

Examen Febrero 2023

Ingeniería de Servidores

Ismael Sallami Moreno

Universidad de Granada
Doble Grado en Ingeniería Informática y ADE
<https://elblogdeismael.github.io>

2.- (1,2 puntos) Cuestiones (0,4 puntos cada pregunta).

- a) ¿Qué tipo de información obtenemos cuando ejecutamos “sar -d”, sin más argumentos? ¿Cuál es la diferencia con respecto a “sar -b”? ¿Y con respecto a “sar -d 10”?

Cuando ejecutamos *sar -d* obtenemos las transferencias de cada unidad de almacenamiento, mientras que con *sar -b* obtenemos las globales de todas las unidades de almacenamiento. Con *sar -d 10* obtenemos la información de las prestaciones de cada disco recopilada de forma interactiva cada 10s.

- b) ¿Qué tipo de información podemos obtener con “perf report”? ¿Y con “perf annotate”?

Con *perf report* se analiza *perf.data* y muestra las estadísticas generales. Mientras que con *perf annotate* se analiza *perf.data* y muestra los resultados a nivel de código ensamblador y código fuente (si está disponible).

- c) Indique qué se entiende por escalabilidad y de qué manera podríamos mejorar la escalabilidad de un monitor de actividad de servidores distribuidos.

La escalabilidad es la capacidad de un sistema para mantener su rendimiento al aumentar la carga o el número de servidores. Para mejorar la escalabilidad de un monitor de actividad de servidores distribuidos se pueden aplicar varias estrategias: reducir la sobrecarga usando muestreo adaptativo, distribuir la monitorización mediante arquitecturas jerárquicas o descentralizadas, filtrar y agregar datos localmente antes de enviarlos, emplear almacenamiento eficiente como bases de datos de series temporales, y paralelizar tareas internas del monitor. Estas medidas permiten que el monitor siga funcionando eficientemente aunque crezca el entorno monitorizado.

3.- (0,8 puntos)

Un programa monohebra consume el 30 % del tiempo en operaciones de disco duro y el resto en operaciones de coma flotante. Tras mejorar 2 veces el procesador que realiza las operaciones de coma flotante, ahora tarda 65 segundos en ejecutarse.

- a) ¿Cuánto tardaba en ejecutarse el programa antes de la mejora? Calcule, de paso, la ganancia en velocidad obtenida (con 2 cifras decimales) y exprese la mejora como tanto por ciento de mejora. **(0,4 puntos)**.

$$T_o = T_{dd} + T_{cf}$$

$$T_{dd} = 0,3 \cdot T_o$$

$$T_{cf} = 0,7 \cdot T_o$$

$$T_m = T_{dd} + \frac{T_{cf}}{2}$$

$$T_m = 65s$$

$$65 = 0,3 \cdot T_o + \frac{0,7 \cdot T_o}{2} \rightarrow T_o = 100s$$

$$\text{Ganancia en velocidad} = \frac{T_o}{T_m} = \frac{100}{65} = 1,5385 \approx 1,54$$

$$\text{Mejora} = (S - 1) \cdot 100 = (1,54 - 1) \cdot 100 = 54 \%$$

- b) ¿Es mejor aplicar esta mejora o triplicar la velocidad del disco duro? Razone la respuesta calculando tanto el nuevo tiempo mejorado como la ganancia en velocidad obtenida (2 decimales). **(0,4 puntos)**.

$$T_{m-nuevo} = \frac{T_{dd}}{3} + T_{cf} = \frac{30}{3} + 70 = 10 + 70 = 80s$$

$$T_{m-antiguo} = 65s$$

$$\text{Ganancia en velocidad} = \frac{T_o}{T_{m-nuevo}} = \frac{100}{80} = 1,25 \rightarrow 25 \%$$

Esta mejora supone el 25 % de mejora, mientras que la mejora del procesador supone un 54 %. Por tanto, es mejor mejorar el procesador.

4.- (1 punto)

La empresa Bi4Group está estudiando dos propuestas, que llamaremos A y B, con el objetivo de actualizar las unidades de estado sólido (SSD) de los computadores de su instalación informática. El ingeniero informático jefe de la empresa ha mandado ejecutar cinco de los programas que utilizan habitualmente en un computador con una unidad de cada propuesta y ha obtenido los tiempos de ejecución que se muestran a continuación:

Programa	tA (s)	tB (s)
1	156	150
2	6	7.5
3	128	125
4	46	43
5	95	89

df	0.20	0.10	0.05	0.02	0.01	0.001	2-tail alpha
1	3.0777	6.3138	12.7062	31.8205	63.6567	636.6192	
2	1.8856	2.9200	4.3027	6.9646	9.9248	31.5991	
3	1.6377	2.3534	3.1824	4.5407	5.8409	12.9240	
4	1.5332	2.1318	2.7764	3.7469	4.6041	8.6103	
5	1.4759	2.0150	2.5706	3.3649	4.0321	6.8688	
6	1.4398	1.9432	2.4469	3.1427	3.7074	5.9588	

