

EXAMEN INTRA INF2500 – Session A-2001 - Réponses

Pondération sur 100 points, possibilité de 22 points extra.

Question 1 (5 pts):

Vous voulez diminuer le trafic entre plusieurs ordinateurs connectés sur le même réseau informatique. Ces ordinateurs partagent leur mémoire de façon à former une grande mémoire disponible et répartie sur le réseau. Que recommandez-vous pour établir un système de haute performance sachant que la conception du réseau ne permet que de petits messages à la fois entre les ordinateurs et spécifier ce que l'on devrait considérer pour déterminer la longueur de chaque message sur le réseau?

- Ajouter des bulles au pipeline à l'intérieur des processeurs pour chaque ordinateur et la longueur du message sera déterminée par la longueur du pipeline;
- Ajouter une autre mémoire cache à l'interface réseautique de chaque ordinateur et la longueur du message sera déterminée par la localité spatiale;
- Diminuer la mémoire cache existante dans chaque ordinateur et la longueur du message pourra être déterminée par la localité temporelle;
- Diminuer la fréquence de l'horloge de chaque ordinateur et la longueur du message pourra être variable;
- Toutes les solutions proposées sont excellentes.

Réponse : b.

Question 2 (5 pts):

Vous disposez de 2 pièces (Integrated Circuits) de mémoire de type SRAM de 64 Kilo-octets (K-bytes) par pièce avec un temps d'accès de 25 nanosecondes (ns). Vous disposez également de 4 pièces de mémoire de type DRAM de 8 Mega-octets (M-bytes) avec un temps d'accès de 120 ns. Il n'y a aucun disque ou autre mémoire disponible. Quelle configuration donnera la meilleure performance?

- Utiliser les 2 SRAMs comme une seule mémoire cache et les 4 DRAMs pour la mémoire principale avec des blocs de 8 M-octets;
- Utiliser les 2 SRAMs comme une seule mémoire cache et les 4 DRAMs pour la mémoire principale avec des blocs de 128 K-octets;
- Utiliser 1 SRAMs comme une mémoire cache de niveau 1 et l'autre SRAM pour une mémoire cache de niveau 2 avec les 4 DRAMs pour la mémoire principale avec des blocs de 64 K-octets;
- Utiliser seulement 1 SRAMs comme une mémoire cache avec les 4 DRAMs pour la mémoire principale avec des blocs de 64 K-octets;

- e. Utiliser seulement les DRAMs comme mémoire principale et ainsi éviter les pénalités causées par les accès manqués à la mémoire cache et ainsi les dimensions des pages ne s'appliquent pas.

Réponse : b.

Question 3 (5 pts):

Le nombre de cycles horloge pour l'exécution de chaque instruction est une moyenne qui dépend des détails de l'organisation de l'ordinateur et du découlement (stream) des instructions. Quelles sont les variables dans les ordinateurs qui peuvent faire varier le nombre de cycles requis pour exécuter une instruction?

1. Un manquement ou échec (miss) au cache;
 2. L'introduction de bulles dans le pipeline pour éviter un aléa de donnée (data hazard);
 3. Un succès (hit) au cache;
-
- a. 1 seulement;
 - b. 2 et 3;
 - c. 1, 2, et 3;
 - d. 3 seulement;
 - e. 1 et 2;

Réponse : e.

Question 4:

En faisant référence aux instructions et format du tableau 2.5 du livre de référence, nous avons les deux instructions suivantes utilisant le mode d'adressage registre. Ce court programme basé sur des instructions à 2 opérandes et qui additionne les valeurs contenues dans les registres R1, R2, et R3, et ceux de R4 et R5 de façon indépendante, causera un aléa (hazard) dans le pipeline.

Question 4A (2 pts): Nommez le type d'aléa rencontré durant l'exécution des 3 lignes d'instructions.

Réponse: Aléa de donnée

Question 4B (7 pts) : Ensuite, sans l'addition de bulles or d'instructions NOP (No OPeration), reformuler le code de façon à éviter des aléas en sachant qu'il y a seulement 5 registres de disponibles (R1 à R5) et que le délais est de une instruction. Vous devez aussi inclure deux des modes d'adressage suivant : immédiat, indexé, directe, autoincrément. Noter que R2 contient une constante = 54, que R3 est une variable mais qu'il y a une autre copie de la même variable à l'adresse de mémoire 1043.

