

Glosario

Camilo Andrés Rivera

19 de Junio de 2015

Ley de Amdahl's

La ley de Amdahl en su forma más literal enuncia que *La mejora obtenida en el rendimiento de un sistema debido a la alteración de uno de sus componentes está limitada por la fracción de tiempo que se utiliza dicho componente*. Sin embargo, aplicado a la computación en paralelo nos dice que el incremento en la velocidad de un programa que se ejecuta de manera paralela, estará limitada por la sección del programa que se ejecute de manera secuencial.

Supercómputo/HPC

El supercómputo (o *High-Performance Computing*) en su definición más cruda es el uso de súper computadoras o computadoras de alta capacidad de procesamiento para la resolución de problemas en un menor tiempo, o que simplemente eran impensables de resolver. Hoy en día este concepto se ha transformado al uso de muchos procesadores de manera paralela para la resolución de problema.

CPU/Socket/Procesador/Core

A pesar de que no siempre estos términos suelen ser intercambiables, en el cómputo en paralelo se refieren al mínimo ente de procesamiento accesible por el usuario; es decir, en cómputo en paralelo se refiere a cada uno de los cerebros que ejecutan tareas de manera independiente.

Tarea

Cada tarea es un programa independiente que se ejecuta en alguno de los cores y posee su propio stack de memoria, sin depender de otras tareas.

Pipeline

El pipeline es un tipo de dependencia de tareas en el cual se tiene un único flujo posible de manera secuencial y cada tarea es ejecutada únicamente al terminar la anterior. Este tipo de dependencia no se puede paralelizar y se asemeja a una cadena de producción.

Memoria compartida

Es un tipo de arquitectura para las computadoras destinadas a trabajar en paralelo en la cual los cores comparten un banco de memoria al que todos pueden acceder mediante direcciones de memoria globales.

Multi-Procesador Simétrico

Éste es un caso específico de las arquitecturas de memoria compartida en la cual todos los procesadores que tienen acceso a la memoria principal son completamente idénticos, tanto en hardware como en software, y tienen los mismos permisos y posibilidades.

Memoria Distribuida

Es otro tipo de arquitectura u organización para el cómputo paralelo en el cual cada procesador posee su propio banco de memoria y sólo éste es capaz de acceder a ella. Para que cada procesador pueda acceder a información contenida en la memoria de otro debe realizar una solicitud para esta, para lo cual se debe tener un protocolo de paso de mensajes, como lo es el MPI.

Comunicaciones

Todo cómputo en paralelo debe tener en cuenta en sus consideraciones que va a existir un conjunto de datos, mensajes e información que se estará intercambiando no sólo entre cada uno de los sockets, pero también con el usuario que se conecta con el supercomputador. Además, para ello se requiere toda una serie de protocolos e infraestructura que lo permite. Todas estas consideraciones hacen parte de las comunicaciones de un sistema paralelo.

Sincronización

Al realizar cómputo en paralelo, diferentes partes del código o del problema en sí se ejecutan distribuídamente en diferentes lugares o procesadores. Dependiendo de la carga que exija cada sección o de las capacidades de cada procesador, lo más esperable es que no todos terminen a tiempo, de manera que se vuelve necesario introducir mecanismos de sincronización para que se pueda continuar con las secciones secuenciales del código correctamente.

Granularidad

La granularidad se refiere a la capacidad de detalle de información que se puede llegar a obtener de una super computadora. Por ejemplo, ya que los grids contienen aportes que no sólo son procesadores, sino recursos completos (procesadores, memoria e infraestructura) su granularidad es más gruesa ya que los elementos son muy grandes. En cambio un clúster, o un SMP, se conoce el estado de cada uno de los elementos, cómo se comportan y se puede solicitar información muy detallada y específica; por lo tanto su granularidad es fina.

Speed-up observado

El speed-up observado en la paralelización de un problema o un código es una medida de la mejoría del rendimiento del problema, y está definido en términos de la métrica utilizada para medir el rendimiento de cada código, por ejemplo el tiempo de ejecución. En dado caso, el speed-up será la razón entre el tiempo de ejecución anterior y el nuevo.

Paralelismo Masivo

El paralelismo masivo se refiere al uso de una gran cantidad de procesadores y recursos (mayor al habitual) para la aplicación de computación paralela, estos estando interconectados por redes no necesariamente especializadas para dicho fin. Un ejemplo de ello es la computación GRID en la cual varios clústers de computadores pueden hacer parte del grid.

Escalabilidad

La escalabilidad, ya sea de un sistema, proceso o código, se refiere a la capacidad (y facilidad) que éste tiene para adaptarse a cambios. Por ejemplo, que un algoritmo funcione igualmente, o sin mayores cambios, para un mayor número de datos, o procesadores. Un supercomputador escalable es aquel que sin mayor problema tiene la capacidad de recibir un mayor número de procesadores o recursos.