Determine de forma razonada si existen diferencias significativas, para un nivel de confianza del 90 %, en el rendimiento de las dos unidades propuestas. AYUDA: La desviación típica muestral de tA-tB es 3,1s. **Nota:** Debe mostrar TODOS los cálculos que haga. Por poner solamente el resultado final, aunque sea correcto, no se obtendrá ninguna puntuación. No hace falta indicar qué propuesta es la mejor si las diferencias fueran significativas.

Debemos de realizar un test t donde la hipótesis nula es:

$$H_0 : \mu_{tA-tB} = 0$$

$$H_1 : \mu_{tA-tB} \neq 0$$

Es decir, que no hay diferencia significativa entre los tiempos de ejecución de las dos propuestas. Calculamos la media de las diferencias:

$$\bar{d} = \frac{(156 - 150) + (6 - 7,5) + (128 - 125) + (46 - 43) + (95 - 89)}{5} = \frac{6 - 1,5 + 3 + 3 + 6}{5} = \frac{16,5}{5} = 3,3s$$

Calculamos el error estándar de la media:

$$SE = \frac{s}{\sqrt{n}} = \frac{3,1}{\sqrt{5}} \approx 1,386s$$

Calculamos el estadístico t:

$$t = \frac{\bar{d} - 0}{SE} = \frac{3,3 - 0}{1,386} \approx 2,38$$

Nota: El valor de \bar{d}_{real} es 0 debido a que estamos suponiendo la hipótesis nula de que no hay diferencia significativa entre los tiempos de ejecución de las dos propuestas.

Si el nivel de confianza es del 90 %, el grado de significatividad de α es 0.10, y como tenemos 5 muestras, los grados de libertad son $df = n - 1 = 5 - 1 = 4$. Consultando la tabla de t de Student, para $df = 4$ y $\alpha = 0,10$, el valor crítico es $t_{crit} = 2,132$. Por lo

tanto, si fuera cierta nuestra hipótesis nula el valor de t debería de estar contenido en el intervalo $[-2,132, 2,132]$. Como no lo está, rechazamos la hipótesis nula y concluimos que hay diferencias significativas entre los tiempos de ejecución de las dos propuestas. *Por tanto, podemos afirmar que existen diferencias significativas en el rendimiento de las dos unidades propuestas al 90 % de confianza.*

5.- (2 puntos)

Un servidor web recibe, por término medio, 4 peticiones de páginas web por segundo. Los valores medios de los tiempos de servicio y de las utilizaciones de los dispositivos que más influyen en el rendimiento de este servidor web se indican en la siguiente tabla:

Dispositivo	Tiempo de servicio (s)	Utilización (%)
CPU	0,01	32
Disco A	0,04	64
Disco B	0,03	36

A partir de la información anterior responda a las siguientes cuestiones (**la respuesta se considerará incorrecta si no se justifica adecuadamente**):

- a) ¿Cuál es el cuello de botella de este servidor? ¿Está el servidor saturado? (0,4 puntos).

El dispositivo de cuello de botella es aquel que tiene la mayor utilización, en este caso el Disco A con un 64 %. El servidor no está saturado porque la utilización del Disco A es menor al 100 %.

- b) ¿Cuántos accesos se hacen, de media, al Disco A por cada página web servida? (0,4 puntos).

Debemos de calcular V_{DiscoA} .

$$\text{Ley de Utilización-Demanda: } U_{DiscoA} = X_0 \cdot D_{DiscoA}$$

$$\text{Fórmula de la Demanda: } D_{DiscoA} = V_{DiscoA} \cdot S_{DiscoA}$$

Como estamos en equilibrio de flujo podemos decir que $X_0 = \lambda_0$.

$$D_{DiscoA} = \frac{U_{DiscoA}}{X_0} = \frac{0,64}{4} = 0,16 \quad \text{De la ley de Utilización-Demanda.}$$

$$V_{DiscoA} = \frac{D_{DiscoA}}{S_{DiscoA}} = \frac{0,16}{0,04} = \boxed{4} \quad \text{De la fórmula de la Demanda.}$$

- c) ¿Qué valor tendría que tener la tasa de llegada para que el cuello de botella fuese otro dispositivo diferente del actual? (0,2 puntos).