ADD R1, R2
ADD R3, R1
ADD R4, R5

Réponse: ADD R1, #54
 ADD R4, R5
 ADD R1, (1043)

Note : Plusieurs étudiants ont donné la réponse exacte, cependant nous avons également accordé les points à ceux qui ont donné une réponse sensiblement semblable à celle-ci.

Question 5:

Nous avons une opération du genre $A + B + C + D$ avec un UAL (ALU) capable d'additionner un maximum de 2 opérandes. La séquence du pipeline est la suivante où à la séquence 1 : la variable A entre dans le pipeline, à la séquence 2 : B entre dans le pipeline et est additionné à A, etc.

1. A
2. $A + B$
3. $(A + B) + C$
4. $(A + B + C) + D$

Question 5A (5 pts): Dans l'exécution de $A + B + C + D$, déterminer le nombre d'envois (forwarding) dans le pipeline et indiquer la ou les numéros de séquences (si applicable) où il y a l'utilisation d'un envoi.

Réponse : Parce que l'UAL peut additionner 2 opérandes à la fois, il y a donc 2 envois, une à la séquence 3 et l'autre à la séquence 4. Cependant, nous avons également accordé tous les points à ceux qui n'ont pas fait attention au fait que l'UAL pouvait additionner 2 opérandes à la fois et qui ont donc indiqué 3 envois, aux séquences 2, 3, et 4.

Question 5B (5 pts): Ensuite, si vous avez une opération du genre $(A + B + C + D) * E$ où la multiplication prend 2 cycles pour exécuter, que recommander vous de faire et pourquoi?

Réponse : Nous avons été très flexible sur cette question car il y a plusieurs possibilités. Nous avons donné tous les points aux réponses plausibles, aussi simple que : ajouter une bulle dans le pipeline pour éviter un aléas structurel et continuer l'utilisation d'envois (forwarding), jusqu'aux réponses beaucoup plus complexes.

Question 6 :

Utiliser le morceau de code suivant :

ADD R3,R4,R5
 LW R2,100(R1)
 AND R6,R2,R7
 SW 100(R1),R6
 SW 200(R0),R3

Question 6A (5 pts). Montrer le comportement temporel de cette séquence d'instructions sur le pipeline entier de DLX **avec envoi généralisé** et en supposant qu'une lecture registre et une écriture registre dans le même cycle machine ne provoque pas d'aléa structurel (envoi registre-registre). Combien de cycles prends cette séquence pour s'exécuter (2 pts)?

Réponse :

	1	2	3	4	5	6	7	8	9	10
ADD R3,R4,R5	LI	DI	EX	MEM	ER					
LW R2,100(R1)		LI	DI	EX	MEM	ER				
AND R6,R2,R7			LI	DI	Susp.	EX	MEM	ER		
SW 100(R1),R6				LI	Susp.	DI	EX	MEM	ER	
SW 200(R0),R3					Susp.	LI	DI	EX	MEM	ER

10 cycles.

Question 6B (5 pts): Afin de réduire la taille du processeur, nous avons retiré le matériel qui calcul le nouveau compteur programme ($NCP = CP + 4$). L'UAL (ALU) sera maintenant utilisé pour faire ce calcul. Reprendre l'exercice a) en tenant compte de cette nouvelle contrainte.

Réponse : (Aléa structurel entre LI et EX)

	1	2	3	4	5	6	7	8	9	10	11	12	13
ADD R3,R4,R5	LI	DI	EX	ME	ER								
LW R2,100(R1)		LI	DI	EX	ME	ER							
AND R6,R2,R7			Susp.	Susp.	LI	DI	EX	ME	ER				
SW 100(R1),R6						LI	DI	EX	ME	ER			
SW							Susp	Susp	LI	DI	EX	ME	ER

200(R0),R3						.	.				M	
------------	--	--	--	--	--	---	---	--	--	--	---	--

13 cycles.

Question 6C (5 pts): Ordonnancez (re-schedule) les instructions de la séquence afin d'optimiser les performances du pipeline.

Réponse : LW R2,100(R1)
 ADD R3,R4,R5
 AND R6,R2,R7
 SW 200(R0),R3
 SW 100(R1),R6

Question 7 (5 pts):

On compile puis on exécute une même tâche sur un ordinateur avec un jeu d'instructions réduit (RISC), puis sur un ordinateur avec un jeu d'instructions complet (CISC).

