**Tec de Monterrey Maestría en Inteligencia Artificial Aplicada**

**Materia: Análisis de grandes volúmenes de datos Clave: TC4034 Nombre: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_ Matrícula: \_\_\_\_\_\_\_\_\_\_\_**

**Título del artículo: BSP vs MapReduce**

**INSTRUCCIONES:**

Lea el artículo científico indicado, complete una copia de este documento y envíela como evidencia de su actividad.

1. **Contenido de la investigación**

¿Qué tipo de contribución intentó hacer el artículo? A continuación, se enumeran varios tipos posibles de contribución. Marque cada uno de los que aplican y justifique su selección utilizando la sección de comentarios.

\_\_\_ 1. Describe una nueva técnica.

\_\_\_\_ 2. Se extiende en mejora una técnica existente.

\_\_\_ 3. Establece propiedades de una técnica o las relaciones entre dos o más empírica o teóricamente.

\_\_\_ 4. Describe una nueva aplicación de una técnica.

\_\_\_ 5. Prueba la validez psicológica de una técnica.

\_\_x\_\_ 6. Combina varias técnicas en un sistema.

\_\_\_ 7. Identifica y propone un nuevo problema.

\_\_\_ 8. Puede ser considerado como un tutorial (por ejemplo un survey).

\_\_\_ 9. Otro (especifique).

Comentarios:

La idea fundamental del trabajo es el discutir la evidente relación entre MapReduce y BSP, ambos soportando algoritmos paralelos de una manera general, intercalando fases y comunicaciones. Ambos algoritmos pueden ser usados para correr en computadoras con bajos recursos, conectados mediante conexiones punto a punto, y ambos pueden usarse para realizar operaciones (pasos/superpasos), presentando dos modelos.

1. Modelo MapReduce en BSP: Este modelo permite la ejecución de algoritmos enfocados par MR, dentro de un entorno BSP, procurando conservar los costes asintóticos. Con las desventajas que MR no permite el almacenamiento de datos entre rondas, requiriendo que los datos si se comuniquen de esa manera, esto implica el aumento en costes de comunicación.

2. Modelo BSP en MapReduce: Este modelo da una descripción algoritmos para simular BSP en el framework MapReduce (conociéndose los tiempos de ejecución de las tareas individuales), mediante el siguiente procedimiento:

* La función MAP se establece como función identidad.
* Cada tarea reduce simula un procesador BSP y con cada etapa reduce simula un solo superstep.
* Cata tarea reduce recibe el input del procesador BSP que se simule, realizando los cálculos y salidas para las tareas reduce
* Cualquier dato en BSP que se almacene en memoria local entre superpasos se almacena en memoria global y se lee en la siguiente ronda

1. **Importancia de la investigación**

Evaluar la importancia de la investigación informada respondiendo las siguientes preguntas. Justifique sus respuestas.

* 1. ¿Qué hipótesis son probadas por la investigación reportada?
  2. ¿Cuáles son los resultados de esta prueba de hipótesis?
  3. Evaluar la importancia de la investigación reportada en una escala de 1 (trivial) a 5 (avance importante).

1. **Corrección e integridad del artículo.**

Evalúe la corrección e integridad del documento vinculando cuál de los siguientes se aplica. Justifique sus respuestas en la sección de comentarios.

\_\_\_ 1. La metodología empleada por los autores no es sensata.

\_\_\_ 2. Hay grandes errores técnicos en el documento.

\_\_\_ 3. Hay omisiones importantes en el documento.

\_\_\_ 4. La investigación informada no está en un estado completo.

\_\_\_ 5. La investigación relacionada no se compara adecuadamente.

\_\_\_ 6. El documento es deficiente de alguna otra manera (especifique).

**Comentarios:**

4. Resumen Describe el mensaje del documento en tus propias palabras. Use 100-200 palabras.