Andrés Calva Valencia A01169639
Marco Isaac Buendía Mejía A01271519
Siegfried Paul Keller Schippner A01375356
ITESM CEM

Límites en rendimiento

Límites asintóticos

Los limites asintóticos establecen un rango de límites optimistas y pesimistas hablando de rendimiento y tiempo de respuesta del sistema en colas sencillas. Surgen a partir de condiciones de cargas ligeras y pesadas.

La información de los limites asintóticos varía dependiendo de si el sistema está abierto (transacciones) o es cerrado (lotes o terminales).

Carga de trabajo por transacciones

El límite del rendimiento nos indica la máxima tasa de arribo que el sistema puede procesar.

(λ)

A partir de que se supere ese límite, se empiezan a generar clientes sin procesar, originando las colas. A este estado se le llama saturación del sistema. Debido a que todo el sistema cae en saturación si cualquier centro, el rendimiento del sistema es tasa de arribo mínimo de todos los centros.

(λ_{sat})

Usamos la siguiente fórmula para calcular los límites por transacción:

$$X(\lambda) \leq \frac{1}{\mathrm{D}_{\max}}$$

Y la siguiente para tiempo de respuesta:

$$D \leq R(\lambda)$$

Carga de trabajo por lotes y terminales

Para realizar las fórmulas para los límites por lotes y terminales hay que tener en cuenta que:

- El rendimiento más pequeño ocurre cuando cada cliente adicional debe formarse detrás de los que están en el sistema.
- El más grande ocurre cuando ningún cliente es atrasado por algún cliente

Entonces, obtenemos que:

$$\frac{N}{D + \left((N-1) * D_{max}\right)} \leq X(N) \leq min\left(\frac{N}{D}, \frac{1}{\mathrm{D}_{\max}}\right)$$

$$\frac{\mathrm{N}}{(N * D) + Z} \leq X(N) \leq min\left(\frac{\mathrm{N}}{\mathrm{D} + \mathrm{Z}}, \frac{1}{\mathrm{D}_{\max}}\right)$$

Y para obtener las fórmulas de tiempo de respuesta, transformamos las de rendimiento usando la ley de la utilización:

$$max(D, N * \mathrm{D}_{\max}) \leq R(N) \leq N * D$$

$$max(D, (N * \mathrm{D}_{\max}) - Z) \leq R(N) \leq N * D$$

Balanceo de cargas

El balanceo de cargas son límites que están basados en sistemas que están balanceados de manera que la demanda de servicio es el mismo en todos los centros.

Para balancear los centros, se realiza lo siguiente:

1. Reemplazar los centros con los nuevos
2. Mover los clientes de un centro a otro, de manera que sus demandas estén balanceadas
3. Agregar el nuevo centro
4. Calcular las nuevas demandas

Límites para sistemas balanceados derivados de Uk(N)=N/(N+k-1)

Para estos límites, la demanda del servicio en cada centro está balanceados.

$$\frac{N}{D + \left((N-1) * D_{max}\right)} \leq X(N) \leq min\left(\frac{1}{\mathrm{D}_{\max}}, \frac{\mathrm{N}}{D + \left((N-1) * \mathrm{D}_{\mathrm{ave}}\right)}\right)$$

$$\frac{N}{\mathrm{D} + \mathrm{Z} + \left(\frac{\left((N-1) * \mathrm{D}_{\max}\right)}{1 + \frac{\mathrm{Z}}{\mathrm{N*D}}}\right)} \leq X(N) \leq min\left(\frac{1}{\mathrm{D}_{\max}}, \frac{\mathrm{N}}{D + Z + \frac{\left((N-1) * \mathrm{D}_{\mathrm{ave}}\right)}{1 + \frac{\mathrm{Z}}{\mathrm{D}}}}\right)$$

$$X(\lambda) \leq \frac{1}{\mathrm{D}_{\max}}$$

$$max\left(N * \mathrm{D}_{\max}, D + \left((N-1) * \mathrm{D}_{\mathrm{ave}}\right)\right) \leq R(N) \leq D + \left((N-1) * \mathrm{D}_{\max}\right)$$

$$max\left((N * \mathrm{D}_{\max}) - Z, D + \frac{\left((N-1) * \mathrm{D}_{\mathrm{ave}}\right)}{1 + \frac{\mathrm{Z}}{\mathrm{D}}}\right) \leq D + \frac{\left((N-1) * \mathrm{D}_{\max}\right)}{1 + \frac{\mathrm{Z}}{\mathrm{N*D}}}$$

$$\frac{\mathrm{D}}{1 - (\lambda * \mathrm{D}_{\mathrm{ave}})} \leq R(\lambda) \leq \frac{\mathrm{D}}{1 - (\lambda * \mathrm{D}_{\mathrm{ave}})}$$

Ejemplo balanceo

Dado un sistema de CPU y dos discos (uno rápido y uno lento), se obtuvieron los siguientes datos:

- **Intervalo de observación (T)**
- **Tiempo activo de CPU (B)**
- **Tiempo activo de disco rápido (B)**
- **Tiempo activo de disco lento (B)**
- **Trabajos completados CPU (C)**
- **Trabajos completados lento (C)**
- **Trabajos completados rápido (C)**
- **Tiempo de reflexión (Z)**

Se necesita que el CPU se reemplace por uno el doble de rápido, balancear demandas entre el disco rápido y lento.

Calculamos demandas, número de visitas y tiempo de servicio por visitas:

$$\mathrm{Dk} = \frac{\mathrm{Bk}}{\mathrm{C}} \quad \mathrm{Vk} = \frac{\mathrm{Ck}}{\mathrm{C}} \quad \mathrm{Sk} = \frac{\mathrm{Bk}}{\mathrm{Ck}}$$

D1 = 2 D2 = 0.5 D3 = 3

V2 = 10 V3 = 100

S2 = 0.05 S3 = 0.03

Reemplazamos CPU por uno el doble de rápido

D1 = 1

Balanceamos cargas:

$$\mathrm{Dk} = \mathrm{Vk{*}Sk}$$

$V_2 + V_3 = 110$

Despejamos:

$$\frac{\mathrm{V_2{*}S_2}}{.05} + \frac{\mathrm{V_3{*}S_3}}{.03} = 110$$

$$\mathrm{D_2}\left[\frac{1}{0.05} + \frac{1}{0.03}\right] = 110$$

$V_2 = 41$ y $V_3 = 69$