### **Rendimiento relativo**

Cuando se dice que un computador es más rápido que otro, ¿qué queremos significar?

El usuario del computador puede decir que un computador es más rápido cuando ejecuta un programa en menos tiempo, mientras que el director de un centro de cálculo puede decir que un computador es más rápido cuando completa más tareas en una hora. El usuario del computador está interesado en reducir el *tiempo de respuesta.* El director del centro de cálculo está interesado en incrementar la *productividad (throughput).*

**Tiempo de respuesta** será el tiempo transcurrido entre el inicio y fin de una tarea, también se llama tiempo de ejecución o latencia.

**Productividad** es la cantidad de trabajos realizados en un determinado tiempo. Por ejemplo el tiempo de una operación de E/S en un disco será el tiempo respuesta y la cantidad de datos que van y vienen en un determinado tiempo dará la productividad del mismo. También se lo llama ancho de banda.

A los términos latencia y ancho de banda se lo usa cuando se habla de sistemas de memorias. Normalmente, los términos «tiempo de respuesta», «tiempo de ejecución» y «productividad» se utilizan cuando se está desarrollando una tarea de cálculo completa.

Cuando se dice que un computador es más rápido que otro se quiere decir que el tiempo de ejecución o de repuesta de una máquina es superior a la otra en una tarea dada, por lo tanto " X es n % más rápida que Y", o sea:

Tiempo de ejecución x = n en porcentaje n\*100= n %

Tiempo de ejecución y

Como el tiempo de ejecuciones es reciproca al rendimiento:

1 .

n = Tiempo de ejecución x = Rendimiento x = Rendimiento y

Tiempo de ejecución y 1 Rendimiento x

Rendimiento y

n = Rendimiento x

Rendimiento y

Si n fuese igual 0,30 se puede expresar “la productividad de X es 30 % superior a la de Y” o " la productividad de X es 1,30 veces superior a la de Y"

**El rendimiento y tiempo de ejecución son recíprocos, por lo tanto incrementar el rendimiento hace decrecer el tiempo de ejecución.**

# Ley de Amdahl

Para mejorar el tiempo de ejecución, una de las mejoras más importantes en el diseño de computadoras es quizás el “acelerar” el caso frecuente o común. Mejorar el evento frecuente en lugar del evento raro ayuda a aumentar el rendimiento.

Además el caso frecuente es más sencillo de resolver y puede ejecutarse más rápido que el evento infrecuente.

Una ley fundamental que cuantifica este principio es la ley de Amdahl.

Esta ley establece que la mejora en el rendimiento al utilizar algún modo de ejecución más rápido está limitada por la fracción de tiempo que se pueda utilizar ese modo rápido.

La ley de Amdahl define la ganancia de rendimiento o aceleración global que se puede lograr al utilizar una característica particular, suponiendo que podemos hacer una mejora en una máquina que mejore su rendimiento, la aceleración global es la relación:

Aceleración global o de rendimiento =

= Rendimiento de la tarea completa utilizando la mejora cuando es posible

Rendimiento de la tarea sin utilizar la mejora

Aceleración global o de rendimiento =

= Tiempo de ejecución sin utilizar la mejora =

Tiempo de ejecución de la tarea completa utilizando la mejora cuando se posible

La aceleración nos indica la rapidez con que se realizará una tarea utilizando una máquina con la mejora respecto a la misma máquina sin la mejora.

La aceleración de pende de dos factores:

1. La fracción del tiempo de cálculo de la máquina original que puede utilizarse para aprovechar la mejora:

Fracción mejorada = Tiempo de mejora =

Tiempo total

esta será siempre menor a 1.

1. La optimización lograda por el modo de ejecución mejorado; es decir cuanto más rápido se ejecutaría la tarea si solamente se utilizara el modo mejorado. Este valor es el tiempo del modo original con respecto al tiempo de modo mejorado y es siempre mayor que 1. Llamaremos **Aceleración mejorada** a este valor.

Ahora el tiempo de ejecución utilizando la máquina original con el modo mejorado será el tiempo empleado utilizando la parte no mejorada de la máquina más el tiempo empleado utilizando la parte mejorada.

