Formulario Examen Final ACSI

Lluis Barca Pons y Jorge González Pascual

20 de junio de 2021

1. Tema 1: Introducción

- **Comparación de prestaciones:**

$$A = \frac{\text{Tiempo de ejecución}_Y}{\text{Tiempo de ejecución}_X} = \frac{Rendimiento_X}{Rendimiento_Y}$$

$$\frac{\text{Tiempo de ejecución}_Y}{\text{Tiempo de ejecución}_X} = \frac{Rendimiento_X}{Rendimiento_Y} = 1{,}0 + \frac{n}{100}$$

- **Ley de Amdahl (con una mejora):**

$$A = \frac{1}{(1-f) + (\frac{f}{k})}$$

- **Ley de Amdahl (con n mejoras):**

$$A = \frac{1}{(1-\sum_{i=1}^{n} f_i) + \sum_{i=1}^{n}(\frac{f}{k})}$$

  - Si $f = 0 \rightarrow A = 1$, es decir, no hay ninguna mejora en el sistema.
  - Si $f = 1 \rightarrow A = k$, es decir, el sistema mejora igual que el componente.

- **Ley de Gustafson:**

$$A = \frac{(T_a + P_x \cdot T_b)}{(Ta + T_b)}$$

2. Tema 2: Monitorización

- Sobrecarga:

$$Sobrecarga = \frac{\text{Tiempo de ejecuición del monitor}}{\text{Intervalo de medida}}$$

3. Tema 3: Referenciación

- Comparación de rendimiento de dos computadores (A vs B):

$$\frac{R_A}{R_B} = \frac{T_B}{T_A} = n$$

Donde n es el número de veces que es mejor el rendimiento de A frente a B.

- Intervalo de confianza:

$$\overline{x} \pm t_{1-\frac{\alpha}{2},n-1} \cdot \frac{S}{\sqrt{n}}$$

- **CPI** *Ciclos por instrucción*
  - Fórmula:

$$CPI = \frac{\text{Ciclos de reloj de CPU usados}}{\text{Instrucciones ejecutadas}} = \frac{\text{Tiempo de ejecución} \cdot \text{Frecuencia de reloj}}{\text{Instrucciones ejecutadas}}$$

  - Tiempo de ejecución:

$$\text{Tiempo de ejecución} = \frac{\text{Instrucciones ejecutadas} \cdot CPI}{\text{Frecuencia de reloj}}, CPI = \sum_{i=1}^{n} t_i \cdot f_i$$

  - Consideraciones:
    - Es un ratio de tiempo. Se podría utilizar la media armónica.
    - Valor medio que interesa minimizar. Valor mínimo es 1.
    - Depende de la organización y arquitectura.
    - Inconveniente: ignora el tiempo imprevisible que hace falta para sincronizar procesador y memoria cache (ciclos de espera, fallos en caché, ...)

- **MIPS** *Million of Instructions per Second*
  - **Nativos**

$$MIPS = \frac{\text{Instrucciones ejecutadas}}{\text{Tiempo de ejecución} \cdot 10^6} = \frac{\text{Frecuencia de reloj de la CPU}}{CPI \cdot 10^6}$$

    - Depende del juego de instrucciones y los MIPS medidos varían entre programas en el mismo computador

    - *"Meaningless Indicator of Processor Speed"*
  - **Relativos**
    Referidos a una máquina de referencia (proceso de normalización).

$$MIPS_{Relativos} = \left(\frac{\text{Tiempo de referencia}}{\text{Tiempo de ejecución} \cdot 10^6}\right) \cdot MIPS_{Referencia}$$

- **MFLOPS** *Million of Floating-Point Operations per Second*
  - **Nativos**

$$MFLOPS = \frac{\text{Operaciones de coma flotante ejecutadas}}{\text{Tiempo de ejecución} \cdot 10^6}$$

    El tiempo de ejecución: es el del programa; incluyendo el tiempo consumido por las instrucciones de enteros

