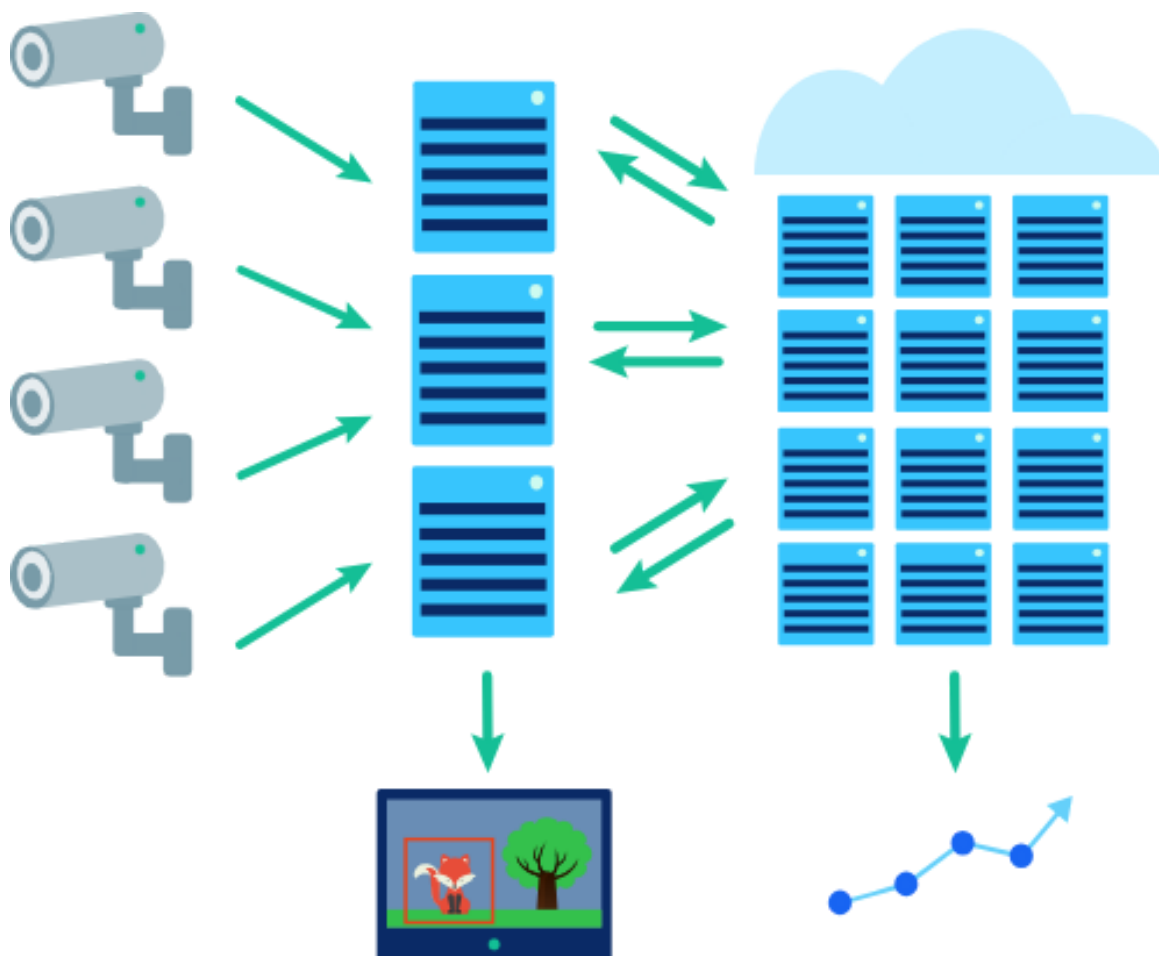


## Computación Distribuida y Paralela



Juan Carlos Cox Fernández  
Adrian Yared Armas de la Nuez

## Contenido

---

<b>1. Definición.....</b>	<b>2</b>
1.1 Tipos de computación paralela.....	2
1.2 Tipos de computación distribuida.....	2
<b>2. Ventajas y desventajas.....</b>	<b>3</b>
2.1 Ventajas.....	3
Mayor Velocidad y Eficiencia.....	3
Escalabilidad.....	3
Eficiencia en el Manejo de Grandes Volúmenes de Datos.....	3
Tolerancia a Fallos.....	3
Ahorro Energético en Algunas Aplicaciones.....	3
2.2 Desventajas.....	4
Complejidad en la Implementación.....	4
Problemas de Comunicación y Latencia.....	4
Mayor Consumo de Recursos Iniciales.....	4
Problemas de Consistencia y Sincronización.....	4
2.3 Tabla comparativa de ventajas y desventajas.....	4
<b>3. Artículos.....</b>	<b>5</b>
3.1 Primer artículo.....	5
3.1.1 Resumen.....	5
3.1.2 Argumentos.....	5
3.1.3 Casos de uso.....	6
3.2 Segundo artículo.....	7
3.2.1 Resumen.....	7
3.2.2 Argumentos.....	7
3.2.3 Casos de uso.....	7
3.3 Tercer artículo.....	8
3.3.1 Resumen.....	8
3.3.2 Definición Hopfield.....	8
3.3.2.1 Paralela.....	8
3.3.2.2 Distribuida.....	8
3.3.3 Argumentos.....	9
3.3.4 Casos de uso.....	9
3.4 Comparativa.....	9
3.4.1 BigData.....	9
3.4.2 IA y Deep Learning.....	10
3.4.3 Ámbito científico y análisis.....	10



## Computación Distribuida y Paralela

<b>3.4.4 Industria e investigación.....</b>	<b>11</b>
<b>3.4.4.1 Aplicaciones en la Investigación.....</b>	<b>11</b>
<b>3.4.4.2 Aplicaciones en la Industria.....</b>	<b>11</b>
<b>4. Bibliografía.....</b>	<b>11</b>

## 1. Definición

La computación paralela y distribuida ha crecido debido a la necesidad de resolver problemas complejos y los avances tecnológicos. Su desarrollo responde a aplicaciones como ecuaciones diferenciales parciales (PDEs), dinámica de fluidos, procesamiento de imágenes y optimización de sistemas interconectados. La computación distribuida también se emplea en redes de sensores y comunicación, donde la descentralización y la confiabilidad son clave.

