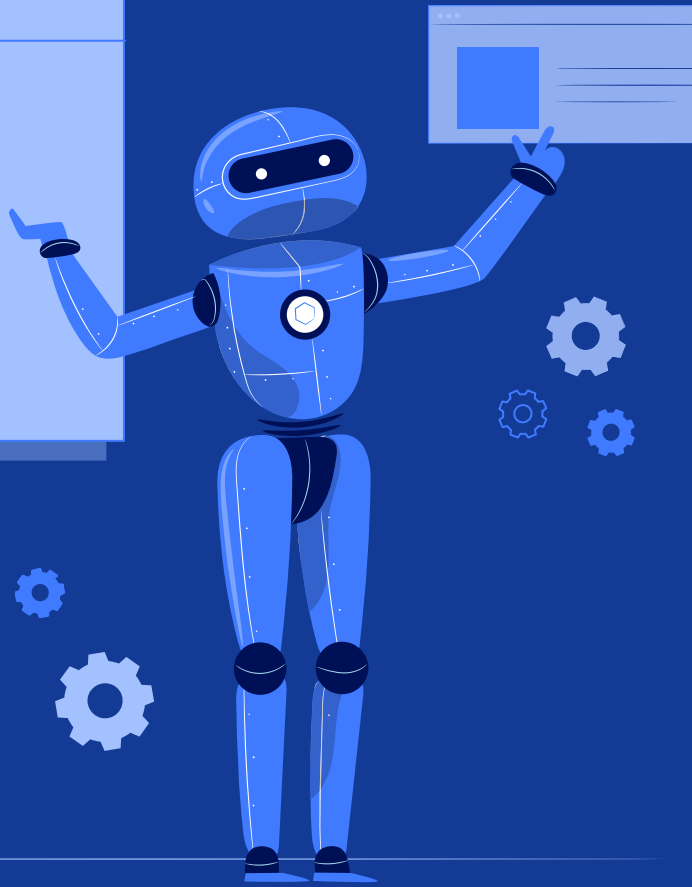
Projet HPC:

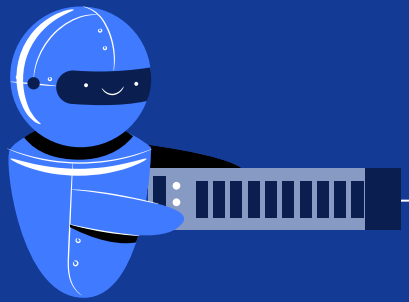
Batch merge and merge path sort

Astrid Legay - Marco Naguib - MAIN5



Sommaire

01 Rappel du sujet



Tri de tableaux

APPLICATIONS

Base de données
Traitement d'image
Théorie des graphes

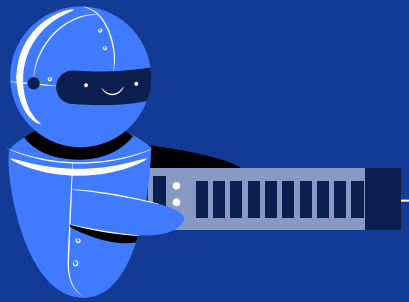
ALGORITHMES

Tri à bulle
Quicksort
Mergesort

MERGESORT

Très parallélisable
Grâce à la méthode
diviser pour régner

02 Etape 1 : mergeSmall_k



ETAPE 1

Pour $|A| + |B| \leq 1024$, écrire un kernel mergeSmall_k qui fusionne A et B avec un block et plusieurs threads.

