

Computación Distribuida y Paralela

Juan Carlos Cox Fernández
Adrián Yared Armas de la Nuez

1

Definición

Computación Distribuida y Paralela



Definición

La computación paralela y distribuida ha crecido debido a la necesidad de resolver problemas complejos y avances tecnológicos. Tales como ecuaciones diferenciales, dinámica de fluidos u optimización de sistemas interconectados

Computación Distribuida

Los sistemas distribuidos conectan procesadores geográficamente separados mediante redes, donde la comunicación puede ser más lenta o menos fiable

Computación Paralela

Los sistemas paralelos agrupan varios procesadores en un mismo entorno con comunicación más rápida y estable, trabajando juntos en una tarea específica.



Tipos de computación paralela

Masivamente paralelas

Miles de procesadores para tareas complejas



Supercomputación

Control centralizado

Instrucciones comunes o autonomía total



Sincronización

Asíncrona → Flexibilidad (Hilos no esperan a otros)

Síncrona → Control (Hilos esperan por otros)



Interconexión de procesadores

Memoria compartida → Sincronización (varios hilos modifican una sola región de memoria)

Paso de mensajes → Comunicación (los procesos envían mensajes)

Híbrido → Flexibilidad (Comunicación entre nodos)





Tipos de computación distribuida

Clúster

Grupo de ordenadores trabajando como uno solo.



Grid

Ordenadores distribuido globalmente trabajando en tareas grandes.



Nube

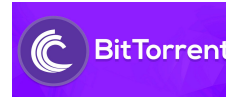
Recursos remotos accesibles por internet



Google Cloud

P2P

Nodos sin servidor central



Niebla

Procesa datos cerca del usuario



Borde

Computación local en dispositivos inteligentes



2

Ventajas

Computación Distribuida y Paralela



Ventajas

Distribuida

- Permite dividir el problema en varias tareas ejecutadas simultáneamente, reduciendo el tiempo total de ejecución en comparación con el procesamiento secuencial
- se pueden agregar más nodos para aumentar la capacidad de cómputo sin necesidad de reemplazar el hardware existente.
- Se pueden procesar grandes conjuntos de datos de manera más eficiente (ejemplo: Big Data con Hadoop y Spark).

Paralela

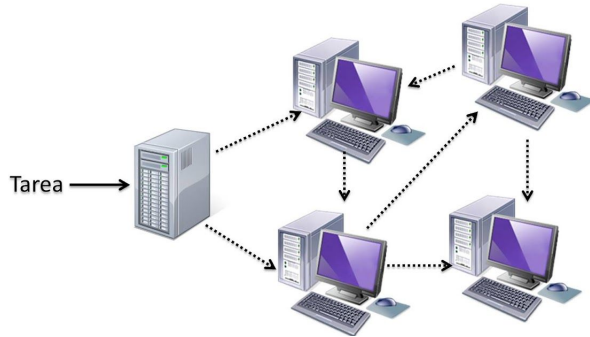
- Se aprovechan múltiples núcleos o máquinas para ejecutar tareas más rápido.
- Permite optimizar el rendimiento al utilizar múltiples procesadores en una sola máquina.
- Útil en aprendizaje automático, simulaciones científicas y modelado de sistemas complejos



Ventajas

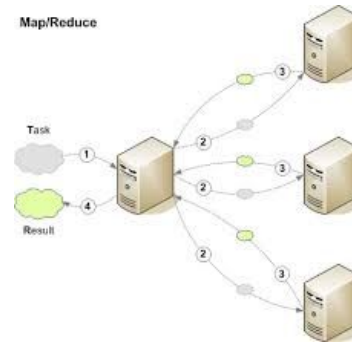
Distribuida

- Pueden continuar funcionando incluso si un nodo falla, mejorando la fiabilidad.
- las tareas pueden ser asignadas dinámicamente a nodos con menor consumo energético.



Paralela

- Algunas arquitecturas paralelas implementan redundancia para evitar fallos críticos.
- distribuir la carga en varios núcleos pequeños puede reducir el consumo total de energía.



3

Desventajas

Computación Distribuida y Paralela



Desventajas

Distribuida

- Requiere algoritmos específicos para la distribución y sincronización de tareas.
- La comunicación entre nodos introduce latencia y sobrecarga en la red.
- Se necesita hardware más potente o una red de ordenadores, lo que puede aumentar costos.
- Garantizar la consistencia de los datos puede ser un desafío

Paralela

- Es más difícil de depurar y mantener debido a la concurrencia y comunicación entre procesos.
- La sincronización de tareas en la computación paralela puede generar cuellos de botella.
- Los algoritmos deben ser diseñados o adaptados para aprovechar el paralelismo, lo que requiere más tiempo de desarrollo.
- El acceso concurrente a memoria compartida puede generar condiciones de carrera o bloqueos.



