POLYTECHNIQUE MONTRÉAL

INF3610
Systèmes embarqués

Laboratoire 2

Soumis par :
Mohamed Ayman Atmani - 2006987
Ali Ameziane Hassani – 2007033
Arnaud Dessain – 1856013

Le 22 avril 2022

Architecture 1

Question 1 solution 1:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	7.256	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
1380101	1380101	1380101	1380101	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1	1380100	1380100	27602	-	-	50	no
+ L2	27600	27600	552	-	-	50	no
++ L3	550	550	11	-	-	50	no

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	-	0	138
FIFO	-	-	-	-
Instance	-	5	348	711
Memory	-	-	-	-
Multiplexer	-	-	-	116
Register	-	-	244	-
Total	0	5	592	965
Available	280	220	106400	53200
Utilization (%)	0	2	~0	1

Detail

Instance

Instance	Module	BRAM_18K	DSP48E	FF	LUT
mmult_hw_fadd_32nbkb_U1	mmult_hw_fadd_32nbkb	0	2	205	390
mmult_hw_fmul_32ncud_U2	mmult_hw_fmul_32ncud	0	3	143	321
Total	2	0	5	348	711

a) **Réponse :**

Les tableaux nous montrent dans l'onglet *Summary -> Latency* que le temps d'exécution total est de 1 380 101 coups d'horloge.

Le tableau 1 d'utilisation des ressources montre que cette solution utilise 5 DSP48E ce qui représente les ressources nécessaire à une seul opération de multiplication accumulation selon le tableau d'instance

b) **Réponse :**

Les tableaux nous montrent dans l'onglet *Loop -> Trip Count* que le nombre d'itérations est de 50 pour L1, 50 pour L2 et 50 pour L3. Comme les boucles sont

imbriquées les unes dans les autres, nous devons multiplier ces valeurs pour obtenir le nombre total de fois (50*50*50) que la ligne multiplication accumulation de la matrice est exécutée ce qui représente un temps d'exécution est en $O(N^3)$.

Le tableau d'utilisation des ressources montre que cette solution utilise 5 DSP48E ce qui représente les ressources nécessaires à une seule opération de multiplication accumulation selon le 2ème tableau.
On a O(const) en ressources vu qu'une seule opération de multiplication accumulation nécessite 5 DSP alors que le nombre de ressources utilisées par notre programme est de 5 ce qui explique O(const) en ressources.

c) **Réponse :**
Oui, cette solution respecte les limites de la carte par rapport aux ressources même si elle ne tire pas profit des ressources disponibles . Il est donc réalisable d'utiliser cette solution sur notre carte.

Question 2 solution 2:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	9.634	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
2757	2757	2757	2757	none

Detail

Instance

Loop

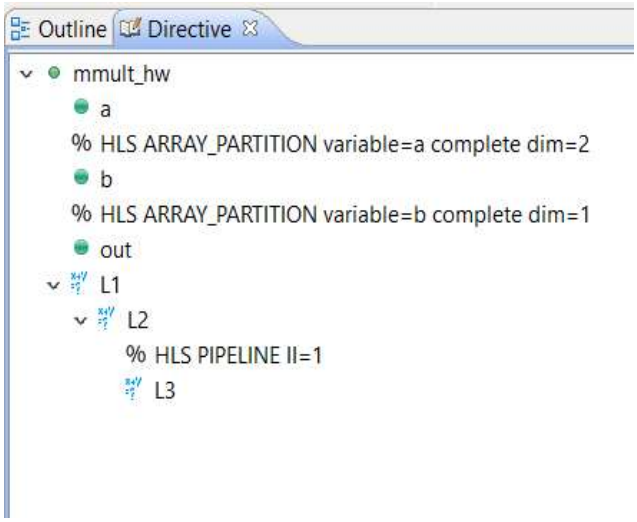
	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1_L2	2755	2755	257	1	1	2500	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	1	-	-
Expression	-	-	0	82
FIFO	-	-	-	-
Instance	-	250	17400	35550
Memory	-	-	-	-
Multiplexer	-	-	-	75
Register	0	-	7029	640
Total	0	251	24429	36347
Available	280	220	106400	53200
Utilization (%)	0	114	22	68

Choix des paramètres:



Justification de paramètre:
Afin d'adhérer au complexité attendu il nous a fallu pipeliner dans un premier lieu la troisième boucle ce qui implique qu'on devrait aussi partitionner notre array afin de permettre à notre programme d'effectuer l'opération de multiplication accumulation en parallèle en faisant des multiplication ligne * colonnes.

