

Ingeniería de Big Data Avanzada

Código: ARC-108

Propuesta de Valor: ARCITURA

Duración: 24 Horas



Este curso se basa en el Módulo 7 al explorar temas de ingeniería avanzada relacionados principalmente al almacenamiento y procesamiento de conjuntos de datos de Big Data.

Específicamente, cubre Big Data avanzado mecanismos de ingeniería, almacenamiento de datos en memoria y procesamiento de datos en tiempo real.

El curso presenta más consideraciones para la construcción de algoritmos MapReduce y también presenta el motor de procesamiento Bulk Synchronous Parallel (BSP), junto con una discusión de procesamiento de datos gráficos.

Los mecanismos de Big Data necesarios para desarrollar pipelines de Big Data, también se exploran las etapas y el proceso de diseño involucrados en la construcción de soluciones de procesamiento de Big Data.



AUDIENCIA

- Cualquier persona que esté interesada en ampliar sus conocimientos en Big Data y deseen cubrir unas necesidades no satisfechas por las tecnologías existentes en el almacenamiento y tratamiento de grandes volúmenes de datos.



PRE REQUISITOS

- No hay requisitos previos.



OBJETIVOS

- Mecanismos avanzados de ingeniería de Big Data.
- Motores de serialización y compresión
- Dispositivos de almacenamiento en memoria
- Cuadrículas de datos en memoria y bases de datos en memoria
- Enfoques de integración de lectura continua, lectura anticipada, escritura simultánea y escritura diferida
- Persistencia polígota
- Explicación, problemas y recomendaciones
- Procesamiento de Big Data en tiempo real
- Volumen de consistencia de velocidad (SCV)
- Procesamiento de flujo de eventos (ESP)
- Procesamiento de eventos complejos (CEP)

- El principio SCV
- Procesamiento general de Big Data en tiempo real y MapReduce
- Diseños avanzados de algoritmos MapReduce
- Motor de procesamiento masivo síncrono paralelo (BSP)
- BSP vs. MapReduce
- BSP síncrono paralelo
- Datos gráficos y procesamiento de datos gráficos usando BSP (superpasos)
- Canalizaciones de Big Data, incluida la definición y las etapas
- Big Data con Extract-Load-Transform (ELT)
- Características de la solución de Big Data, consideraciones de diseño y proceso de diseño. Duración: 1 día



CERTIFICACIÓN DISPONIBLE

- Certificado oficial **ARCITURA**.
- Este curso es parte de la ruta de certificaciones de:
 - Consultor de Big Data Certificado



CONTENIDO

1. MECANISMOS AVANZADOS DE INGENIERÍA DE BIG DATA
2. MOTORES DE SERIALIZACIÓN Y COMPRESIÓN
3. DISPOSITIVOS DE ALMACENAMIENTO EN MEMORIA
4. CUADRÍCULAS DE DATOS EN MEMORIA Y BASES DE DATOS EN MEMORIA
5. LECTURA COMPLETA, LECTURA ANTICIPADA, ESCRITURA DIRECTA Y ESCRITURA DETRÁS DE LOS ENFOQUES DE INTEGRACIÓN
6. PERSISTENCIA POLÍGLOTA
7. EXPLICACIÓN, PROBLEMAS Y RECOMENDACIONES
8. PROCESAMIENTO DE BIG DATA EN TIEMPO REAL
9. VOLUMEN DE COHERENCIA DE VELOCIDAD (SCV)
10. PROCESAMIENTO DE FLUJO DE EVENTOS (ESP)
11. PROCESAMIENTO DE EVENTOS COMPLEJOS (CEP)
12. EL PRINCIPIO SCV
13. PROCESAMIENTO GENERAL DE BIG DATA EN TIEMPO REAL Y MAPREDUCE

14. DISEÑOS AVANZADOS DE ALGORITMO MAPREDUCE
15. MOTOR DE PROCESAMIENTO PARALELO SINCRÓNICO (BSP) A GRANEL
16. BSP VS. MAPREDUCE
17. BSP SYNCHRONOUS PARALLEL
18. GRAPH DATA Y GRAPH DATA PROCESSING USANDO BSP (SUPERSTEPS)
19. BIG DATA PIPELINES, INCLUIDAS LA DEFINICIÓN Y LAS ETAPAS
20. BIG DATA CON EXTRACCIÓN, CARGA Y TRANSFORMACIÓN (ELT)
21. CARACTERÍSTICAS DE LA SOLUCIÓN DE BIG DATA, CONSIDERACIONES DE DISEÑO Y PROCESO DE DISEÑO



BENEFICIOS

- Al finalizar este curso tendrás los conocimientos en el almacenamiento y procesamiento de conjuntos de datos de Big Data.