Encerchez les énoncés vraies :

- a) La tâche s'exécute plus rapidement sur la machine CISC.
- b) Un plus grand nombre d'instructions dynamiques seront exécutés par la machine CISC.
- c) La machine RISC est plus simple à concevoir.
- d) Le compilateur conçu pour la machine RISC est plus complexe.

Réponse : c)

Question 8:

Voici la liste des instructions d'un ordinateur fictif :

4 instructions UAL :	ADD Rd,Rs
	SUB Rd,Rs
	MULT Rd, Rs
	DIV Rd,Rs
2 instructions de chargement :	LOAD Rd, [Rs]
	LOAD Rd, Immédiat
1 instruction de rangement :	STORE [Rd], Rs
1 instruction de branchement :	BEQZ Rs, Immédiat

L'ordinateur possède 8 registres 32 bits et les valeurs immédiates sont encodées sur 32 bits.

Dans un premier temps, on utilise **un seul format** d'instruction de longueur fixe et à champ fixe. Énumérez les différents champs d'une instruction ainsi que le nombre de bits accordés à chaque champ. Quelle est la longueur totale d'une instruction ?

Réponse:

3 bits	3 bits	3 bits	32 bits
Code – Op	Rd	Rs	Immédiat

Longueur : 41 bits (ou 6 octets)

Question 8A (5 pts): Afin de réduire la longueur des programmes, on utilise **deux formats** d'instructions de longueur fixe à champ fixe. Décrivez ces deux formats.

Réponse:

Pour les 6 instructions qui n'utilisent pas de valeur immédiate :

3 bits	3 bits	3 bits
Code – Op	Rd	Rs

Longueur : 9 bits (ou 2 octets)

Pour les 2 instructions qui utilisent des valeurs immédiates :

3 bits	3 bits	32 bits
Code – Op	Rd/Rs	Immédiat

Longueur : 38 bits (ou 5 octets)

Question 8B (5 pts): Si on suppose qu'un programme typique utilise une proportion égale de chaque instruction, combien vaut le rapport Longueur du programme avec 2

formats ?

Longueur du programme avec 1 format

Réponse:

Longueur du programme avec 1 format:

$N_i * 41$ bits

OU

$N_i * 6$ octets

Longueur du programme avec 2 formats :

$$(6 / 8) Ni * 9 \text{ bits} + (2 / 8) Ni * 38 \text{ bits} = 16,25 Ni \text{ bits}$$

OU

$$(6 / 8) Ni * 2 \text{ octets} + (2 / 8) Ni * 5 \text{ octets} = 2,75 Ni \text{ octet}$$

Donc le rapport est

$$16,25 Ni / 41 Ni = 0,396 \text{ (Calculs en bits)}$$

OU

$$2,75 Ni / 6 Ni = 0,458 \text{ (Calculs en octets)}$$

Question 9:

Nommez **le ou les** modes d'adressage utilisés dans chaque situation.

Question 9A (2 pts): Ajouter 8 au contenu du registre R7.

Réponse: Immédiat et Registre

Question 9B (2 pts): Ajouter le contenu de l'adresse 300 sur le dessus d'une pile.

Réponse: Direct (ou Absolu) et Auto-Incrément

Question 9C (2 pts): Copier le contenu d'une variable locale à l'adresse spécifiée par le registre R9.

Réponse: Déplacement (ou basé) et Registre Indirect

Question 10 (5 pts):

Encerchez **le ou les** principes qui sont à la base de la hiérarchie mémoire:

- a) Rendre rapide le cas courant
- b) Le plus petit le plus vite
- c) La localité des références
- d) La loi d'Amdahl

Réponse: b) et c)

Question 11 – Suites d'essais et normalisation

Voici les temps mesurés pour 2 programmes. Ces programmes exécutent 100 millions d'opérations flottantes.

	Machine A	Machine B	Machine C
Programme 1	2 s	10 s	40 s
Programme 2	500 s	50 s	40 s
Total	502 s	60 s	80 s

Question 11A (5 pts): Calculer le débit en MFLOPS de chaque programme

Question 11B (5 pts): Calculer les moyennes arithmétiques, géométriques et harmoniques des MFLOP pour chaque machine

Question 11C (5 pts): Laquelle des 3 moyennes correspond le mieux aux performances de temps. Pourquoi ?

