

Aplicaciones de la Computación Paralela por medio de Clusters

Propuesta de Investigación

Sonia Juliana Mora Niño
Ingeniería de Sistemas
Facultad de Ingeniería de Sistemas
smora9@unab.edu.co

Sebastian Roa Prada (Tutor)
Ingeniería Mecatrónica
Facultad de Ingenierías Fisicomecánicas
sroa@unab.edu.co

Universidad Autónoma de Bucaramanga

Resumen

Esta investigación busca desarrollar en la Universidad Autónoma de Bucaramanga una plataforma que permita resolver problemas de computación numérica empleando elementos de procesamiento de manera simultánea, por medio de la técnica de computación paralela. Con este fin se propone el desarrollo e implementación de "clusters", los cuáles mejoran el rendimiento en los cálculos numéricos intensivos aprovechando la disponibilidad de más de una computadora.

Abstract

This research seeks the development at Universidad Autónoma de Bucaramanga of a platform to allow solving numerical computing problems employing processing elements simultaneously, by means of the parallel computing technique. To reach this goal, this article proposes the development and implementation of computer clusters, which improve the performance during intensive numerical calculations, in cases where more than one computer is available to carry out a given work.

Área de conocimiento

El tema de la investigación está relacionado con el área de Ingenierías, más específicamente con Ingeniería de Sistemas y Mecánica Computacional.

Palabras clave

Paralelismo, cluster, multiprocesador

1. Introducción

La computación paralela es el nombre que se le da al procesamiento de datos de manera simultánea. Se procesa la información "en paralelo", es decir, dividiendo una tarea en áreas, problemas o trabajos específicos, y luego derivando cada una de estas partes a diferentes ordenadores o partes dentro de uno de ellos, se aumenta la velocidad y eficiencia en los procesos. Parece el nombre de una tecnología experimental, pero es la norma hoy en día, dado el creciente requerimiento en procesamiento del software y aplicaciones actuales en general. Por ejemplo si se está trabajando en un computador de doble núcleo, el equipo emplea la computación paralela; asimismo si se tiene un Playstation 3, este emplea el chip Cell de IBM, que también pertenece a esta categoría.

Existen otras formas de computación paralela como el empleo de redes de computadoras para llevar a cabo una gran tarea o grupo de tareas, completando operaciones en segundos o minutos, cosa que a

una sola computadora le tomaría mucho más tiempo; a este grupo de computadoras interconectadas trabajando como si fuera una sola se le llama "Cluster", grilla o red.

Estructura del Artículo: En la Sección 2 se presenta el objetivo general y objetivos específicos. La Sección 3 ofrece la metodología de investigación propuesta. La Sección 4 muestra los referentes teóricos. La Sección 5 corresponde al cronograma. Y la Sección 6 pertenece a los resultados esperados.

2. Objetivos

2.1 Objetivo general

Implementar el paralelismo por medio de los clusters en las instalaciones de la Universidad Autónoma de Bucaramanga.

2.2 Objetivos específicos

- Montar el cluster en la Universidad Autónoma de Bucaramanga.
- Investigar todo el diseño y configuración que un cluster implica.
- Adquirir todo el conocimiento adecuado en el intercambio de datos y comunicación de un cluster.
- Emplear los atributos que brindan los clusters de ordenadores de forma adecuada para cualquier propósito.

3. Metodología de investigación

Como metodología en esta investigación se empezaría por el planteamiento del problema a resolver.

Problema: El uso de computadoras es muy frecuente y más en una institución educativa y la falta de tiempo para poder ejecutar varios procesos, nos lleva a la búsqueda de una posible solución y es por medio de las aplicaciones que nos brinda la computación paralela aprovechando además que la difusión creciente de los sistemas paralelos ha provocado que la programación paralela haya adquirido gran relevancia en la actualidad.

Búsqueda del estado del arte: En primera instancia se conseguirán las fuentes bibliográficas necesarias que provean los fundamentos técnicos y científicos de los clusters y las técnicas de computación paralela

Análisis de la información recopilada: Consiste en apropiarse y entender el funcionamiento e implementación de clusters, así como también su aplicación en la solución de problemas físicos.

4. referentes teoricos

El paralelismo se ha empleado durante muchos años, sobre todo en informática de alto rendimiento, pero el interés ha crecido últimamente debido a las limitaciones físicas que impiden la ampliación de frecuencias. La computación paralela es una técnica de programación en la que muchas instrucciones se ejecutan simultáneamente. Utiliza múltiples elementos de proceso simultáneamente para resolver un problema. Esto se logra mediante la división del problema en partes independientes de modo que cada elemento de proceso puede ejecutar su parte del algoritmo de forma simultánea con los otros. Los elementos de procesamiento pueden ser muy diversos e incluyen recursos como un solo equipo con varios procesadores, varios ordenadores en red, hardware especializado, o cualquier combinación de los anteriores.