Los sistemas paralelos agrupan varios procesadores en un mismo entorno con comunicación rápida y estable, trabajando juntos en una tarea específica. En cambio, los sistemas distribuidos conectan procesadores geográficamente separados mediante redes, donde la comunicación puede ser más lenta o menos confiable debido a fallos o variaciones en los enlaces.

### 1.1 Tipos de computación paralela

Existen diferentes tipos de computaciones paralelas:

- **Masivamente paralelas:** Se refiere a sistemas con miles de procesadores, típicos en supercomputación.
- **Control centralizado o descentralizado:** Desde instrucciones comunes hasta autonomía total. Como el caso del map Reduce de hadoop.
- **Sincronización síncrona o asíncrona:** La primera facilita el control, mientras que la segunda es más flexible. Un ejemplo de asíncrona podría ser cuda de nvidia y síncrona OpenMP.
- **Interconexión de procesadores:**
  - **Memoria compartida** (varios hilos modifican una sola región de memoria), como el caso de OpenMP
  - **Paso de mensajes** (los procesos envían mensajes), como el caso de los mpi (message passing interface)
  - **Modelos híbridos.** Aportan flexibilidad (Comunicación entre nodos).

### 1.2 Tipos de computación distribuida

Los principales tipos de computación distribuida son:

- **Clúster:** Grupo de ordenadores trabajando como uno solo.
- **Grid:** Ordenadores distribuido globalmente trabajando en tareas grandes, como el caso de setihome.
- **Nube:** Recursos remotos accesibles por internet como Google Cloud.
- **P2P:** Nodos sin servidor central, como BitTorrent.
- **Niebla:** Procesa datos cerca del usuario, por ejemplo, el 5G.
- **Borde:** Computación local en dispositivos inteligentes (coches autónomos).
- 

## 2. Ventajas y desventajas

La computación distribuida y paralela se utilizan para mejorar el rendimiento de las aplicaciones al dividir tareas entre múltiples procesadores o nodos.

