificación:

- Permite tracking en tiempo real
- Optimización eficiente de rutas
- Análisis histórico de eficiencia

2.7. Respuesta Caso 7: Plataforma de Juegos Online

MongoDB (Base de Datos Documental) para:

- Perfiles de jugadores
- Inventarios
- Estadísticas

Redis (Base de Datos Clave-Valor) para:

- Chat en tiempo real
- Estado de partidas
- Rankings temporales
- Matchmaking

Neo4j (Base de Datos de Grafos) para:

- Relaciones entre jugadores
- Análisis de comportamiento

Justificación:

- Alta velocidad para operaciones en tiempo real
- Escalabilidad para picos de usuarios
- Análisis social para matchmaking

2.8. Respuesta Caso 8: Sistema de Salud

MongoDB (Base de Datos Documental) para:

- Historias clínicas
- Resultados de estudios
- Prescripciones

Redis (Base de Datos Clave-Valor) para:

- Gestión de turnos
- Caché de datos frecuentes
- Estado de pacientes

Neo4j (Base de Datos de Grafos) para:

- Relaciones entre diagnósticos
- Análisis de tratamientos
- Seguimiento de derivaciones

Justificación:

- Alta disponibilidad 24/7
- Seguridad y consistencia de datos
- Análisis complejo de historias clínicas

2.9. Respuesta Caso 9: Plataforma IoT Smart City

Cassandra (Base de Datos Columnar) para:

- Datos de sensores
- Métricas ambientales
- Históricos de tráfico

Redis (Base de Datos Clave-Valor) para:

- Estado actual de sensores
- Control de semáforos
- Alertas en tiempo real

Neo4j (Base de Datos de Grafos) para:

- Análisis de patrones de tráfico
- Optimización de rutas

- Relaciones entre servicios

Justificación:

- Manejo eficiente de grandes volúmenes de datos
- Alta velocidad para decisiones en tiempo real
- Capacidad de análisis predictivo

2.10. Respuesta Caso 10: Sistema Financiero

MongoDB (Base de Datos Documental) para:

- Perfiles de clientes
- Historial de transacciones
- Documentación

Redis (Base de Datos Clave-Valor) para:

- Procesamiento de transacciones
- Detección de fraude en tiempo real
- Caché de datos frecuentes

Cassandra (Base de Datos Columnar) para:

- Logs de auditoría
- Históricos de operaciones
- Reportes regulatorios

Justificación:

- Garantiza consistencia en transacciones
- Alta disponibilidad y tolerancia a fallos
- Cumplimiento regulatorio
- Escalabilidad para millones de operaciones