Tiempo de ejecución nuevo = Tiempo de ejecución antiguo x

( 1 – Fracción mejorada) + Fracción mejorada

Aceleración mejorada

La aceleración global es la relación de los tiempos de ejecución:

Aceleración global = Tiempo de ejecución antiguo =

Tiempo de ejecución nuevo

= 1 =

( 1 – Fracción mejorada ) + Fracción mejorada

Aceleración mejorada

Un corolario de la Ley de Amdahl es que si una mejora solo es utilizable por una fracción de una tarea, no podemos aumentar la velocidad de la tarea más que el reciproco de 1 menos esa fracción.

Esta ley puede servir como guía para ver como una mejora aumenta el rendimiento y como distribuir los recursos para mejorar la relación costo/rendimiento.

### **Rendimiento de la CPU**

El tiempo de CPU es la medida del rendimiento del computador: un computador que realiza la misma cantidad de trabajo en el mínimo tiempo es el más rápido. Este tiempo incluye los tiempos necesarios para completar una tarea incluyendo acceso a disco, acceso a memorias, actividades de E/S, gastos del sistema operativo. Por ejemplo en el caso de multiprogramación la CPU trabaja sobre otro programa mientras espera la E/S y puede no minimizar el tiempo transcurrido en un programa, por lo tanto se debe buscar otro término que lo exprese mejor. El tiempo que ve el usuario es el tiempo transcurrido del programa, no el de la CPU. Por lo tanto el tiempo de CPU reconoce esta distinción y mide el tiempo que la CPU está calculando sin incluir tiempos de E/S o de otros programas. El tiempo de la CPU se puede dividir en **tiempo de CPU del usuario** y tiempo empleado por el sistema operativo para realizar tareas del programa y que llamaremos **tiempo de CPU del sistema.**

Lo que nos interesa en este momento es el rendimiento de la CPU o sea el tiempo de CPU del usuario.

Las computadoras tienen relojes que trabajan a una frecuencia constante. Estos eventos de tiempo se llaman ciclos, ciclos de reloj, pulsos de reloj, etc. (La frecuencia es la inversa del tiempo de un ciclo). Entonces se referencia las tareas por la cantidad y duración de estos ciclos.

El tiempo de la CPU se puede definir:

Tiempo de la CPU = Ciclos de reloj de CPU para un programa x Duración del ciclo de reloj

Otra forma es

Tiempo de CPU = Ciclos de reloj de CPU para un programa

Frecuencia de reloj

Dado que los tiempos de dispositivos de E/S son independientes de los ciclos de reloj, es necesario tomar otro parámetro para medir el tiempo de CPU. Por lo tanto si conocemos el número de ciclos de reloj y el recuento de instrucciones podemos calcular el número medio de *ciclos de reloj por instrucción* (CPI),

CPI = Ciclos de reloj para un programa

Recuento de instrucciones

Entonces:

Tiempo de CPU = Recuento de instrucciones x CPI x Duración del ciclo de reloj

o

Tiempo de CPU = Recuento de instrucciones x CPI =

Frecuencia de reloj

En el caso que distinto tipo de instrucciones tenga distintas cantidades de ciclos por tipo de instrucción el CPI total será

Tiempo de CPU = T x Σ Recuento de instrucciones tipo i x CPIi (i = tipo de

instrucción)

CPI = Σ CPIi x Fi

donde CPIi = ciclo por instrucción de tipo i

Fi = frecuencia de aparición de instrucciones tipo i

Buscando una medida estándar del rendimiento se han creados parámetros como los MIPS (millones de instrucciones por segundo), para un programa dado. Estos MIPS se definen como:

Tiempo de ejecución = Recuento de instrucciones

MIPS x 10⁶

MIPS = Recuento de instrucciones = 1 = f .

Tiempo de ejecución x 10⁶ CPI x T x 10⁶ CPI X 10⁶

El problema que tiene este tipo de medidas es que los MIPS son dependientes de repertorio de instrucciones de las computadoras y estos difieren de una a otras.

Por lo tanto la medida más fiable consistente del rendimiento es el tiempo de ejecución de los programas.