### 2.1 Ventajas

#### Mayor Velocidad y Eficiencia

- **Distribución de carga:** Permite dividir el problema en varias tareas ejecutadas simultáneamente, reduciendo el tiempo total de ejecución en comparación con el procesamiento secuencial.
- **Mejor uso del hardware:** Se aprovechan múltiples núcleos o máquinas para ejecutar tareas más rápido.

#### Escalabilidad

- En sistemas distribuidos, se pueden agregar más nodos para aumentar la capacidad de cómputo sin necesidad de reemplazar el hardware existente.
- La computación paralela permite optimizar el rendimiento al utilizar múltiples procesadores en una sola máquina.

#### Eficiencia en el Manejo de Grandes Volúmenes de Datos

- Se pueden procesar grandes conjuntos de datos de manera más eficiente (ejemplo: Big Data con Hadoop y Spark).
- Útil en aprendizaje automático, simulaciones científicas y modelado de sistemas complejos.
-

### Tolerancia a Fallos

- Los sistemas distribuidos pueden continuar funcionando incluso si un nodo falla, mejorando la fiabilidad.
- Algunas arquitecturas paralelas implementan redundancia para evitar fallos críticos.

### Ahorro Energético en Algunas Aplicaciones

- En entornos distribuidos, las tareas pueden ser asignadas dinámicamente a nodos con menor consumo energético.
- En comparación con un solo procesador de alto rendimiento, distribuir la carga en varios núcleos pequeños puede reducir el consumo total de energía.

## 2.2 Desventajas

### Complejidad en la Implementación

- Requiere algoritmos específicos para la distribución y sincronización de tareas.
- Es más difícil de depurar y mantener debido a la concurrencia y comunicación entre procesos.

### Problemas de Comunicación y Latencia

- En sistemas distribuidos, la comunicación entre nodos introduce latencia y sobrecarga en la red.
- La sincronización de tareas en la computación paralela puede generar cuellos de botella.

### Mayor Consumo de Recursos Iniciales

- Se necesita hardware más potente o una red de ordenadores, lo que puede aumentar costos.
- Los algoritmos deben ser diseñados o adaptados para aprovechar el paralelismo, lo que requiere más tiempo de desarrollo.

### Problemas de Consistencia y Sincronización

- En sistemas distribuidos, garantizar la consistencia de los datos puede ser un desafío (ejemplo: bases de datos distribuidas).
- En paralelismo, el acceso concurrente a memoria compartida puede generar condiciones de carrera o bloqueos.

### 2.3 Tabla comparativa de ventajas y desventajas

Criterio	Computación Secuencial	Computación Paralela	Computación Distribuida
<b>Velocidad</b>	Lenta, ya que solo usa un núcleo.	Más rápida en hardware con múltiples núcleos.	Depende de la red y los nodos disponibles.
<b>Escalabilidad</b>	Limitada por la velocidad de un solo procesador.	Buena en hardware con múltiples núcleos.	Alta, se pueden agregar más nodos.
<b>Tolerancia a fallos</b>	Si el sistema falla, la ejecución se detiene.	Puede tener redundancia en algunos casos.	Alta, ya que otros nodos pueden continuar la ejecución.
<b>Complejidad de programación</b>	Simple, lógica secuencial.	Moderada, requiere manejo de hilos y concurrencia.	Alta, implica coordinación entre múltiples sistemas.
<b>Consumo de energía</b>	Puede ser alto si se usa un procesador potente.	Puede ser más eficiente en algunos casos.	Variable, depende de la distribución de carga.
<b>Latencia</b>	Baja, ya que no hay comunicación entre procesos.	Puede haber latencia si los hilos compiten por recursos.	Alta, debido a la comunicación entre nodos.