a) **Réponse :**
Les tableaux nous montrent dans l'onglet *Summary -> Latency* que le temps d'exécution total est de 2 757 coups d'horloge.
Le tableau d'utilisation des ressources montre que cette solution utilise 251 DSP48E.

b) **Réponse :**
Les tableaux nous montrent dans l'onglet *Loop -> Trip Count* que le nombre d'itérations est de 2500. Comme notre n = 50 et que 50*50 = 2500, nous avons donc que le temps d'exécution est en O(n^2) .
Le tableau d'utilisation des ressources montre que cette solution utilise 251 DSP48E.
On a O(N) en ressources vu qu'une seule opération de multiplication accumulation nécessite 5 DSP et que $5 * 50 = 250 \simeq 251$.

c) **Réponse :**
Cette solution n'est pas implémentable sur notre carte car elle requiert 251 DSP48E. Or, la carte en contient 220 ce qui n'est pas suffisant pour soutenir les besoins de cette solution.

Question 3 solution 3 :

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	9.634	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
2071	2071	2071	2071	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1_L2	2069	2069	222	1	1	1849	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	1	-	-
Expression	-	-	0	83
FIFO	-	-	-	-
Instance	-	215	14964	30573
Memory	-	-	-	-
Multiplexer	-	-	-	75
Register	0	-	6097	640
Total	0	216	21061	31371
Available	280	220	106400	53200
Utilization (%)	0	98	19	58

Réponse :

La valeur que nous avons trouvé est N = 43. Pour N = 43, nous avons 216 DSP48E utilisés. Pour N = 44, nous dépassons la limite de 220 DSP48E.

Pour cette solution, nous avons 2071 coup d'horloges pour le temps d'execution.

Question 4 solution 4:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	8.024	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
42	42	42	42	none

Detail

- Instance
- Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1	40	40	36	1	1	6	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	-	0	25
FIFO	-	-	-	-
Instance	-	180	12528	25596
Memory	-	-	-	-
Multiplexer	-	-	-	48
Register	0	-	5550	1056
Total	0	180	18078	26725
Available	280	220	106400	53200
Utilization (%)	0	81	16	50

choix des paramètres:

Outline | Directive

```
mmult_hw
  a
  % HLS ARRAY_PARTITION variable=a complete dim=2
  b
  % HLS ARRAY_PARTITION variable=b complete dim=0
  out
  % HLS ARRAY_PARTITION variable=out complete dim=2
  L1
    % HLS PIPELINE II=1
    L2
      L3
```

Justification des paramètres:

afin d'obtenir les une complexité de O(n) il faut pipeliner la 2ème boucle afin de garder seulement la première boucle avec n itérations ainsi il faudrait aussi faire un array partition afin de pouvoir d'effectuer l'opération en parallèle la matrice B sera complètement partitionné tandis que les matrices A et C sera partitionnés selon une seule la deuxième dimension.

a) **Réponse :**
   Cette solution prend 42 cycles d'horloge.
   Elle requiert 180 DSP48E

b) **Réponse :**
   Nous avons trouvé que le N le plus grand que nous pouvions avoir avec cette solution est N = 6. Pour la complexité temporelle, le nombre total d'itération est de 6 donc on rencontre notre objectif O(N) .

   Ensuite, nous avons établi que nous avions besoin de 5 DSP48E pour l'opération multiplication-accumulation. Comme nous avons n = 6, $6 * 6 * 5 = 180$ ce qui correspond au nombre de DSP48E utilisé pour cette solution. Cela démontre donc que la complexité en ressources est en $O(N^2)$.

<u>Architecture 2</u>
Question 5 solution 1:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	7.614	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
1380101	1380101	1380101	1380101	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1	1380100	1380100	27602	-	-	50	no
+ L2	27600	27600	552	-	-	50	no
++ L3	550	550	11	-	-	50	no

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	-	0	138
FIFO	-	-	-	-
Instance	-	4	200	152
Memory	-	-	-	-
Multiplexer	-	-	-	116
Register	-	-	180	-
Total	0	4	380	406
Available	280	220	106400	53200
Utilization (%)	0	1	~0	~0