Réponses :

(basé sur 1.14)

Correction

(a) 5 points : 3 pour les réponses, 2 pour les calculs.

(b) 5 points : 3 pour les réponses, 2 pour les calculs.

	Machine A	Machine B	Machine C
Programme 1	2	10	40
Programme 2	500	50	40
Total	502	60	80
MFLOPS 1	50	10	2,5
MFLOPS 2	0,2	2	2,5
Moyenne Arithmétique	25,1	6,0	2,5
Moyenne Géométrique	3,2	4,5	2,5
Moyenne Harmonique	0,4	3,3	2,5
MFLOPS (temps total)	0,4	3,3	2,5

(c) La moyenne harmonique est plus appropriée car les MFLOPS sont une mesure de débit.

Correction :

5 points si la réponse indique qu'il faut utiliser la formule des moyennes harmoniques pour des débits, (que sont des MFLOPs).

Question 12 (10 pts) – Équation de performance de l'UC

Vous faites un changement à l'unité centrale pour réduire le nombre de chargements/rangements associés aux accès multiples à des données séquentielles.

- ? T cycle non optimisé est 10% plus rapide
- ? 30% des instructions dans la version non optimisée sont des chargements et des rangements
- ? la version optimisée exécute 50% des chargements/rangements de la version non optimisée
- ? Toutes les instructions prennent un cycle d'horloge.

Quelle version est plus rapide? Justifier avec des calculs.

Réponses :

T Cycle optimisé / T cycle non optimisé	1,1
Répartition LOAD/STORE non optimisé	30%
LOAD/STORE optimisé / LOAD/STORE non optimisé	1/2
Ratio de NI non optimisé / NI optimisé	1,176470588
Ratio de Temps non optimisé/Temps Optimisé	107%

Autres réponse :

$$\text{Temps Non optimisé} = (90\% * T \text{ cycle optimisé}) * NI \text{ Non Optimisé}$$

$$\text{Temps Optimisé} =$$

$$T \text{ cycle optimisé} * (70\% + 30\%/2)* NI \text{ Non Optimisé} =$$

$$(85\% * T \text{ cycle optimisé}) * NI \text{ Non Optimisé}$$

Autre réponse :

$$\text{Temps Non optimisé} = T \text{ cycle non optimisé} * NI \text{ Non Optimisé}$$

$$\text{Temps Optimisé} = T \text{ cycle non optimisé} * 110\% * (70\% + 30\%/2)* NI \text{ Non Optimisé} =$$

$$(85\% * 110\% * T \text{ cycle non optimisé}) * NI \text{ Non Optimisé} = 0.935 \text{ Temps non optimisé}$$

Correction :

2 points pour les équations; 2 points pour la justification; 6 points pour des calculs corrects

Question 13 – Notations des instructions

Référez à la figure 2.21 et aux notations de la page XVI (ou par exemple de la figure 2.23).

Pour le fragment de code suivant, donnez

- (a) le type d'instruction (I, R, J) (**3 pts**)
- (b) la valeurs numérique des champs d'opérandes rs1, rs2, rd et valeur immédiate (**3 pts**)
- (c) la signification des instructions (**4 pts**)

(ce code est donné au problème 3.1 du livre)

INSTRUCTION	Type	rs1	rs2	Rd	Valeur immédiate	Signification
Boucle : LW R1, 0(R2)						
ADDI R1,R1,#1						
SW 0(R2),R1						
ADDI R2,R2,#4						
SUB R4,R3,R2						
BNZ R4,Boucle						

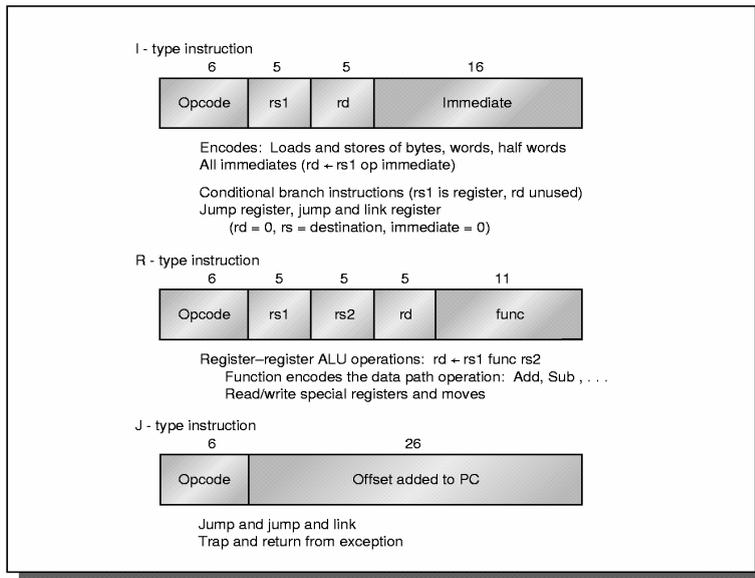


FIGURE 2.21 Instruction layout for RLV

Réponses :