En la actualidad, la computación paralela está siendo utilizada en multitud de campos para el desarrollo de aplicaciones y el estudio de problemas que requieren gran capacidad de cómputo, bien por el gran tamaño de los problemas que abordan o por la necesidad de trabajar con problemas en tiempo real. De esta forma, el paralelismo en la actualidad, además de constituir diversas líneas abiertas de intensa labor investigadora, puede encontrarse en infinidad de aplicaciones en campos muy variados, entre los que se destacan:

- El modelado predictivo y simulación: se realiza mediante extensos experimentos de simulación por computador que con frecuencia acarrearán computaciones a gran escala para obtener la precisión y el tiempo de respuesta deseado. Entre estos modelados están la previsión meteorológica numérica y la oceanografía.
- El desarrollo industrial también reclama el uso de computadores para progresar en el diseño y automatización de proyectos de ingeniería, la inteligencia artificial y la detección remota de los recursos terrestres. En este campo se destaca: la inteligencia artificial y automatización.
- Investigación médica: En el área médica los computadores rápidos son necesarios en tomografía asistida, diseño de corazones artificiales, diagnóstico hepático, estimación de daños cerebrales y estudios de ingeniería genética.

Los ordenadores paralelos pueden ser clasificados según el nivel en el que el hardware es compatible con el paralelismo, los ordenadores multinúcleo y multiproceso tienen varios elementos de procesamiento en una sola máquina, mientras que los clusters, los MPP y los grids emplean varios ordenadores para trabajar en la misma tarea.

Un grupo de computadoras (cluster) consiste en un conjunto de ordenadores conectados que trabajan en conjunto para que en muchos aspectos puedan ser vistos como un solo sistema.

Los componentes de un grupo suelen estar conectados entre sí a través de rápidas redes de área local, cada nodo que ejecuta su propia instancia de un sistema operativo. Los clusters de ordenadores surgieron como resultado de la convergencia de una serie de tendencias de la computación, incluyendo la disponibilidad de microprocesadores de bajo coste, redes de alta velocidad, y el software para un alto rendimiento de computación distribuida. Los Clusters de ordenadores tienen una amplia gama de aplicación y despliegue, que van desde las agrupaciones de pequeñas empresas con un puñado de nodos a algunas de las más rápidas supercomputadoras en el mundo, como el equipo K.

Además se pueden configurar para diferentes propósitos que van desde negocios de propósito general, tales como las necesidades de apoyo de servicios web, de computación intensiva y cálculos

científicos. En cualquier caso, el grupo de ordenadores puede utilizar un enfoque de alta disponibilidad.

5. Cronograma

Actividad	Abr May	Jun Jul	Ago Sep	Oct Nov	Dic
Recolección de información general de "clusters"	X				
Recopilación de información de computación paralela y sus aplicaciones		X			
Recopilación del estado del arte de computación paralela por medio de "clusters" aplicado a problemas de computación numérica			X		
Análisis de la información recopilada				X	
Entrega de reporte					X

6. Resultados esperados

Como resultado principal sería lograr cumplir con los objetivos propuestos, para alcanzar una evolución con la aplicación del paralelismo en la Universidad.

Poder aprovechar adecuadamente el potencial de los sistemas paralelos, con el estudio de técnicas de análisis y diseño de algoritmos paralelos, y estudiarse con un nivel de abstracción similar al que se utiliza en la programación secuencial.

Obtener el conocimiento necesario que permita más adelante poner a disposición de entidades externas interesadas un cluster para resolver problemas numéricos intensivos aprovechando las ventajas de la computación paralela.

7. Referencias

- [1] Bader, David; Robert Pennington (June 1996). "Cluster Computing: Applications". Georgia Tech College of Computing.
- [2] abcNetwork-Based Information Systems: First International Conference, NBIS 2007 ISBN 3-540-74572-6 page 375
- [3] Almasi, GS and A. Gottlieb (1989). Highly Parallel Computing. Benjamin-Cummings publishers, Redwood City, CA.
- [4] SV Adve et al. (November 2008). "Parallel Computing Research at Illinois: The UPCRC Agenda" (PDF).
- [5] Universidad de Alicante, Dpto. De Ciencia de la Computación e Inteligencia Artificial.

<http://www.dccia.ua.es/cp/index.htm>