    - Basado en operaciones y no en instrucciones
    - Consideración: El juego de instrucciones en coma flotante varía de una arquitectura a otra y, dentro de una misma arquitectura, cada instrucción tiene un tiempo distinto que puede variar según los operandos.
  - **Normalizados** *(Relativos)*

$$MFLOPS_{\text{Nomralizados}} = \frac{N \text{ Operaciones de complejidad } x + N \text{ Operaciones de complejidad } y}{\text{Tiempo de ejecución} \cdot 10^6}$$

    El tiempo de ejecución: es el del programa; incluyendo el tiempo consumido por las instrucciones de enteros

    - Consideran la complejidad de las operaciones en coma flotante.
      - Suma, resta, multiplicación, comparación, negación: poco costosas
      - División, raíz cuadrada: costosas
      - Trigonométricas: muy costosas

- **Medias:**
  - **Media Aritmética**
    - Fórmula:

$$S_A = \frac{\sum_{i=1}^{n} x_i}{n}$$

    - Fórmula ponderada *(% dedicado)*:

$$S_{A,W} = \frac{\sum_{i=1}^{n} p_i \cdot x_i}{n}$$

- Tiempos: Primero se hace la media aritmética y luego se normaliza. NO al revés.
- Porcentajes (%)
- Consideraciones:
  - Con valores extremos: Primero se eliminan los extremos y después se hace la media aritmética.

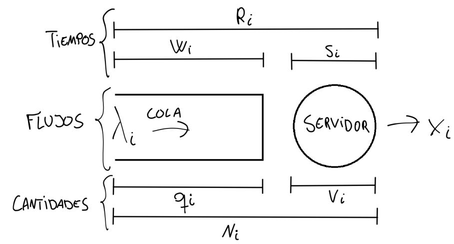
- **Media Armónica**
  - Fórmula:
$$S_H = \frac{n}{\frac{1}{x_i} + ... + \frac{1}{x_n}}$$
  - Fórmula ponderada:
$$S_{H,W} = \frac{1}{\frac{p_i}{x_i} + ... + \frac{p_n}{x_n}}$$
  - Ratios ($[x]/[y]$)
  - Posibles casos:
    - Frecuencias o velocidades

- **Media Geométrica**
  - Fórmula:
$$S_G = \sqrt[n]{x_1 \cdot x_2 \cdot x_2 \cdot ... \cdot x_n}$$
  - Fórmula ponderada:
$$S_{G,W} = (x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdot ... \cdot x_b^{\alpha_b} \cdot)^{\frac{1}{\alpha_1 + .. + \alpha_n}}, \textit{siendo } \alpha \textit{ la podrición}$$
  - No tiene unidad de medida.
  - Cuando se dispone de índices de rendimiento, se puede emplear la media geométrica, **pero a riesgo de reordenar los resultados.**

- Consideraciones:
  - La media geométrica no resulta apropiada para promediar tiempos ni ratios.
  - Pero sin embargo, su uso se puede usar para resumir valores normalizados porque mantiene un orden consistente cuando se comparan las prestaciones de varios computadores, aunque a veces este orden puede resultar erróneo. El adjetivo consistente significa en este contexto que el orden es el mismo.

- **Consideraciones Generales**
  - La máquina de referencia no puede ser parte del experimento de comparación.
  - La mejor alternativa a seguir consiste en hacer una media y posteriormente normalizar el valor obtenido.