## 3. Artículos

### 3.1 Primer artículo

Link del artículo:

[https://fastercapital.com/es/contenido/Computacion-distribuida-y-paralela--creacion-de-una-ventaja-competitiva--el-papel-de-la-computacion-distribuida-y-paralela-en-las-empresas.html?utm\\_source=chatgpt.com#Introducci-n-a-la-computaci-n-distribuida-y-paralela](https://fastercapital.com/es/contenido/Computacion-distribuida-y-paralela--creacion-de-una-ventaja-competitiva--el-papel-de-la-computacion-distribuida-y-paralela-en-las-empresas.html?utm_source=chatgpt.com#Introducci-n-a-la-computaci-n-distribuida-y-paralela)

#### 3.1.1 Resumen

El artículo "Computación distribuida y paralela: creación de una ventaja competitiva" explora cómo estas arquitecturas informáticas se han convertido en componentes esenciales para el desarrollo estratégico de las empresas en diversas industrias. Destaca que la integración de sistemas distribuidos y paralelos no solo potencia la capacidad de procesamiento, sino que también promueve un uso eficiente de los datos y recursos, impulsando a las empresas hacia un crecimiento significativo.

### 3.1.2 Argumentos

1. **Escalabilidad y flexibilidad:** Permiten a las empresas ampliar sus recursos computacionales en consonancia con su crecimiento, adaptándose a demandas variables sin comprometer el rendimiento.
2. **Rentabilidad:** Al optimizar el uso de recursos y reducir la necesidad de hardware especializado, estas arquitecturas contribuyen a una disminución en los costos operativos.
3. **Procesamiento de datos mejorado:** Facilitan el manejo eficiente de grandes volúmenes de datos, permitiendo análisis en tiempo real y una toma de decisiones más informada.
4. **Aceleración de la innovación:** Al reducir los tiempos de procesamiento, las empresas pueden desarrollar y lanzar nuevos productos y servicios más rápidamente.
5. **Confiabilidad y redundancia:** La naturaleza distribuida de estos sistemas ofrece mayor tolerancia a fallos, asegurando continuidad en las operaciones incluso ante fallas individuales de componentes.
6. **Colaboración global:** Facilitan el trabajo conjunto entre equipos ubicados en diferentes regiones geográficas, mejorando la eficiencia y cohesión en proyectos internacionales.

### 3.1.3 Casos de uso

- **Gigante minorista durante temporadas pico:** Implementa un sistema de bases de datos distribuidas para gestionar millones de transacciones simultáneamente sin afectar el rendimiento.
- **Empresa de análisis financiero:** Utiliza algoritmos paralelos para ejecutar simulaciones y evaluaciones de riesgos de manera más eficiente, reduciendo tiempo y costos energéticos.
- **Empresa de logística:** Emplea computación distribuida para procesar y analizar grandes conjuntos de datos en tiempo real, optimizando rutas de entrega basadas en patrones de tráfico y ubicaciones de paquetes.
- **Empresa de biotecnología:** Aplica procesamiento paralelo para acelerar la secuenciación genómica, reduciendo el tiempo necesario para la investigación médica y el descubrimiento de nuevos medicamentos.
- **Proveedor de almacenamiento en la nube:** Se beneficia de la confiabilidad de los sistemas distribuidos, asegurando disponibilidad continua de datos incluso ante fallos de algunos nodos.
- **Equipo de diseño multinacional:** Colabora en proyectos de simulación compartidos, permitiendo contribuciones simultáneas desde diferentes zonas horarias gracias a la infraestructura distribuida.



