ARQUITECTURAS BIG DATA, APACHE SPARK Y KAFKA

Actividad de transferencia

**Definiendo Tecnologías**

En los proyectos de Big Data y fundamentalmente en aquellos proyectos que son de Big Data Streaming, es decir, de tiempo casi-real, hemos visto que existen arquitecturas por un lado y tecnologías por otro que permiten de alguna forma dar respuesta eficiente a estos proyectos.

Entre las arquitecturas vimos la arquitectura Lambda y la Arquitectura Kappa. Entre las tecnologías, vimos Apache Spark y Apache Kafka. No obstante, es frecuente que en los proyectos de big data se combinen diversas tecnologías.

La actividad de esta semana consiste en armar un **cuadro comparativo**, para que puedas conocer otras tecnologías también muy utilizadas y, sobre todo, para revisar algunas diferencias sutiles que hay entre ellas.

Vamos a comparar las siguientes tecnologías:

* Apache Storm
* Apache Spark
* Apache Samza

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Aplicación** | **¿Para qué se utiliza?** | **Algunas ventajas** | **Proyectos ideales para implementar esta tecnología** | **Ejemplo de empresa que lo utiliza** |
| **Apache Storm** | (Debes completar esta celda con la síntesis de lo que hayas encontrado) | (Debes completar esta celda con la síntesis de lo que hayas encontrado) | (Debes completar esta celda con la síntesis de lo que hayas encontrado) | Yahoo, Netflix, etc…(lo que encuentres) |
| **Apache Spark** | …. |  |  |  |
| **Apache Samza** | ….. |  |  |  |

**Fuente de datos consultada:**

* Aquí deberás colocar en bullets, los sitios de los cuales has obtenido la información

La entrega será un documento WORD o PDF con el cuadro comparativo. Debes de estar lo más completo que puedas y, sobre todo, debes colocar las referencias de dónde has obtenido la información.

NOTA: Recuerda utilizar recursos como Google Académico, para acceder a papers o documentos académicos que puedan tener validez para educación.