INSTRUCTION	Type	rs1	rs2	rd	immediat	Signification
Boucle : LW R1, 0(R2)	I	2	1	X	0	Reg[1] ← Mem[0+Reg[2]]
ADDI R1,R1,#1	I	1	1	X	1	Reg[1] ← Reg[1]+1
SW 0(R2),R1	I	2	1	X	0	Mem[Reg[2]+0] ← Reg[1]
ADDI R2,R2,#4	I	2	2	X	4	Reg[2] ← Reg[2]+4
SUB R4,R3,R2	R	3	2	4	X	Reg[4] ← Reg[3]-Reg[2]
BNEZ R4,Boucle	I	4	X	X	-24	Si (Reg[4]≠0) CP ← Boucle (NCP-24)

Correction :

Types : 2 points ; 1 à 2 fautes : 1 point
rs1, rs2, rd : 3 points; 1 à 3 fautes : 2 points; 4 fautes : 1 point
Immédiat : 2 points; 1 à 2 fautes : 1 point

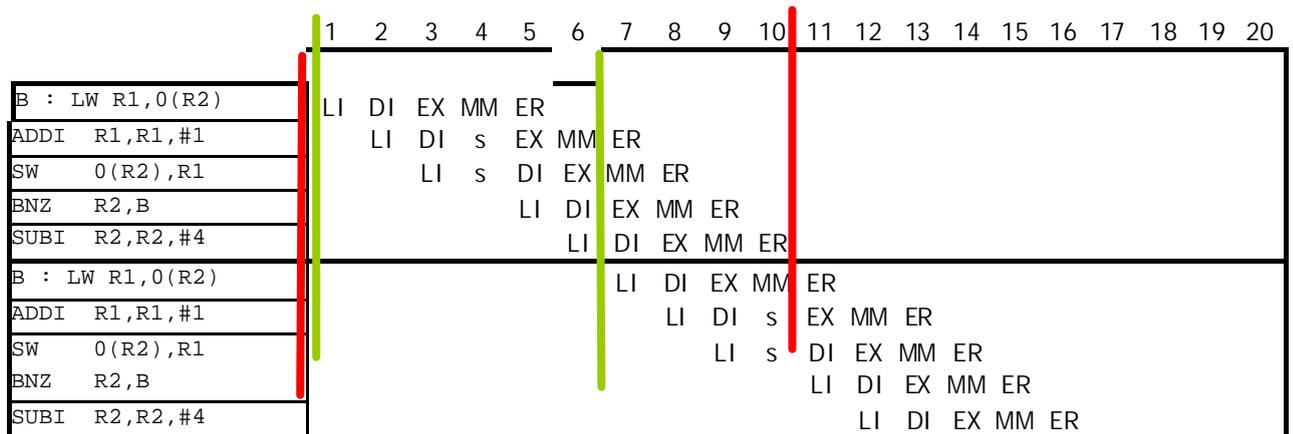
2 points pour le nombre de cycles; 1 point pour avoir déterminé 25 itérations; 1 point pour la formule

Question 15 (5 pts)

Montrez le comportement temporel de cette séquence avec matériel d'envoi DLX, et le branchement différé sans prédiction. Si la valeur initiale de R2=100, quel est le nombre total de cycles pour exécuter cette séquence?

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B : LW R1,0(R2)																				
ADDI R1,R1,#1																				
SW 0(R2),R1																				
BNZ R2,B																				
SUBI R2,R2,#4																				

Réponse :



nb itérations	26
cycles avant la boucle	0
cycles dans la boucle	6
cycles après la boucle	4
total	160

Pondération :

3 points pour le diagramme temporel; -1 point par erreur
 2 points pour le nombre de cycles; 1 point pour avoir déterminé 26 itérations; 1 point pour la formule