Se ha mencionado que el cuello de botella actual es el Disco A, con una utilización del 64 %. Partiendo de la fórmula de la utilidad máxima de un dispositivo:

$$X_0^{max} = \frac{1}{D_{DiscoA}} = \frac{1}{V_{DiscoA} \cdot S_{DiscoA}}$$

Vemos que aquí la tasa de llegada no influye en nada, por ende, podemos terminar el apartado diciendo que no influye en nada la tasa de llegada para que el cuello de botella sea otro dispositivo diferente del actual.

- d) ¿Qué tiempo de servicio debería tener el dispositivo cuello de botella actual para multiplicar por dos la productividad máxima del servidor? Demuestre numéricamente la respuesta. (0,4 puntos).

Antes de resolver el ejercicio podemos pensar que es extraño que un solo dispositivo pueda multiplicar por dos la productividad máxima del servidor, pero vamos a resolverlo.

Sabemos que el cuello de botella actual es el Disco A, con una utilización del 64 % y un tiempo de servicio de 0.04s. La productividad máxima del servidor se calcula como:

$$X_0^{max} = \frac{1}{D_{DiscoA}} = \frac{1}{V_{DiscoA} \cdot S_{DiscoA}} = \frac{1}{4 \cdot 0,04} = 6,25tr/s$$

Si actuamos solo sobre el Disco A para multiplicar por dos la productividad máxima del servidor, su demanda se ve aumentada, lo que se traduce en una disminución de su tiempo de servicio a la mitad, haciendo que el cuello de botella sea el Disco B.

De esta manera se limita la productividad máxima del servidor a:

$$X_0^{max} = \frac{1}{D_{DiscoB}} = \frac{1}{U_{DiscoB}/X_0} = \frac{1}{0,36/4} = 11,11tr/s$$

Por tanto, actuando solo sobre el tiempo de servicio del Disco A no podemos llegar a la productividad máxima del servidor que se pide.

- e) Calcule el cambio en R_0^{min} y en X_0^{max} del servidor si eliminásemos el Disco A e hiciéramos, por tanto, que todos los accesos a disco se tuvieran que hacer al Disco B (basta con que indique los valores de R_0^{min} y X_0^{max} , antes y después del cambio). (0,6 puntos).

Antes del cambio

Sabiendo que $D_i = U_i/X_0$ y que $V_i = D_i/S_i$, calculamos la tabla de D_i y V_i de cada dispositivo:

Dispositivo	D_i (s)	V_i (accesos/s)	U_i (%)
CPU	0.08	8	32
Disco A	0.16	4	64
Disco B	0.09	3	36

Sabemos que:

$$R_0^{min} = \sum_{i=1}^n D_i = D_{CPU} + D_{DiscoA} + D_{DiscoB} = 0,08 + 0,16 + 0,09 = 0,33s$$

Y que:

$$X_0^{max} = \frac{1}{D_{DiscoA}} = \frac{1}{V_{DiscoA} \cdot S_{DiscoA}} = \frac{1}{4 \cdot 0,04} = 6,25tr/s$$

Después del cambio

Tras eliminar el Disco A, todos los accesos que antes se dirigían a él (4 por página) ahora se redirigen al Disco B, que ya tenía 3 accesos por página. Por tanto, la nueva razón de visita del Disco B será:

$$V_{\text{DiscoB}} = 4 + 3 = 7$$

Y su tiempo de servicio sigue siendo:

$$S_{\text{DiscoB}} = 0,03 \text{ s}$$

La tabla de dispositivos tras el cambio es:

Dispositivo	S_i (s)	V_i
CPU	0.01	8
Disco B	0.03	7

Calculamos la nueva demanda de servicio de cada dispositivo:

$$D_{\text{CPU}} = 8 \cdot 0,01 = 0,08 \quad \text{y} \quad D_{\text{DiscoB}} = 7 \cdot 0,03 = 0,21$$

Por tanto, el nuevo tiempo de respuesta mínimo es:

$$R_0^{\min} = D_{\text{CPU}} + D_{\text{DiscoB}} = 0,08 + 0,21 = \boxed{0,29 \text{ s}}$$

Y la nueva productividad máxima del servidor será:

$$X_0^{\max} = \frac{1}{D_{\text{DiscoB}}} = \frac{1}{0,21} = \boxed{4,76 \text{ tr/s}}$$

Resumen final:

	Antes del cambio	Después del cambio
R_0^{\min}	0,33 s	0,29 s
X_0^{\max}	6,25 tr/s	4,76 tr/s

Nota: No podemos decir que al eliminar el Disco A, la utilidad del Disco B sea del 100 % (suma) porque no hemos calculado la utilización del Disco B tras el cambio. Sin embargo, podemos deducir que la utilización del Disco B ha aumentado, ya que ahora tiene más accesos por página. Al igual que con las *demandas*.