

Durée : 1h30 — Documents autorisés

1– On veut rechercher des **mots** dans un **texte** :

14pts ▷ le texte est contenu dans un tableau `char *T` de N caractères :

a	b	a	a	b	b	g	g	g	b	j	l	...	k	c	l	c	m	o	d
---	---	---	---	---	---	---	---	---	---	---	---	-----	---	---	---	---	---	---	---

▷ le texte à rechercher est contenu dans le tableau `char *motif`, sa longueur est **variable**, indiquée par la variable `int taille_motif`:

m	o	t	i	f
---	---	---	---	---

⏟
taille_motif

▷ le tableau `int positions` va contenir les positions du motif dans le texte.

La valeur -1 sera la valeur d'initialisation d'une position non déterminée.

On va fixer la taille du tableau T à $N = 2^{20} = 1048576$.

Questions :

a. Donnez une version **séquentielle en C** d'une fonction réalisant la **recherche du motif** dans le texte et (1,5pts)
qui retourne sa **position** dans le texte.

On ne recherchera que la première position du motif dans le texte.

On voudrait maintenant déterminer **en parallèle** avec CUDA, **toutes les positions** où se trouve le motif dans le texte.

b. Décrivez comment vous allez réaliser la **recherche du motif en parallèle** ? (1,5pts)

Comment les **positions** de chaque motif trouvé vont être enregistrés ?

Est-ce que des **conflits** entre les *threads* peuvent survenir ?

Vous indiquerez :

- ◇ le travail réalisé par chaque *thread* et les données qu'elle utilise ;
- ◇ la taille de la grille et des blocs.

c. Dans votre solution, est-ce que **les accès à la mémoire globale** de la carte graphique sont **optimaux** ? (2pts)

Est-ce que l'utilisation du type `char` est gênant ? Quelle solution proposez vous ?

d. Donnez une version CUDA **n'utilisant pas la mémoire partagée**. (3pts)

e. Est-ce que l'utilisation de la **mémoire partagée** peut être intéressante ? (2pts)

Comment pouvez vous utiliser la mémoire partagée pour rechercher les **différentes positions** du motif ?

Est-ce intéressant de pouvoir ajuster la taille de la mémoire partagée **dynamiquement** ?

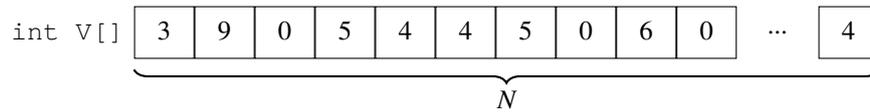
Est-ce que l'existence des « *banks* » est gênante pour votre travail ?

f. Donnez une version CUDA utilisant la **mémoire partagée**. (4pts)



2– On veut rechercher la valeur **min** et **max** d'un tableau `int V[N]` de valeurs entières.

6pts



On choisira $N = 1048576$.

Questions :

- a. Décrivez une solution parallèle en CUDA : (2pts)
- ◇ Est-ce que l'utilisation de la **mémoire partagée** est intéressante ? Expliquez comment vous pouvez exploiter.
 - ◇ Est-ce que des **conflits** peuvent survenir entre les différentes *threads* ?
 - ◇ Comment allez **vous répartir le travail** entre les différentes threads ?
Vous indiquerez la grille et les blocs que vous utiliserez.
 - ◇ Est-ce que les accès à la mémoire globale sont **optimaux** ?
- b. Donnez un programme CUDA réalisant la recherche du **min** et du **max** dans le tableau `V[]` utilisant la **mémoire partagée**. (4pts)