## 3.2 Segundo artículo

Link del artículo:

[https://books.google.es/books?id=n\\_Q5EAAAQBAJ&printsec=frontcover&hl=es#v=onepage&q&f=false](https://books.google.es/books?id=n_Q5EAAAQBAJ&printsec=frontcover&hl=es#v=onepage&q&f=false)

### 3.2.1 Resumen

El libro Parallel and Distributed Computation: Numerical Methods es una referencia clave en computación paralela y distribuida, centrada en métodos numéricos. Analiza algoritmos para estos sistemas, considerando su teoría y aplicaciones. Examina la convergencia, estabilidad y escalabilidad de los métodos, además de estrategias de sincronización y comunicación entre nodos, esenciales para su implementación.

### 3.2.2 Argumentos

1. **Eficiencia computacional:** La implementación de algoritmos en entornos paralelos y distribuidos permite una reducción significativa en los tiempos de ejecución, especialmente en problemas de gran escala.
2. **Escalabilidad:** Estos enfoques facilitan la resolución de problemas que superan la capacidad de un único procesador, distribuyendo la carga de trabajo entre múltiples unidades de procesamiento.
3. **Flexibilidad en la resolución de problemas:** La arquitectura distribuida ofrece diversas estrategias para abordar problemas complejos, adaptándose a diferentes estructuras y requisitos computacionales.
4. **Robustez y tolerancia a fallos:** Los sistemas distribuidos pueden continuar operando adecuadamente incluso si algunos componentes fallan, mejorando la confiabilidad general del sistema.

### 3.2.3 Casos de uso

- **Solución de sistemas de ecuaciones lineales:** Se presentan métodos paralelos para resolver grandes sistemas de ecuaciones, como los algoritmos de descomposición y técnicas iterativas, que son esenciales en simulaciones científicas y modelado.
- **Optimización numérica:** El libro aborda algoritmos distribuidos para problemas de optimización, aplicables en áreas como la economía, ingeniería y aprendizaje automático, donde es crucial manejar grandes conjuntos de datos y funciones complejas.
- **Procesamiento de señales e imágenes:** Se exploran técnicas paralelas para el filtrado y análisis de señales, mejorando la eficiencia en aplicaciones como el procesamiento de imágenes médicas y reconocimiento de patrones.

- **Simulaciones en dinámica de fluidos:** Utiliza métodos numéricos paralelos para modelar el comportamiento de fluidos, relevante en la ingeniería aeroespacial y estudios meteorológicos, donde se requiere un procesamiento intensivo de datos.
- **Análisis de grandes redes:** Implementa algoritmos distribuidos para examinar propiedades de redes extensas, como las sociales o de comunicación, permitiendo un análisis más rápido y eficiente de su estructura y dinámica.

### 3.3 Tercer artículo

Link del artículo:

<https://dl.acm.org/doi/abs/10.5555/39096#sec-recommendations>

#### 3.3.1 Resumen

El artículo "A Graph-Based Recommender System for Digital Library", aborda la mejora de los sistemas de recomendación en bibliotecas digitales mediante un enfoque híbrido que combina técnicas basadas en contenido y colaborativas. Los autores proponen un sistema de recomendación basado en grafos que utiliza un algoritmo de red de Hopfield para explotar asociaciones de alto grado entre libros y usuarios, con el objetivo de mejorar la precisión y el recall de las recomendaciones.

#### 3.3.2 Definición Hopfield

Las redes de Hopfield son un tipo de red neuronal recurrente utilizada para la memoria asociativa y la optimización combinatoria.

##### 3.3.2.1 Paralela

Dado que cada neurona en la red de Hopfield se actualiza en función de la actividad de otras neuronas, existen dos enfoques de actualización:

- **Asíncrono:** Se actualiza una neurona a la vez de manera secuencial.
- **Síncrono (paralelo):** Todas las neuronas se actualizan simultáneamente.

##### 3.3.2.2 Distribuida

Las redes de Hopfield pueden dividirse en subredes y distribuirse en múltiples nodos de cómputo:

- **Descomposición espacial:** Dividir la red en regiones y asignarlas a distintos procesadores
- **Descomposición de datos:** Cada nodo almacena una parte de las conexiones sinápticas y colabora en la actualización de los estados.

