

Evaluación Sumativa II:

Unidad 2: Arquitectura de Big Data.

Implementación de una arquitectura de Big Data 30%.

Objetivo General

El estudiante debe diseñar e implementar conceptualmente una solución de Big Data basada en un caso real o simulado, utilizando arquitectura de alta disponibilidad, herramientas en Cloud, modelos analíticos y planificación del proyecto.

Criterios de Evaluación:

- 2.1.1. Diseña una arquitectura de alta disponibilidad para una solución de Big Data.
- 2.1.2. Aplica herramientas en Cloud que permitan utilizar un ambiente de Big Data.
- 2.1.3. Aplica modelos descriptivos y predictivos, sobre ambiente de Big Data.
- 2.1.4. Planifica las tareas de acuerdo al tiempo, objetivos y características del equipo para el logro de metas establecidas.

Formato de entrega:

Modalidad: Individual

Extensión máxima: 10 páginas

Formato: PDF o Word

Letra recomendada: Arial o Calibri 11

Formato de entrega archivo: Apellido_Nombre_BigData.pdf

Interlineado: 1.5

ESTRUCTURA DEL INFORME

1. Arquitectura de Alta Disponibilidad

Objetivo: Diseñar una arquitectura Big Data resiliente y escalable. Esto es en base al tema escogido durante la clase, para luego ir desarrollando en el informe.

Debe incluir:

- **Descripción general:** ¿Para qué tipo de datos o escenario se implementa la arquitectura?
- **Componentes mínimos:**
 - Capa de ingesta (ej: Kafka, NiFi, Flume).
 - Almacenamiento distribuido (HDFS, S3, Lakehouse).
 - Procesamiento (Spark, MapReduce, Databricks).
 - Capa de análisis o consulta (Hive, Presto, Impala).
 - Debe incluir el total de capas para considerarse el punto correcto.
- **Alta disponibilidad:**
 - Replicación de datos.
 - Balanceo de carga.
 - Tolerancia a fallos (clustering o redundancia).
- **Diagrama simple** (2D o en bloques).

2. Herramientas Cloud para Big Data

Objetivo: Aplicar tecnologías Cloud para soportar la arquitectura.

Debe incluir:

- **Proveedor seleccionado:** AWS, Azure, GCP u otro.
- **Servicios utilizados:** Ejemplos posibles:
 - AWS: EMR, S3, Glue, Redshift, Lambda.
 - Azure: Synapse, HDInsight, Data Lake, Databricks.
 - GCP: Dataproc, BigQuery, Storage, Composer.
- **Justificación técnica:**
Por qué se eligen esos servicios (costos, escalabilidad, facilidad, rendimiento).

- **Esquema o tabla de recursos configurados.**

3. Modelos Descriptivos y Predictivos

Objetivo: Explicar cómo aplicar modelos analíticos dentro de la arquitectura.

Debe incluir:

- **Modelo descriptivo:**
 - Tipo (ej: clustering, dashboards, análisis histórico).
 - Caso de uso (ej: segmentación de usuarios, reporte de tendencias).
- **Modelo predictivo:**
 - Tipo (ej: regresión, árboles de decisión, forecasting).
 - Caso de uso (ej: predicción de demanda, detección de anomalías).
- **Herramientas:** Spark MLlib, TensorFlow, BigQuery ML, Databricks, etc.
- No se exige implementación real, pero sí un ejemplo o flujo.

4. Planificación del Proyecto

Objetivo: Organizar las actividades necesarias para implementar la solución.

Debe incluir:

- **Cronograma general:**
Diagrama de Gantt, tabla o fases con fechas estimadas.
- **Objetivos por etapa:**
Ejemplo: Diseño, configuración, análisis, pruebas.
- **Roles hipotéticos del equipo:**
Aunque el trabajo es individual, se debe imaginar un equipo (ej: analista de datos, ingeniero cloud, científico de datos).
- **Gestión del tiempo y riesgos:**
Ej: plazos críticos, dependencias, recursos.

5. Conclusiones y cierre

Incluye:

- Síntesis del proyecto.
- Beneficios de la arquitectura propuesta.
- Posibles mejoras o desafíos.

6. Bibliografía o fuentes

Formato libre (norma APA).

IMPORTANTE:

La evaluación se calificará con una escala de apreciación de 120 puntos basada en los niveles: Muy Bueno, Bueno, Logrado, Parcialmente Logrado y No Logrado, con exigencia del 60%.

La entrega es individual

Nota 7: 120 puntos

Nota 4: 71 puntos

Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota
0.0	1.0	10.0	1.4	20.0	1.8	30.0	2.3	40.0	2.7	50.0	3.1	60.0	3.5
1.0	1.0	11.0	1.5	21.0	1.9	31.0	2.3	41.0	2.7	51.0	3.1	61.0	3.5
2.0	1.1	12.0	1.5	22.0	1.9	32.0	2.3	42.0	2.8	52.0	3.2	62.0	3.6
3.0	1.1	13.0	1.5	23.0	2.0	33.0	2.4	43.0	2.8	53.0	3.2	63.0	3.6
4.0	1.2	14.0	1.6	24.0	2.0	34.0	2.4	44.0	2.8	54.0	3.3	64.0	3.7
5.0	1.2	15.0	1.6	25.0	2.0	35.0	2.5	45.0	2.9	55.0	3.3	65.0	3.7
6.0	1.3	16.0	1.7	26.0	2.1	36.0	2.5	46.0	2.9	56.0	3.3	66.0	3.8
7.0	1.3	17.0	1.7	27.0	2.1	37.0	2.5	47.0	3.0	57.0	3.4	67.0	3.8
8.0	1.3	18.0	1.8	28.0	2.2	38.0	2.6	48.0	3.0	58.0	3.4	68.0	3.8
9.0	1.4	19.0	1.8	29.0	2.2	39.0	2.6	49.0	3.0	59.0	3.5	69.0	3.9
Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota	Puntaje	Nota		
70.0	3.9	80.0	4.5	90.0	5.1	100.0	5.8	110.0	6.4	120.0	7.0		
71.0	4.0	81.0	4.6	91.0	5.2	101.0	5.8	111.0	6.4				
72.0	4.0	82.0	4.6	92.0	5.3	102.0	5.9	112.0	6.5				
73.0	4.1	83.0	4.7	93.0	5.3	103.0	5.9	113.0	6.6				
74.0	4.1	84.0	4.8	94.0	5.4	104.0	6.0	114.0	6.6				
75.0	4.2	85.0	4.8	95.0	5.4	105.0	6.1	115.0	6.7				
76.0	4.3	86.0	4.9	96.0	5.5	106.0	6.1	116.0	6.8				
77.0	4.3	87.0	4.9	97.0	5.6	107.0	6.2	117.0	6.8				
78.0	4.4	88.0	5.0	98.0	5.6	108.0	6.3	118.0	6.9				
79.0	4.4	89.0	5.1	99.0	5.7	109.0	6.3	119.0	6.9				