

# Conceptos básicos y modelos NoSQL PEC1

## Ejercicio 1 (30%)

A partir de la lectura de los apuntes (locuciones de los vídeos) de los temas I y II se pide responder de manera concisa (una página y media en total) a las siguientes preguntas:

1. ¿Qué significa que las bases de datos NoSQL orientadas hacia agregados favorecen los esquemas de crecimiento horizontal (o escalabilidad horizontal)?
2. Explica las ventajas del modelo de procesamiento map-reduce.
3. Explica brevemente qué es la persistencia polígota y cómo se refleja en un proyecto.
4. Explica las razones por las que un modelo relacional no es una buena opción cuando se deben procesar flujos de datos.
5. ¿Qué diferencias hay entre el modelo relacional y el modelo en grafo en cuanto a la representación de las relaciones?

## Ejercicio 2 (30%)

A partir de la lectura de los apuntes (locuciones de los vídeos) de los temas I y II indica si te parecen ciertas o falsas las siguientes afirmaciones.

Para cada una de las afirmaciones indica si es cierta o falsa, justificando la respuesta mediante lo que has leído en los materiales. En cada justificación deberá indicarse la cita de los apuntes, vídeo o libros en la que se sustenta.

**No serán válidas las respuestas que no se justifiquen.**

Se valorará la concisión (una página y media para las 5 afirmaciones como máximo).

### Afirmación 1

En un modelo de agregación, la estructuración de agregados de un mismo tipo no puede variar en ningún caso. Es decir, todos los agregados del mismo tipo deben seguir la misma estructura.

### Afirmación 2

El modelo en grafo es igual de fácil de escalar que los modelos agregados.

### Afirmación 3

En un proyecto donde se debe priorizar la disponibilidad frente a la consistencia de los datos, son igual de recomendables una base de datos relacional que una base de datos NoSQL.

### Afirmación 4

Los modelos de agregación son una buena elección en el caso de modelos conceptuales donde predominen asociaciones jerárquicas, asociaciones 1 a 1 y relaciones de tipo parte-todo.

### Afirmación 5

El modelo documental es un caso particular del modelo clave-valor, pero más flexible que el clave-valor.

## Ejercicio 3 (20%)

La biblioteca de la UOC ha diseñado un sistema para conocer los hábitos de las personas que solicitan préstamos de sus libros. Para ello se quiere explotar la información que se almacena cada vez que un usuario solicita el préstamo de un recurso de la biblioteca.

Actualmente la información se encuentra almacenada en una base de datos relacional de tipo Oracle. La dirección de la biblioteca ha solicitado ayuda al equipo docente de la asignatura de Arquitectura de bases de datos no tradicionales, y de acuerdo al tipo de consultas que les gustaría realizar, el equipo docente ha recomendado realizar una migración de la información a una base de datos de tipo documental.

Además, se ha acordado que los estudiantes de esta asignatura ayudarán a diseñar los tipos de documentos más eficientes para las siguientes dos consultas que desean realizar:

Consulta 1: Teniendo en cuenta la titulación en la que están matriculados los estudiantes, interesa conocer información acerca de los estudiantes que solicitaron un préstamo por cada mes y titulación. El objetivo es hacer un estudio de estacionalidad, por lo tanto, no consideraremos el año, sólo consideraremos los meses. Concretamente para cada mes y por cada titulación, se desea recuperar los datos (dni, nacionalidad, país donde vive) de los estudiantes de esa titulación, así como, por cada estudiante, la lista de recursos que tomaron en préstamo. De cada recurso se quiere obtener el título, isbn, tipo de recurso y asignatura de la titulación para la cual está recomendado el recurso prestado (se supone que todo los préstamos que solicitan los estudiantes están relacionados con alguna asignatura de la titulación que están cursando).

Consulta 2: También interesa conocer información de los recursos prestados agrupados por las asignaturas y titulación en las que están recomendados. Concretamente para cada titulación y por cada asignatura (una asignatura queda especificada en el plan de

estudios indicando su nombre y el curso en el que se imparte) que forman parte del plan de estudios de la misma, se desea recuperar los datos principales de cada recurso prestado: título, isbn, tipo de recurso y una lista con la información (dni, nacionalidad, país donde vive) de cada uno de los estudiantes que solicitaron el préstamo del recurso.