Detail

Instance

Instance	Module	BRAM_18K	DSP48E	FF	LUT
mmult_hw_hadd_16nbkb_U1	mmult_hw_hadd_16nbkb	0	2	109	116
mmult_hw_hmul_16ncud_U2	mmult_hw_hmul_16ncud	0	2	91	36
Total	2	0	4	200	152

DSP48

a) Le temps d'exécution total est de 1 380 101 cycles d'horloge et 4 DSP48E sont utilisés.

b) Les résultats sont semblables à ceux de l'architecture no 1 pour la solution 1. Dans un premier temps, le temps d'exécution est prend le même nombre de cycle

d'horloge Pour ce qui est des ressources, nous utilisons 4 DSP48E contrairement à 5 pour la première architecture. l'utilisation n'a pas été coupé en deux

c) Oui, Cette solution est envisageable, pour ce qui est des ressources, la carte peut largement subvenir aux besoins des 4 DSP48E nécessaires pour cette architecture.

Question 6 solution 2:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	9.634	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
2757	2757	2757	2757	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1_L2	2755	2755	257	1	1	2500	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	1	-	-
Expression	-	-	0	82
FIFO	-	-	-	-
Instance	-	200	10000	7600
Memory	-	-	-	-
Multiplexer	-	-	-	75
Register	0	-	3829	640
Total	0	201	13829	8397
Available	280	220	106400	53200
Utilization (%)	0	91	12	15

Justifiez le choix des paramètres

La justification est la même que celle de la partie 1.

a) Le temps d'exécution est de 2757 cycles et 201 DSP48E sont utilisés.

b) Pour l'architecture 1 solution 2, nous utilisions 251 DSP48E pour N = 50 ce qui dépassait déjà les 220 DSP48E disponibles. Pour l'architecture 2, nous utilisons 201 DSP48E pour N = 50. Il nous reste donc 19 DSP48E disponibles. Alors, cette

architecture pourrait prendre un N plus grand que 50 étant donné qu'il reste des ressources inutilisées.

Question 7 solution 3:

Target device: xc7z020clg484-1

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	9.634	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
3193	3193	3193	3193	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1_L2	3191	3191	277	1	1	2916	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	1	-	-
Expression	-	-	0	82
FIFO	-	-	-	-
Instance	-	216	10800	8208
Memory	-	-	-	-
Multiplexer	-	-	-	75
Register	0	-	4105	640
Total	0	217	14905	9005
Available	280	220	106400	53200
Utilization (%)	0	98	14	16

Detail

Ici, nous avons trouvé que le temps d'exécution total des de 3193 cycles d'horloge. 217 DSP48E sont utilisés. La valeur de N qui maximise l'utilisation des ressources sans dépasser les ressources disponibles pour cette solution est de 54.

Justifiez le choix des paramètres:

même justification que la partie 1

Question 8 solution 4:

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	-	0	25
FIFO	-	-	-	-
Instance	-	196	9800	7448
Memory	-	-	-	-
Multiplexer	-	-	-	48
Register	0	-	5444	1472
Total	0	196	15244	8993
Available	280	220	106400	53200
Utilization (%)	0	89	14	16

Detail

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	7.614	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
49	49	49	49	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1	47	47	42	1	1	7	yes

Le temps d'exécution total est de 49 cycles d'horloge. 196 DSP48E sont utilisés. Cette solution utilise N = 7.

Question 9 solution 4:
L'architecture 1 utilisait 180 DSP48E pour N = 6, tandis que cette architecture utilise 196 DSP48E pour N = 7. Le ratio pour l'architecture 1 est donc de 30 DSP48E par unité de N et 28 DSP48E par unité de N pour l'architecture 2. Il n'y a donc pas de très grande différence entre les deux solutions. Toutefois, pour des grandes valeurs de N, la différence pourrait représenter un gain significatif.

<u>Architecture 3</u>
Question 10 solution 1:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	8.510	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
505101	505101	505101	505101	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1	505100	505100	10102	-	-	50	no
+ L2	10100	10100	202	-	-	50	no
++ L3	200	200	4	-	-	50	no

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	3	0	197
FIFO	-	-	-	-
Instance	-	-	-	-
Memory	-	-	-	-
Multiplexer	-	-	-	95
Register	-	-	237	-
Total	0	3	237	292
Available	280	220	106400	53200
Utilization (%)	0	1	~0	~0

Detail

a) Cette architecture requiert 505 101 cycles d'horloge et 3 DSP48E.

b) Pour la complexité temporelle, le nombre de cycles d'horloge a été réduit par un facteur plus grand que 2. Ensuite, nous utilisons 3 DSP48E alors que nous en avions 5 pour l'architecture 1 et 4 pour l'architecture 3.
c) Cette solution est réalisable sur la carte puisqu'elle requiert seulement 3 DSP48E sur les 220 disponibles.

