

La respuesta debe escribirse en el hueco existente a continuación de cada pregunta con letra clara. Cada respuesta incorrecta, ilegible o vacía no suma ni resta. En el caso de preguntas teóricas se valorará la capacidad de síntesis.

1o Una API REST utiliza un único hilo de ejecución para servir peticiones. Esta API se despliega en sendas máquinas bajo demanda en la nube de AWS. La tabla siguiente muestra las características de cada tipo de máquina y el tiempo de respuesta de una única petición a la API, en milisegundos.

	Memoria (GiB)	Núcleos	Coste (\$/hora)	t_1	t_2	t_3	t_4	t_5
t3.xlarge	16	4	0.408	94.92	95.34	93.45	96.23	92.12
a1.4xlarge	32	16	0.1664	92.65	93.67	91.89	95.24	92.17

a— (0.25 puntos) ¿Cuál es la aceleración de la máquina *a1.4xlarge* respecto a la máquina *t3.xlarge* al servir una única petición? Redondea a dos decimales.

b— (0.5 puntos) ¿Cuál es la productividad máxima de cada máquina medida en peticiones por segundo? Redondea a un único decimal.

t3.xlarge: a1.4xlarge:

Supongamos que fuese necesario alcanzar una productividad mínima de 40000 peticiones por segundo y solo pudiesen contratarse máquinas del mismo tipo. c— (0.75 puntos) ¿Cuál sería el tipo de máquina elegido para minimizar costes y cuántas instancias serían necesarias?

Tipo de máquina: Núm. instancias:

2o En un servidor dedicado al procesamiento de imágenes se instala una CPU superescalar con un ancho de emisión 6 a una frecuencia de 2GHz. Para medir el rendimiento del servicio de procesamiento se ejecuta un benchmark que obtiene de media 5 segundos de tiempo de respuesta. Analizando la ejecución del benchmark se obtiene que la CPU se utiliza un 55% del tiempo, la GPU un 35% y el resto corresponde a tiempo de entrada/salida.

a— (0.75 puntos) ¿Cuál es la aceleración obtenida si se sustituye la GPU por una 3 veces más rápida? Escribe la expresión matemática utilizada para el cálculo.

b— (0.25 puntos) ¿Cuál sería el tiempo de respuesta del benchmark con la mejora anterior?

3o Se ejecuta el siguiente programa MIPS64 sobre una microarquitectura segmentada básica con una unidad de multiplicación/división de enteros de 5 ciclos no segmentada y un predictor de saltos siempre no tomado. No están soportadas la evaluación agresiva de saltos ni las excepciones precisas.

Se acompaña un fichero Excel de apoyo para facilitar la realización de este ejercicio. Recuerda escribir tus respuestas en las hojas del examen.

```
daddi r5, r0, 9 daddi r3, r0, -
3 begin:
    beq r5, r0, end dadd r5,
    r5, r3 dmul r1, r3, r4 j
    begin
end:
    sltu r4, r1, r4
```

a— (1 punto) Enumera las dependencias de datos que existen en el programa, especificando el tipo junto con las instrucciones y registros involucrados. Ejemplo de respuesta: RAW => daddi, ori : r7 // dsub, xori : r3.

RAW =>

WAW =>

WAR =>

b— (1.5 puntos) Rellena todas las filas del siguiente cronograma (sin añadir nuevas columnas) con la evolución de las primeras instrucciones del programa sobre el *pipeline*.

Instrucción \ Ciclo	1	2	3	4	5	6	7	8	9	10	11	12	13	14
daddi r5, r0, 9	IF													

c— (1 punto) Indica las duraciones en ciclos de reloj de las primeras detenciones de cada tipo que se producen al ejecutar el código. Para los tipos RAW, WAW y estructural debe indicarse además la instrucción que se detiene, mientras que para el tipo control la instrucción que la provoca. Si un tipo de detención no aparece, indícalo con N/A. Ejemplo de respuesta, RAW: xor 1 ciclo.

RAW:

WAW:

Estructural:

Control:

d— (1 punto) ¿Cuántos ciclos de reloj son necesarios para ejecutar el código anterior? ¿Cuál es el CPI ignorando el transitorio inicial? Indica la expresión matemática utilizada para obtener el CPI. Redondea a dos decimales.

Ciclos:

CPI:

e— (0.5 puntos) Indica las dos primeras rutas de reenvío que se activarían si todas las posibles estuviesen implementadas en esta microarquitectura. Ejemplo de respuesta: Salida EX daddi → Entrada EX dmul.

f— (1 punto) Si se implementase renombrado de registros en esta microarquitectura, ¿cuáles serían los 4 primeros renombrados que se producen. Para el renombrado se dispone de una cola FIFO que originalmente contiene los registros rr32 a rr63 y a los que se van añadiendo los registros disponibles

Ejemplo de respuesta para esta pregunta: Registro r7 se asigna a rr43.

g— (0.5 puntos) Teniendo en cuenta únicamente el salto condicional, para el programa completo, ¿cuántas veces en total acierta la predicción el predictor siempre no tomado? ¿Y cuántas veces en total falla la predicción?

Núm. aciertos:

Núm. fallos:

h— (0.5 puntos) Si el predictor de saltos siempre no tomado se sustituye por un predictor dinámico de 2 bits con el valor 01 por defecto del historial (*weak not taken*), ¿cuántos ciclos de reloj se detendría el *pipeline* en la ejecución del programa completo debido a la instrucción beq y a la instrucción j?

beq:

j:

i— (0.5 puntos) Si el programa se ha cargado a partir de la dirección de memoria 3020h, ¿cuáles serán los valores de las entradas en la tabla BHT+BTB correspondientes a las instrucciones de salto al finalizar la ejecución del programa?

Dirección	Destino	Historial