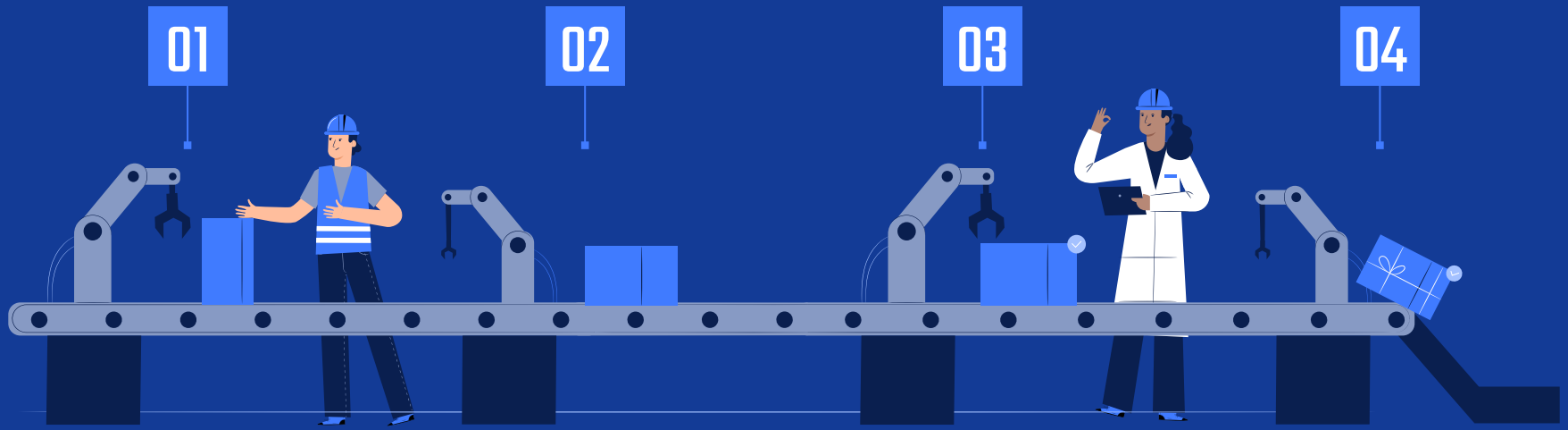
Logique

01 Ecriture séquentiel A et B

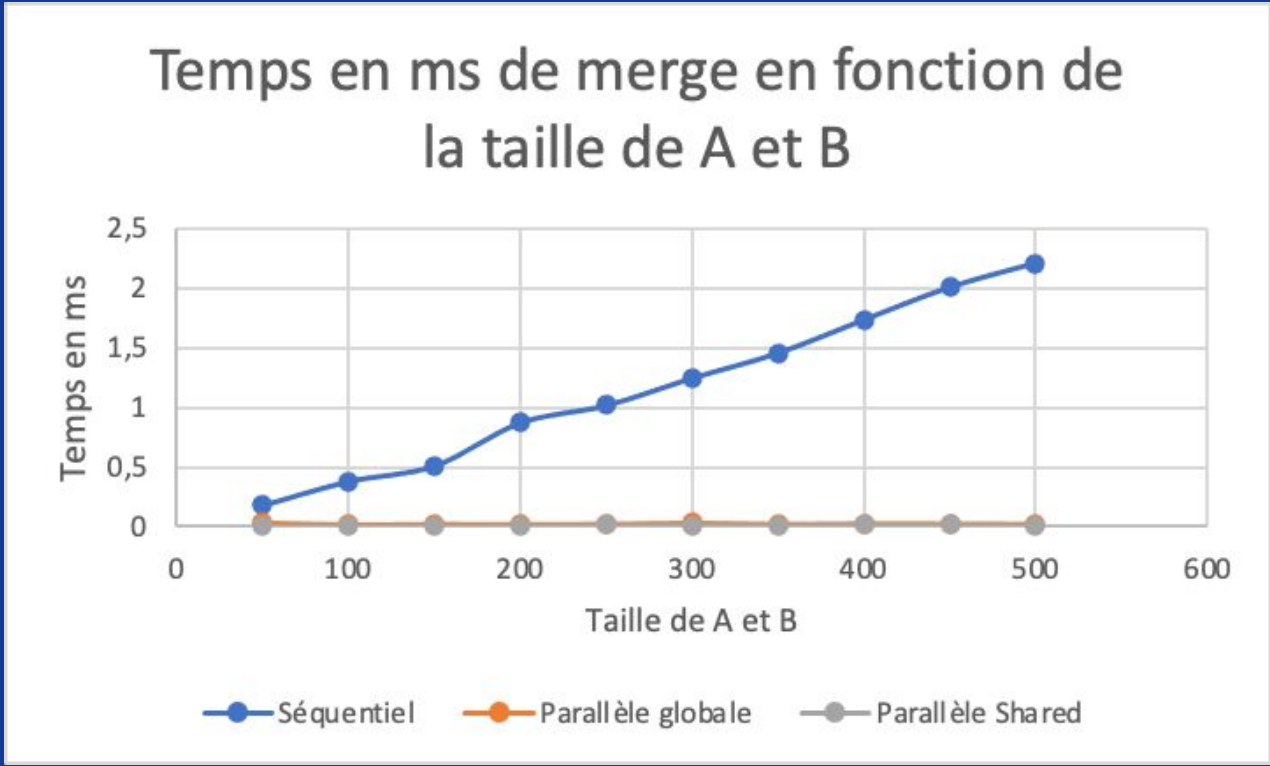
02 Ecriture des fonctions utiles

03 Ecriture en Cuda de MergeSmall_k sur mémoire globale et shared