4. Tema 4: Introducción al análisis operacional



- **Variables básicas**
  - T Duración del periodo de medida (*time*)
  - A_i Número de trabajos que llegan (*arrivals*)
  - C_i Número de trabajos que se van (*completions*)
  - B_i Tiempo de ocupación (*busy time*)
- **Variables deducidas**
  - U_i Utilización (*utilization*)
  - λ_i Tasa de llegadas (*arrival rate*)
  - X_i Productividad (*throughput*)
  - S_i Tiempo de servicio (*service time*)
  - V_i Razón de visita (*visit radio*)
  - D_i Demanda de servicio (*service demand*)
- **Otras variables de una estación**
  - R_i Tiempo de respuesta (*response time*)
  - W_i Tiempo de espera en cola (*waiting time*)
  - N_i Trabajos en toda la estación
  - Q_i Trabajos en cola de espera (*waiting customers*)
- **Variables del sistema básicas**
  - A_0 Número de trabajos que llegan (*arrivals*)
  - C_0 Número de trabajos que se van (*completions*)
- **Variables deducidas de sistema**
  - λ_0 Tasa de llegadas (*arrival rate*)
  - X_0 Productividad (*throughput*)

- **Ecuaciones**

$$\lambda_i = \frac{A_i}{T} \qquad V_i = \frac{C_i}{C_0}$$

$$X_i = \frac{C_i}{T} \qquad D_i = V_i \cdot S_i$$

$$S_i = \frac{B_i}{C_i} \qquad N_i = Q_i + U_i$$

$$U_i = \frac{B_i}{T} \qquad R_i = W_i + S_i$$

- **Hipótesis del equilibrio de flujo**
  - La tasa de llegada coincide con la tasa de salida $\lambda_i = X_i$
  - El número de trabajos que llegan con el que sale $A_i = C_i$
  - $A_i == C_i \rightarrow \lambda_i == X_i$
- **Ley de flujo forzado**
  - $X_i = X_0 \cdot V_i$
  - Decimos que la Ley de Flujo es forzada porque igualar la tasa de llegada de trabajos con la tasa de salida es, en sí mismo, algo difícil de que suceda. Por eso mismo se realiza con intervalos de observación suficientemente largos, para que se cumpla el supuesto equilibrio de flujo.
- **Ley de Little**
  - Relaciona el número de trabajos en el sistema con el tiempo de permanencia y su productividad o tasa de llegada.
  - Nos permite juntar tiempos, flujos y cantidades.

$$N = Flujo \cdot Tiempo$$

$$N_i = X_i \cdot R_i \qquad Q_i = X_i \cdot W_i$$

- **Ley de la Utilización**

$$U_i = \frac{B_i}{T} = \frac{C_i B_i}{T C_i} = X_i S_i \qquad U_i = X_i S_i$$

- **Ley general del tiempo de respuesta**
  - En general $R \neq R_1 + R_2 + ... + R_K = \sum_{i=1}^{K} R_i$
  - En particular $R = V_1 \cdot R_1 + V_2 \cdot R_2 + ... + V_K \cdot R_K = \sum_{i=1}^{K} V_i \cdot R_i$

- **Ley del tiempo de respuesta interactiva**

$$R = (\frac{N}{X}) - Z$$

- **Fórmulas ejercicios**
  - **Utilización:**

    $$U_i = X_i \cdot S_i$$

  - **Tiempo de respuesta:**
    - Tiempo de respuesta:

      $$R_0 = \sum V_i \cdot R_i$$

    - Tiempo de respuesta especifico:

      $$R_i = W_i + S_i$$

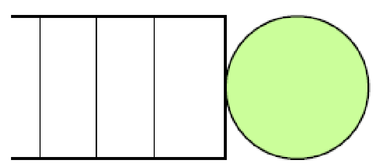
      - De una estación:

        $$R_i = \frac{S_i}{1 - U_i}$$

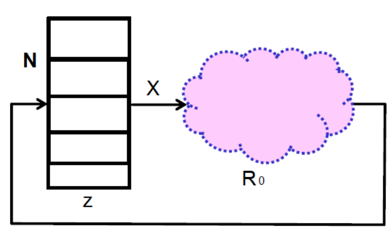
  - Ley de flujo forzado:

    $$X_i = X_0 \cdot V_i$$

- **Tipos de modelos:**
  - Modelo abierto:



  - Modelo cerrado:



    - Ley de Little:
      - $U_i = X_i \cdot S_i$
      - $N = X \cdot (R + Z)$

5. Tema 5: Aplicación del análisis operacional

- **Resolución de redes colas**
  - Utilización de cada estación: $U_i = X_i S_i = \lambda \cdot V_i \cdot S_i$
  - Tiempo de respuesta de cada estación: $R_i = \dfrac{S_i}{1 - U_i}$
  - Tiempo de respuesta del sistema: $R_o = \sum_{i=1}^{k} V_i R_i = \sum_{i=1}^{k} (\dfrac{V_i S_i}{1 - U_i})$
- **Cuello de botella**
  - $U_i = X_i S_i = X \cdot V_i \cdot S_i = X \cdot D_i$
- **Sistema equilibrado**
  - $U_i = U_j$
- **Límites optimistas: sistemas abiertos**
  - Tiempo de respuesta: $R_{opt} = \sum_{i=1}^{K} D_i = D$
  - Productividad: $X_{opt} = \dfrac{1}{D_b}$

    Limitada por el dispositivo cuello de botella (Utilización = 1)
- **Límites optimistas: sistemas cerrados**
  - Tiempo de respuesta:

$$R_{opt} = Max\{\sum D_i, N \cdot D_b - Z\}$$

$$f(x) = X \cdot D_b - Z$$

  - Productividad:

$$X_{opt} = Min\{\frac{1}{D_b}, \frac{N}{\sum D_i + Z}\}$$

- **Punto teórico de saturación**
  - $N^* = \lceil \frac{D+Z}{D_b} \rceil$

6. Tema 6: Caracterización de la carga

- **Clustering: escalado**
  - $x'_{ik} = \dfrac{x_{ik} - \overline{x}_k}{s_k}$
- **Clustering: distancia métrica**
  - Distancia euclídea entre dos componentes: $d = \left\{ \displaystyle\sum_{k=1}^{n} (x_{ik} - x_{jk})^2 \right\}^{0,5}$
  - Distancia euclídea ponderada: $d = \displaystyle\sum_{k=1}^{n} \left\{ a_k (x_{ik} - x_{jk})^2 \right\}^{0,5}$
  - Distancia chi-cuadrado: $d = \displaystyle\sum_{k=1}^{n} \left\{ \frac{(x_{ik} - x_{jk})^2}{x_{ik}} \right\}$

7. Tema 7: Planificación de la capacidad

- **Regresión Lineal**
  - Fórmula de la ecuación lineal: $y = a + b \cdot x$
  - $b = \dfrac{\displaystyle\sum_{i=0}^{n} x_i \cdot y_i - n \cdot \overline{x} \cdot \overline{y}}{\displaystyle\sum_{i=0}^{n} x_i^2 - n \cdot \overline{x}^2}$
  - $a = \overline{y} - b \cdot \overline{x}$
  - Aplicaciones
    - Datos no estacionales que muestran una tendencia
    - Datos históricos que muestran un patrón de evolución lineal
  - Consideraciones: No es necesario tener datos anteriores para predecir los futuros.
- **Medias Móviles**
  - $f_{t+1} = \dfrac{y_t + y_{t-1} + ... + y_{t-n+1}}{n}$
  - El error se calcula con la siguiente ecuación: $\dfrac{1}{n} \displaystyle\sum_{t=1}^{n} (y_t - f_t)^2$
  - Aplicaciones
    - Datos casi estacionarios.
    - Predicciones a corto plazo.

- Consideraciones
    - La exactitud alcanzada es normalmente alta
    - Una serie temporal se considera estacionaria cuando no hay cambio sistemático ni en la media ni en la varianza.

- **Suavizado Exponencial**
    - $f_{t+1} = f_t + \alpha(y_{t+1} - f_t)$
    - Aplicaciones
        - En tendencias históricas.
        - Datos estacionales que no muestran una tendencia sistemática.
    - Consideraciones: Realiza una media ponderada de las observaciones pasadas y la presente para predecir la futura.

Universitat de les Illes Balears