Tabla de comparación

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

4

Artículos

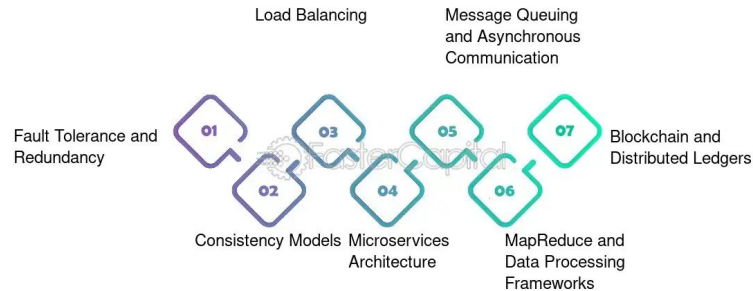
Computación Distribuida y Paralela



Artículo 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.

Key Concepts and Technologies in Distributed Systems





Artículo 1: Argumentos

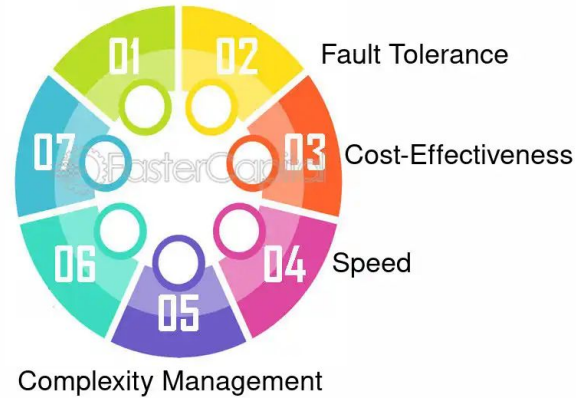
Distributed Computing in the Era of Big Data

1. Escalabilidad y flexibilidad
2. Rentabilidad
3. Procesamiento de datos mejorado
4. Aceleración de la innovación
5. Confiabilidad y redundancia
6. Colaboración global

Scalability

Data Locality

Resource Sharing





Artículo 1: Casos de uso

1. Gigante minorista durante temporadas pico
2. Empresa de análisis financiero
3. Empresa de logística





Artículo 2: 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.

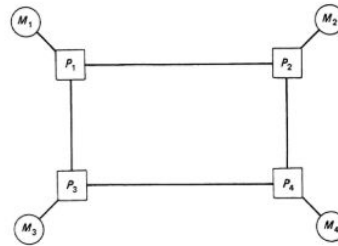
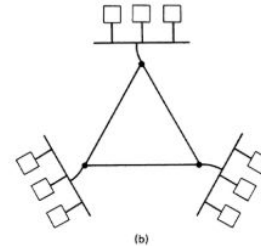
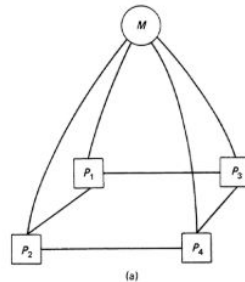


Figure 1.1.2 An interconnection network joining a set of processors P_i , each one endowed with its own local memory M_i .





Artículo 2: Argumentos

1. Eficiencia computacional
2. Escalabilidad
3. Flexibilidad en la resolución de problemas
4. Robustez y tolerancia a fallos

Aprovechar el poder de la computación en la nube para aplicaciones distribuidas

Consejos para adoptar la
computación en la nube para
aplicaciones distribuidas

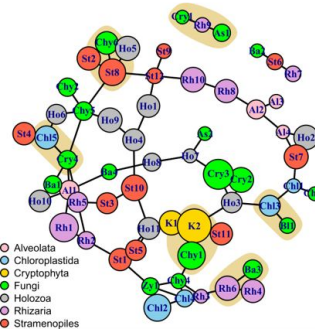
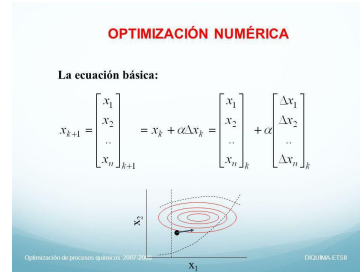
Estudio de caso





Artículo 2: Casos de uso

1. Solución de sistemas de ecuaciones lineales
2. Optimización numérica
3. Simulaciones en dinámica de fluidos
4. Análisis de grandes redes

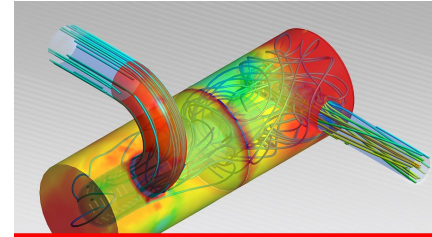


REGULACIÓN

SISTEMAS de Ecuaciones

$$\begin{cases} x + y = 3 \\ 2x + y = 2 \end{cases} \quad \begin{cases} 2x + y = 1 \\ -2x + 3y = 19 \end{cases}$$