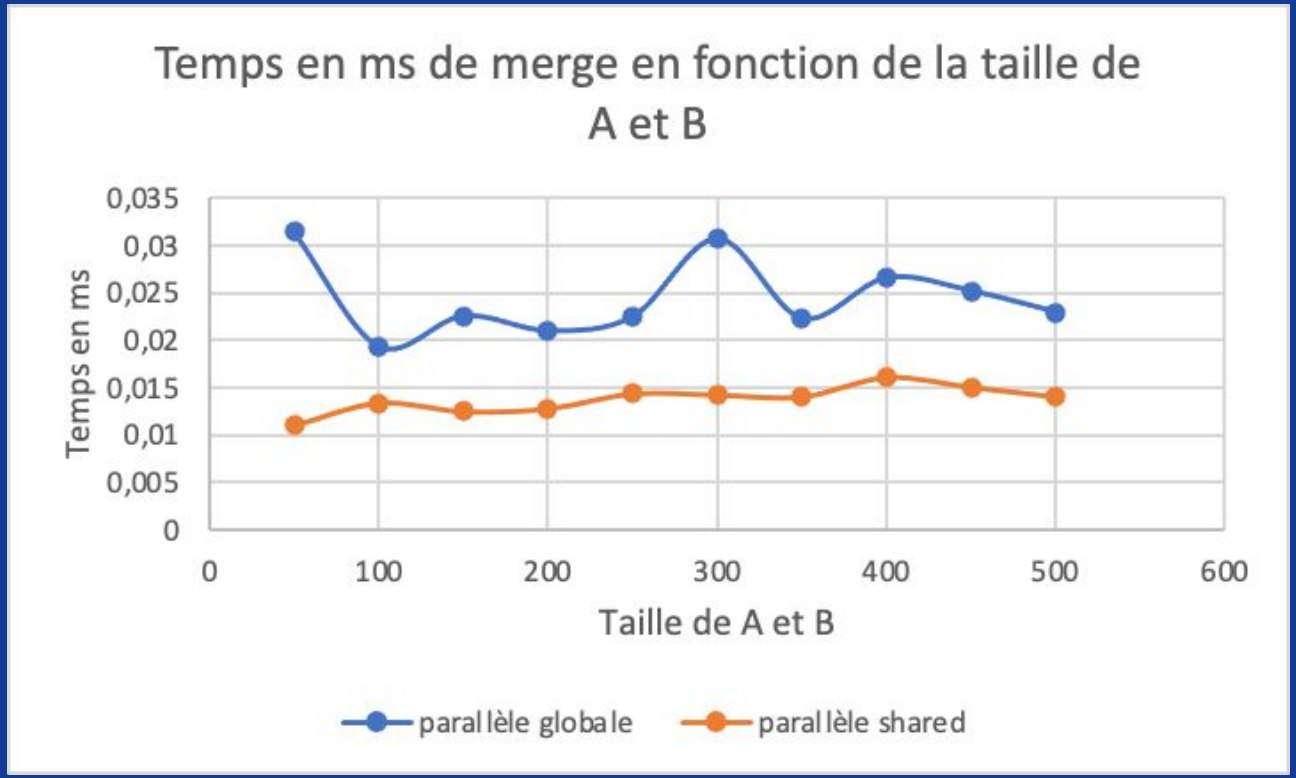
04 Tests et mesures



Résultats

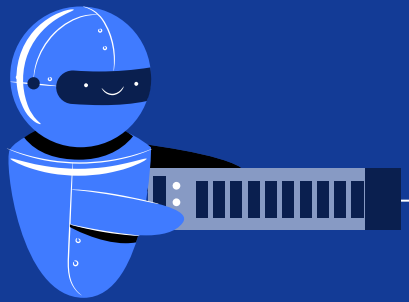

Temps en ms de merge en fonction de la taille de A et B
Temps en ms
2,5
2
1,5
1
0,5
0
0
100
200
300
400
500
600
Taille de A et B