### 3.3.3 Argumentos

1. **Mejora de la calidad de las recomendaciones:** Al combinar enfoques basados en contenido y colaborativos, el sistema puede proporcionar recomendaciones más precisas y relevantes para los usuarios.
2. **Explotación de asociaciones complejas:** El uso de un algoritmo de red de Hopfield permite identificar y aprovechar relaciones de alto grado entre libros y usuarios, enriqueciendo el proceso de recomendación.
3. **Evaluación empírica positiva:** Pruebas preliminares indican que el sistema propuesto mejora tanto la precisión como el recall en comparación con métodos tradicionales de recomendación.

### 3.3.4 Casos de uso

**Bibliotecas digitales académicas:** Implementación del sistema para recomendar artículos y libros relevantes a investigadores y estudiantes, facilitando el descubrimiento de recursos académicos pertinentes.

**Plataformas de venta de libros en línea:** Utilización del enfoque para sugerir a los clientes libros que podrían interesarles, basándose en sus preferencias y en las de usuarios con gustos similares.

**Sistemas de gestión de información personal:** Aplicación del sistema para organizar y recomendar documentos y recursos dentro de grandes colecciones personales o corporativas, mejorando la eficiencia en la gestión de información.

## 3.4 Comparativa

### 3.4.1 BigData

En cuanto a procesamiento de grandes volúmenes de datos (BigData), los tres artículos destacan que los volúmenes de datos actuales han crecido de manera exponencial y requieren tecnologías que permitan su procesamiento eficiente: El libro de Bertsekas y Tsitsiklis enfatiza el uso de métodos numéricos paralelos y distribuidos para manejar problemas a gran escala, lo cual es esencial en el análisis de Big Data.

El artículo sobre computación en empresas menciona cómo las compañías utilizan bases de datos distribuidas para gestionar millones de transacciones en tiempo real. El artículo sobre sistemas de recomendación muestra un caso específico donde grandes volúmenes de datos de usuarios y libros se procesan con algoritmos basados en grafos para mejorar la personalización.

Estos ejemplos demuestran que sin computación distribuida y paralela, el análisis de datos a gran escala sería inviable en muchas aplicaciones empresariales y científicas.

### 3.4.2 IA y Deep Learning

La IA y el aprendizaje profundo dependen de enormes cantidades de datos y poder de cómputo, lo que hace imprescindible el uso de arquitecturas distribuidas y paralelas.

- **El libro de Bertsekas y Tsitsiklis** trata algoritmos de optimización y métodos numéricos aplicados al entrenamiento eficiente de modelos de IA.
- **El artículo sobre computación en empresas** menciona cómo la computación distribuida impulsa la IA en sectores como el financiero, logístico y de biotecnología, mejorando la toma de decisiones basada en datos.
- **El artículo sobre sistemas de recomendación** demuestra un uso práctico de IA mediante algoritmos basados en grafos y redes neuronales para generar recomendaciones personalizadas.

Dado que los modelos de IA modernos requieren procesamiento masivo, el uso de sistemas distribuidos y paralelos permite acelerar entrenamientos e inferencias, haciendo posible aplicaciones como asistentes virtuales, visión por computadora y detección de fraudes en tiempo real.

### 3.4.3 Ámbito científico y análisis.

En el ámbito científico y análisis de datos complejos, la computación distribuida y paralela es fundamental para simulaciones científicas que requieren cálculos de alta precisión en diversos campos.

- **El libro de Bertsekas y Tsitsiklis** destaca su aplicación en dinámica de fluidos, resolución de ecuaciones diferenciales y modelado de fenómenos físicos complejos.
- **El artículo sobre computación en empresas** menciona su impacto en la biotecnología y el análisis financiero, donde la simulación de escenarios complejos es clave para la investigación y la toma de decisiones.
- **El artículo sobre sistemas de recomendación**, aunque no trata simulaciones científicas, aborda el análisis de redes de gran escala, lo que tiene aplicaciones en la investigación de estructuras de datos complejas.

