

Examen para SBD01.

INTENTO 1.

Pregunta 1.

Una de las siguientes no es una de las ventajas clave que pueden conseguirse mediante el uso de clusters de máquinas:

- a. Alta disponibilidad.
- b. **Menor consumo energético.**
- c. Alto rendimiento.
- d. Escalabilidad.

Pregunta 2.

Una de las siguientes afirmaciones es cierta en relación a procesamiento paralelo de tareas:

- a. **Algunas tareas no pueden paralelizarse para que se ejecuten más rápido usando varias máquinas.**
- b. Cualquier tarea siempre tiene al menos una parte que puede ejecutarse en paralelo para ganar tiempo.
- c. Todas las tareas pueden paralelizarse para que se ejecuten más rápido usando varias máquinas.
- d. Si conseguimos paralelizar una tarea, la potencia equivalente de cómputo es la suma de las potencias de cada nodo del cluster.

Pregunta 3.

La capa de colección en una arquitectura para Big Data es:

- a. Aquella de la que se obtienen los datos para visualización.
- b. **Aquella en la que se pasa de los datos que vienen de diversas fuentes a un conjunto de datos unificado y ya casi listo para ser usado.**
- c. Aquella en la que se gestiona el almacenamiento de los datos.
- d. Aquella en se determina para cada dato si va a ser utilizado o no para realizar analíticas.

Pregunta 4.

De entre las 5 Vs que definen el reto de Big Data, la V de velocidad tiene que ver con:

- a. La velocidad a la que se degrada la información.
- b. La alta velocidad que necesitan tener los microprocesadores con los que se tratan los datos.
- c. **La cantidad de datos que siguen generándose constantemente.**
- d. El bajo tiempo que transcurre desde que se genera el dato hasta que es adquirido por el sistema.

Pregunta 5.

Big Data suele traducirse al castellano como:

- a. Grandes Datos.
- b. Megadatos.
- c. **Macrodatos.**
- d. Gran Información.

Pregunta 6.

Las 4 características del principio ACID son:

- a. Atomicidad, Concurrencia, Aislamiento y Durabilidad.
- b. Atomicidad, Consistencia, Independencia y Durabilidad.
- c. **Atomicidad, Consistencia, Aislamiento y Durabilidad.**
- d. Aislamiento, Concurrencia, Independencia y Durabilidad.

Pregunta 7.

De entre las 5 Vs que definen el reto de Big Data, la V de volumen tiene que ver con:

- a. El elevado volumen de espacio necesario para ubicar las máquinas que conforman un clúster lo suficientemente potente.
- b. El hecho de que en cada posible fichero de vídeo o audio el sonido puede haberse guardado a diferente volumen.
- c. **La gran cantidad de bytes de datos que es necesario almacenar y tratar.**
- d. La complejidad distribuir una enorme cantidad de datos en los distintos volúmenes o unidades de almacenamiento.

Pregunta 8.

Una de las siguientes afirmaciones es cierta:

- a. Tanto procesamiento en paralelo como distribuido se realizan dentro de la misma máquina.
- b. El procesamiento distribuido se realiza dentro de la misma máquina y el paralelo en varias máquinas.
- c. Tanto procesamiento en paralelo como distribuido se realizan en varias máquinas.
- d. **El procesamiento en paralelo se realiza dentro de la misma máquina y el distribuido en varias máquinas.**

Pregunta 9.

Una de las siguientes no es una de las capas típicas de una arquitectura para Big Data:

- a. Capa de almacenamiento.
- b. Capa de visualización.
- c. **Capa de consolidación.**
- d. Capa de ingestión.

Pregunta 10.

Las tecnologías de Big Data se complementan con otras disciplinas como la Minería de Datos para obtener valor de los datos. ¿Verdadero o falso?

Seleccione una:

Verdadero

Falso

INTENTO 2.

Pregunta 1.

El acrónimo OLTP designa:

- a. Un sistema de ficheros distribuido en los nodos de un clúster.
- b. Un sistema que está orientado a analítica descriptiva.
- c. Un sistema que está orientado a generación de modelos predictivos.
- d. **Un sistema que está orientado a transacciones.**

Pregunta 2.