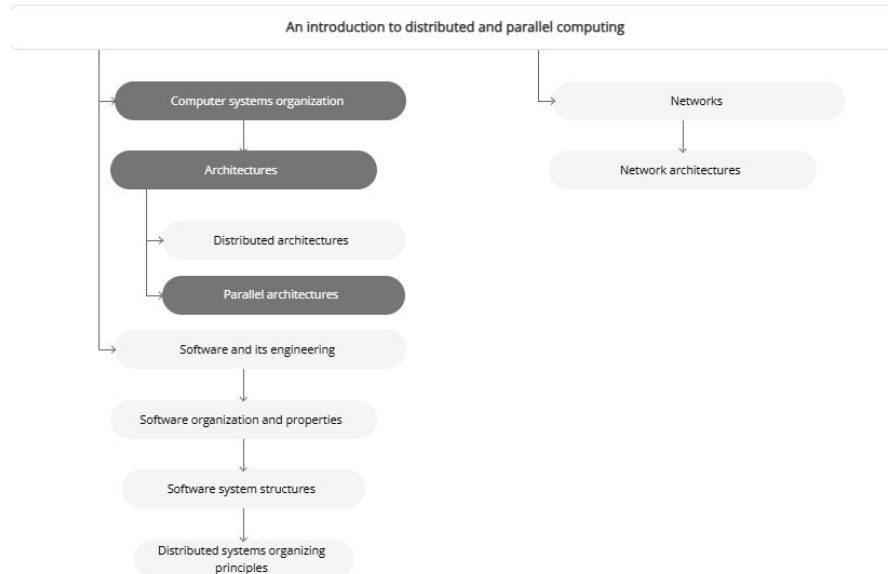
Comprobar si $x = -2$ e $y = 5$ es **SOLUCIÓN**





Artículo 3: 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.

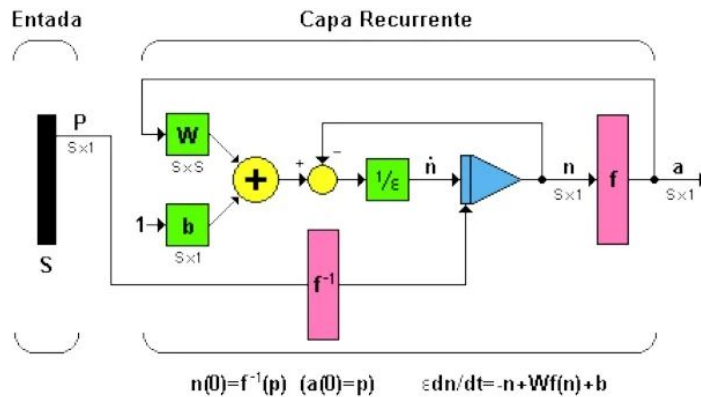




Hopfield

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

Red Hopfield.-Estructura





Artículo 3: Argumentos

Predictions for Distributed and Parallel Computing

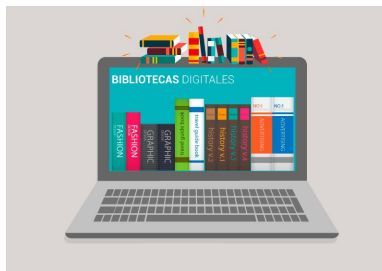
1. Mejora de la calidad de las recomendaciones
2. Explotación de asociaciones complejas
3. Evaluación empírica positiva





Artículo 3: Casos de uso

1. Bibliotecas digitales académicas
2. Plataformas de venta de libros en línea
3. Sistemas de gestión de información personal





Comparativa

Big data

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

Deep Learning e IA

En cuanto a los artículos, se señala que, debido a que los modelos de IA modernos requieren un procesamiento masivo, el uso de sistemas distribuidos y paralelos permite acelerar tanto el entrenamiento como las inferencias. Esto hace posible aplicaciones avanzadas, como asistentes virtuales, sistemas de visión por computadora y detección de fraudes en tiempo real.



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.

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.



Bibliografía

Enlaces de artículos utilizados:

[Definiciones](#)

[Computación distribuida y paralela: creación de una ventaja Competitiva](#)

[Parallel and Distributed Computation: Numerical Methods](#)

[A Graph-Based Recommender System for Digital Library](#)

Otros enlaces:

[An introduction to distributed and parallel computing](#)

[Computación Distribuida basada en Objetivos](#)

[Introducción a la computación paralela](#)



Si tienes alguna pregunta sobre la exposición es el momento.