Question 11 solution 2:
Justifiez le choix des paramètres

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	9.634	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
2506	2506	2506	2506	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1_L2	2504	2504	6	1	1	2500	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	1	-	-
Expression	-	150	0	2699
FIFO	-	-	-	-
Instance	-	-	-	-
Memory	-	-	-	-
Multiplexer	-	-	-	75
Register	0	-	5498	96
Total	0	151	5498	2870
Available	280	220	106400	53200
Utilization (%)	0	68	5	5

Detail

Justification des paramètres:
même justification que l'architecture 1

a) Le temps d'exécution est de 2506 cycles et cette solution utilise 151 DSP48E

b) Pour l'architecture 1 solution 2, nous utilisions 251 DSP48E pour N = 50 ce qui dépassait déjà les 220 DSP48E disponibles. Pour l'architecture 2, nous utilisons 151 DSP48E pour N = 50. Il nous reste donc 69 DSP48E disponibles. Alors, cette architecture pourrait prendre un N plus grand que 50 étant donné qu'il reste des ressources inutilisées.

Question 12 solution 3:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	9.634	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
5336	5336	5336	5336	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1_L2	5334	5334	7	1	1	5329	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	1	-	-
Expression	-	219	0	3993
FIFO	-	-	-	-
Instance	-	-	-	-
Memory	-	-	-	-
Multiplexer	-	-	-	75
Register	0	-	8112	96
Total	0	220	8112	4164
Available	280	220	106400	53200
Utilization (%)	0	100	7	7

Detail

Nous avons obtenu N = 73. Le temps d'exécution est de 5334 cycles d'horloge et les 220 DSP48E sont utilisés.

Justification des paramètres:
même justification que l'architecture 1
Question 13 solution 4:

Performance Estimates

Timing (ns)

Summary

Clock	Target	Estimated	Uncertainty
ap_clk	10.00	8.510	1.25

Latency (clock cycles)

Summary

Latency		Interval		
min	max	min	max	Type
13	13	13	13	none

Detail

Instance

Loop

	Latency			Initiation Interval			
Loop Name	min	max	Iteration Latency	achieved	target	Trip Count	Pipelined
- L1	11	11	5	1	1	8	yes

Utilization Estimates

Summary

Name	BRAM_18K	DSP48E	FF	LUT
DSP	-	-	-	-
Expression	-	192	0	3268
FIFO	-	-	-	-
Instance	-	-	-	-
Memory	-	-	-	-
Multiplexer	-	-	-	48
Register	0	-	7313	64
Total	0	192	7313	3380
Available	280	220	106400	53200
Utilization (%)	0	87	6	6

