

PREGUNTAS EXAMEN CC TEMA 5

Las respuestas que salen en cada pregunta son las **correctas**

1. **La tecnología blockchain sólo sirve, esencialmente, para implementar criptomonedas**
 - a. Falso
2. **¿Qué podemos aseverar sobre Hadoop?**
 - a. Sólo es aplicable para procesos en batch
 - b. Su velocidad es inferior a la de Spark en la mayoría de las aplicaciones
3. **Si el formato en que están representados los datos de mi problema (almacenados en HDFS) no es en pares (clave, valor) el problema no puede ser resuelto con MapReduce**
 - a. Falso
4. **El principal concepto innovador de Spark es el de Resilient Distributed Dataset (RDD), que no es más que una implementación distribuida del concepto de dataframe en R, pero que soporta exactamente las mismas operaciones en formato distribuido.**
 - a. Falso
5. **Escriba el nombre de la unidad de información que es mil millones de veces mayor que 1 TB (una única palabra)**
 - a. Zettabyte
6. **El programador de aplicaciones MapReduce**
 - a. Debe concentrar la creatividad en el diseño de las funciones Map y Reduce
7. **Señala los principales componentes de Spark:**
 - a. Spark SQL
 - b. Spark Streaming
 - c. MLLib
 - d. GraphX
 - e. Spark Core
8. **MapReduce es un paradigma de programación caracterizado por:**
 - a. Ser fácil de usar
 - b. Robusto frente a fallos de los nodos del clúster
9. **El enfoque tradicional para resolver problemas con cantidades ingentes de datos se basa en el uso de recursos “High Performance Computing”. ¿Cuál es la relación entre HPC y MapReduce?**
 - a. HPC siempre obtiene las soluciones más eficientes computacionalmente, pero el coste de desarrollo suele ser más elevado
 - b. Por su facilidad, MapReduce es accesible para un mayor número de programadores
10. **En su primera versión, Big Data se definía por tres Vs, correspondientes a:**
 - a. Volumen
 - b. Velocidad
 - c. Variedad

- 11. Big Data se refiere al estudio y diseño de sistemas y métodos para computación de conjuntos de datos cuyas propiedades suponen un reto para las capacidades de los computadores convencionales para analizar dichos datos y extraer valor no trivial de los mismos.**
- a. Falso
- 12. ¿Se puede implementar una aplicación blockchain sobre un servicio DaaS con un SGBD relacional?**
- a. Falso
- 13. Supongamos que disponemos de tantos recursos computacionales como se desee, entonces, ¿se debe acelerar el cálculo de un proceso MapReduce usando el mayor número posible de procesos Map?**
- a. No siempre, dado que podemos encontrarnos con particiones con un número de datos demasiado bajo (poco representativos)
- 14. ¿Cuál es el factor de aceleración (speedup) que, habitualmente, se obtiene en un algoritmo al codificarlo con Spark frente al mismo algoritmo en Hadoop?**
- a. 100
- 15. Seleccione las herramientas que forman parte del ecosistema Apache Hadoop**
- a. Pig
 - b. Hive
 - c. Hbase
 - d. ZooKeeper
- 16. A partir de la versión 2.0 Hadoop introdujo Yarn (Data Operating System) como una forma de mejorar el uso y rendimiento de HDFS**
- a. Falso
- 17. Seleccione los tipos de problemas que no sean (fácilmente) resolubles con Hadoop:**
- a. Procesos iterativos
 - b. Procesos interactivos
 - c. Procesamiento de grados
 - d. Ajedrez
- 18. De entre las técnicas clásicas de diseño de algoritmos, la más parecida en filosofía a MapReduce es la descomposición propia de la programación dinámica**
- a. Falso
- 19. Sobre Big Data podemos indicar:**
- a. Todas las demás respuestas son falsas. (Ha supuesto, fundamentalmente, una revolución en el contexto de los negocios, pero no de la tecnología; sólo contempla el procesamiento de datos con un volumen masivamente grande; sólo se aplica a métodos de analítica de datos y extracción de conocimiento)
- 20. Hadoop File System (HDFS)**
- a. Es robusto y puede usarse con múltiples marcos de trabajo MapReduce
- 21. ¿Se puede desplegar una infraestructura para Big Data usando contenedores?**
- a. Sí
- 22. En relación con MapReduce:**
- a. No es aplicable a todo tipo de problemas.

- b. Es aplicable a todos los conjuntos de datos que puedan representarse como pares <clave, valor>
- 23. La principal ventaja de Spark sobre Hadoop fue el cambio de lenguaje de programación elegido para su implementación**
- a. Falso
- 24. ¿Cuáles son las principales interfaces para usar HDFS?**
- a. Interfaz web, puerto 50070
 - b. Línea de órdenes: mandatos, Hadoop o HDFS
 - c. API de programación
- 25. La tecnología blockchain es tan versátil que permite reordenar los bloques de las cadenas por distintos criterios.**
- a. Falso
- 26. Seleccione de la siguiente lista todas las características asociadas a la tecnología de bloques de cadenas.**
- a. Inmutabilidad
 - b. Ofrece garantías en la privacidad
 - c. Elimina intermediarios
 - d. Es un sistema abierto y transparente
- 27. Un contrato inteligente es un modelo de relación comercial entre empresas soportado por un sistema de Inteligencia Artificial.**
- a. Falso
- 28. Para mantener la seguridad completa de la información en una cadena de bloques, esta debe ser completamente pública:**
- a. Falso
- 29. ¿En qué lenguaje de programación está implementado Spark?**
- a. Scala