De entre las 5 Vs que definen el reto de Big Data, la V de valor tiene que ver con:

- a. El alto precio que hay que pagar por los datos.
- b. Cómo de útiles son los **datos para la institución**, empresa o persona que los usa.
- c. El valor que se puede obtener revendiendo datos a terceros.
- d. La valentía que implica el acometer este tipo de proyectos.

Pregunta 3.

Un procesamiento distribuido implica:

- a. Procesamiento transaccional.
- b. Distintos procesos para un mismo trabajo ejecutándose dentro del mismo procesador.
- c. Procesamiento para analítica.
- d. Distintos procesos para un mismo **trabajo ejecutándose** en distintas máquinas.

Pregunta 4.

La sabiduría se obtiene del conocimiento del mismo modo que los datos se obtienen de la información. ¿Verdadero o falso?

Seleccione una:

Verdadero

Falso

Pregunta 5.

De las siguientes, una no es una operación con datos que se consiga gracias a tecnologías Big Data:

- a. **Ser capaces de reproducir ficheros de audio o vídeo.**
- b. Aplicar técnicas de minería de datos para crear modelos predictivos con cantidades de datos que no caben en una única máquina.
- c. Almacenar los datos de un modo distribuido y replicado para conseguir una mayor disponibilidad de los mismos.
- d. Tratar los datos de forma distribuida empleando varias máquinas que trabajan en paralelo.

Pregunta 6.

Un dataset no debería contener ficheros de texto plano. ¿Verdadero o falso?

Seleccione una:

Verdadero

Falso

Pregunta 7.

Según el teorema CAP:

- a. **No podemos conseguir a la vez consistencia, disponibilidad y tolerancia al particionamiento.**
- b. No podemos conseguir a la vez consistencia, aislamiento y tolerancia a particionamiento.
- c. Podemos conseguir a la vez consistencia, disponibilidad y tolerancia al particionamiento.
- d. La base de datos no puede almacenar más información que la suma del espacio de almacenamiento de sus nodos.

Pregunta 8.

Un almacén de datos está pensado para poder realizar funciones transaccionales.

¿Verdadero o falso?

Seleccione una:

Verdadero

Falso

Pregunta 9.

De los siguientes, uno no está entre los aportes o beneficios que podemos obtener gracias a poder tratar con gran cantidad de datos:

- a. **Editar imágenes.**
- b. Dar soporte a la toma de decisiones.
- c. Realizar descubrimientos científicos.
- d. Identificar nuevos mercados.

Pregunta 10.

Una de los siguientes es el orden de etapas desde que ocurren los eventos hasta que de ellos se genera valor:

- a. **Eventos->Datos->Información->Conocimiento->Sabiduría->Valor**
- b. Eventos->Datos->Información->Sabiduría->Conocimiento->Valor
- c. Eventos->Información->Datos->Conocimiento->Sabiduría->Valor
- d. Eventos->Datos->Conocimiento->Información->Sabiduría->Valor

Intento 3

Pregunta 1

En una base de datos diseñada según el principio BASE:

- a. Se consigue atomicidad, consistencia, aislamiento y durabilidad.
- b. **El sistema es básicamente disponible, de estado blando y de consistencia eventual.**
- c. El sistema es básicamente disponible, de estado blando y siempre consistente.
- d. No podemos almacenar datos que incluyan algún tipo de compresión.

Pregunta 2

Los sistemas para Big Data de hoy en día son capaces de trabajar con cantidades de datos de:

- a. Como máximo 1 terabyte (TB).
- b. **Petabytes (PB) o incluso exabytes (EB).**
- c. Cualquier cantidad de datos que nos podamos imaginar.
- d. Como máximo 1 petabyte (PB).

Pregunta 3

De entre las 5 Vs que definen el reto de Big Data, la V de veracidad tiene que ver con:

- a. La dificultad en encontrar fuentes de información de pago que realmente sean fiables.
- b. **El hecho de que los datos no siempre cuentan con la calidad deseada o no son totalmente fieles a la realidad.**
- c. El hecho de que las fuentes Open Data, que son las más económicas, suelen contener información falsa.
- d. El hecho de que muchos de los post que encontramos en redes sociales contienen información falsa.

Pregunta 4

El lenguaje comúnmente empleado para interactuar con bases de datos se llama: SQL.

Pregunta 5

Sólo para una de las siguientes estrategias de procesamiento se necesita tener mucho tiempo disponible para procesar los datos: Por lotes.