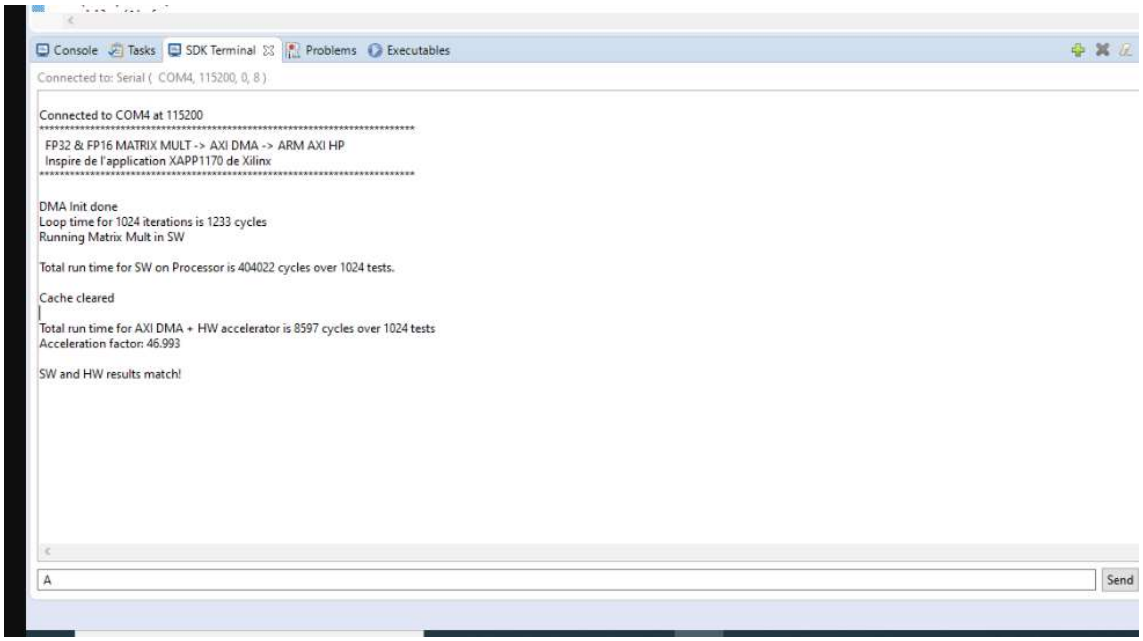
Detail

Le temps d'exécution est de 13 cycles d'horloge et cette solution utilise 192 DSP48E.

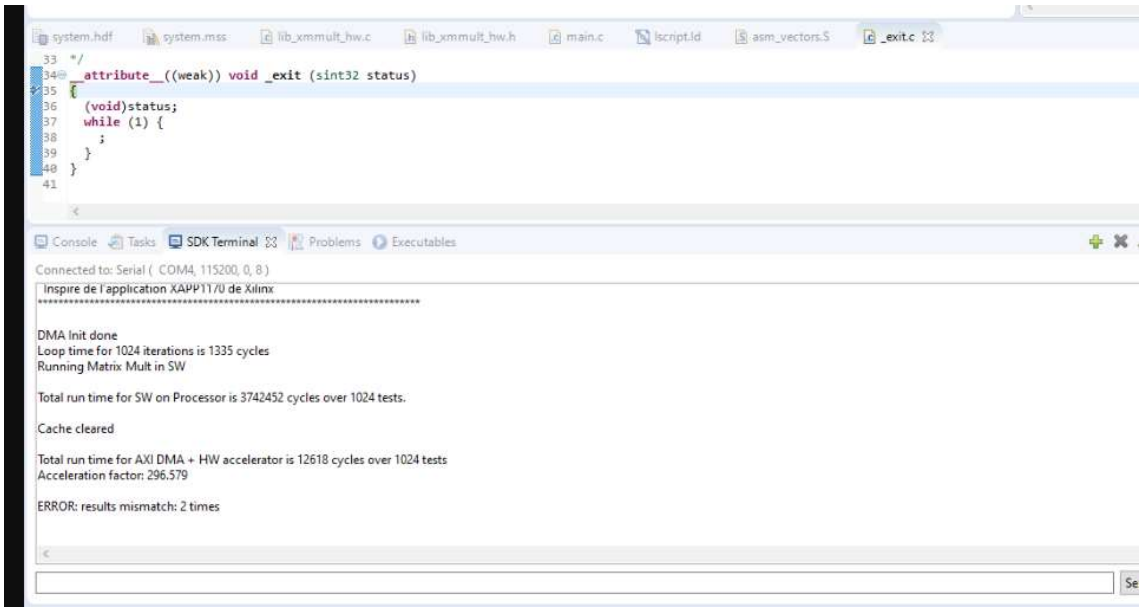
Question 14:
le n trouvé est une augmentation de 1 et par rapport aux architectures 1 et 2 ce gain peut ne pas semblé significatifs pour des dimensions petites mais en effet ca nous permet de faire 8*8*8 - 7*7*7 = 512 - 343 = 69 opérations de plus par rapport à l'architecture et 296 de plus par rapport au float donc le gain et comme même significatif et il le sera d'autant plus en augmentant la dimension

Question 15:
La fonction void wrapped_mmult_hw() nous permet l'échange de données entre le DMA et le HLS IP Core les optimisations faites sur cette fonction nous permet d'envoyer une valeur de la matrice par cycle d'horloge (II= 1) ce qui nous permet de faire un échange rapide.

<u>Virgule flottante.</u>



Half



Question 16:

Oui certainement, dans le cas où l'on a pas besoin d'utiliser une très grande précision il serait préférable d'utiliser les halves puisque cela va permettre de faire significativement plus d'opérations surtout dans des traitement d'images ou des réseaux de neurones ou les calculs sont très intensifs.