Cada consulta requiere un tipo concreto de documento. En este sentido, se pide indicar:

- Una representación gráfica del documento propuesto utilizando un diagrama de cajas anidadas como el que se explica en los apuntes sobre "Diseño de agregados" suministrado en el aula virtual. Como alternativa se puede presentar el documento en formato JSON pero se debe elegir entre una representación u otra.
- Una breve explicación de la estructura del documento y una justificación del porqué de su estructura.

## Ejercicio 4 (20%)

Las limitaciones que presenta el modelo relacional fue una de las razones que hicieron que surgieran las bases de datos NoSQL. Sin embargo, las limitaciones de los sistemas NoSQL también motivaron otras bases de datos denominadas NewSQL, que son sistemas que adoptan el modelo relacional para ofrecer algunas de sus ventajas junto con algunas de las mejoras que proporcionan las bases de datos NoSQL. Para saber más de las bases de datos NewSQL se propone leer los apartados 1, 2 y 3 del artículo titulado "What's Really New with NewSQL?" y los apartados 1, 3 y 4 del artículo titulado "NewSQL Through the Looking Glass".

Una vez leídos los artículos de referencia, se propone buscar un caso de aplicación de una base de datos NoSQL y otro caso de aplicación de una base de datos NewSQL. A continuación, contesta a las siguientes preguntas por cada caso de aplicación (1 página como máximo para cada caso):

1. Indica el enlace al caso analizado.
2. Describe el problema de persistencia de datos que se ha resuelto.
3. Justifica las razones por las que es recomendable la base de datos que se ha utilizado como solución.
4. Justifica las razones por las que no sería recomendable utilizar otro tipo de bases de datos diferentes a la utilizada en la solución.
5. Indica las referencias extra utilizadas para desarrollar el ejercicio.

No se puede reproducir de forma textual ninguna frase de las referencias utilizadas. A continuación, se muestran algunos ejemplos. **Estos ejemplos no pueden utilizarse para resolver el ejercicio (si se usan alguno de estos ejemplos la pregunta estará anulada), deben buscarse otros diferentes:**

- Uso de un [modelo en grafo para analizar los papeles de Panamá](#).
- Uso de un [modelo documental para ofrecer un servicio de contenidos personalizado y en tiempo real](#).
- Uso de TIDB [para análisis de datos en tiempo real](#).

## Criterios de valoración

Los apartados 1 y 2 tienen un peso del 30% cada uno, y los apartados 3 y 4 tienen un peso del 20% cada uno. Se valorará, para cada apartado, la validez de la solución y la claridad de la argumentación. Cualquier solución no justificada se considerará incompleta.

## Formato y fecha de entrega

Tenéis que enviar la PEC al buzón de Entrega y registro de EC disponible en el aula (apartado Evaluación). El formato del archivo que contiene vuestra solución puede ser .pdf, .odt, .doc y .docx. Para otras opciones, por favor, contactar previamente con vuestro profesor colaborador. El nombre del fichero debe contener el código de la asignatura, vuestro apellido y vuestro nombre, así como el número de actividad (PEC1). Por ejemplo nombreapellido1\_nosql\_pec1.docx. La fecha límite para entregar la PEC1 es el **17 de octubre**.

## Propiedad intelectual

Al presentar una práctica o PEC que haga uso de recursos ajenos, se tiene que presentar junto con ella un documento en que se detallen todos ellos, especificando el nombre de cada recurso, su autor, el lugar donde se obtuvo y su estatus legal: si la obra está protegida por el copyright o se acoge a alguna otra licencia de uso (Creative Commons, licencia GNU, GPL etc.). El estudiante tendrá que asegurarse que la licencia que sea no impide específicamente su uso en el marco de la práctica o PEC. En caso de no encontrar la información correspondiente tendrá que asumir que la obra está protegida por el copyright.

Será necesario, además, adjuntar los ficheros originales cuando las obras utilizadas sean digitales, y su código fuente, si así corresponde.