Gracias a estas tecnologías, los científicos pueden realizar estudios más detallados en menos tiempo, desde la predicción del clima hasta la investigación en nuevos materiales y medicamentos.

### 3.4.4 Industria e investigación

#### 3.4.4.1 Aplicaciones en la Investigación

- **Procesamiento de señales e imágenes:** Utilizado en reconocimiento de patrones e imágenes médicas.
- **Análisis de grandes redes** – Investigaciones en redes sociales, comunicación y estructuras complejas.
- **Empresa de biotecnología** – Aceleración de la secuenciación genómica y descubrimiento de medicamentos.
- **Bibliotecas digitales académicas** – Recomendación de artículos y libros para facilitar el acceso a recursos científicos.

#### 3.4.4.2 Aplicaciones en la Industria

- **Empresa de logística** – Optimización de rutas de entrega en tiempo real.
- **Proveedor de almacenamiento en la nube** – Disponibilidad continua de datos con sistemas distribuidos.
- **Empresa de análisis financiero** – Evaluación de riesgos y simulaciones con algoritmos paralelos.
- **Gigante minorista durante temporadas pico** – Gestión eficiente de millones de transacciones simultáneas.

## 4. Bibliografía

[https://www.researchgate.net/profile/Jose-Aguilar-51/publication/267367623\\_Introduccion\\_a\\_la\\_Computacion\\_Paralela/links/544f7d6a0cf29473161ce888/Introduccion-a-la-Computacion-Paralela.pdf](https://www.researchgate.net/profile/Jose-Aguilar-51/publication/267367623_Introduccion_a_la_Computacion_Paralela/links/544f7d6a0cf29473161ce888/Introduccion-a-la-Computacion-Paralela.pdf)

<https://dspace.itsjapon.edu.ec/jspui/bitstream/123456789/407/1/computacion-distribuida-basada-en-objetivos.pdf>

[https://books.google.es/books?hl=es&lr=&id=n\\_Q5EAAQBAJ&oi=fnd&pg=PR3&dq=parallel+computation+and+distributed+computation+&ots=efDyKgyxp9&sig=5w52e8FNGgOrnhj6S7oCMmc\\_kDE&redir\\_esc=y#v=onepage&q=parallel%20computation%20and%20distributed%20computation&f=false](https://books.google.es/books?hl=es&lr=&id=n_Q5EAAQBAJ&oi=fnd&pg=PR3&dq=parallel+computation+and+distributed+computation+&ots=efDyKgyxp9&sig=5w52e8FNGgOrnhj6S7oCMmc_kDE&redir_esc=y#v=onepage&q=parallel%20computation%20and%20distributed%20computation&f=false)

<https://dl.acm.org/doi/abs/10.5555/39096#sec-recommendations>

[https://www.redalyc.org/journal/2570/257051186010/?utm\\_source=chatgpt.com](https://www.redalyc.org/journal/2570/257051186010/?utm_source=chatgpt.com)



## Computación Distribuida y Paralela

[https://fastercapital.com/es/contenido/Computacion-distribuida-y-paralela--creacion-de-una-ventaja-competitiva--el-papel-de-la-computacion-distribuida-y-paralela-en-las-empresas.html?utm\\_source=chatgpt.com#Computacion-distribuida-en-la-era-del-Big-Data](https://fastercapital.com/es/contenido/Computacion-distribuida-y-paralela--creacion-de-una-ventaja-competitiva--el-papel-de-la-computacion-distribuida-y-paralela-en-las-empresas.html?utm_source=chatgpt.com#Computacion-distribuida-en-la-era-del-Big-Data)

Definiciones:

[https://books.google.es/books?id=n\\_Q5EAAQBAJ&printsec=frontcover&hl=es#v=onepage&q&f=false](https://books.google.es/books?id=n_Q5EAAQBAJ&printsec=frontcover&hl=es#v=onepage&q&f=false)

Utilizamos ChatGPT para que nos hiciera una selección de artículos de entre 10 y 12 para seleccionar nosotros según lo que necesitáramos y de ahí vienen los artículos mencionados