Séquentiel
Parallèle globale
Parallèle Shared

Comparaison de mémoire globale et shared



03 Etape 2 :

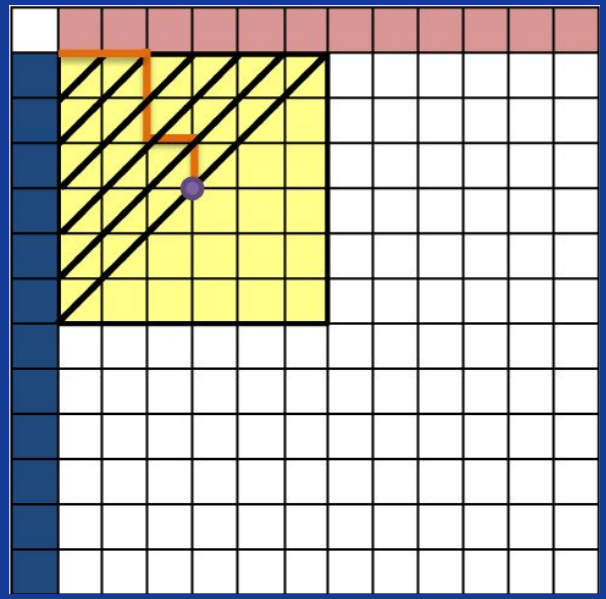
pathBig_k et mergeBig_k

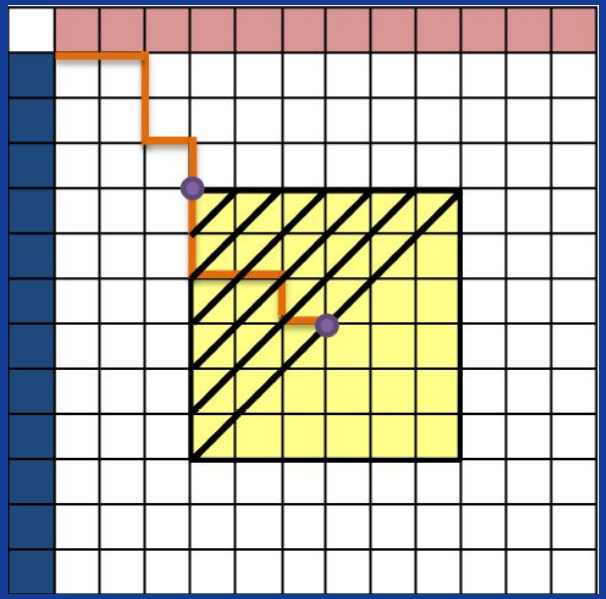


ETAPE 2

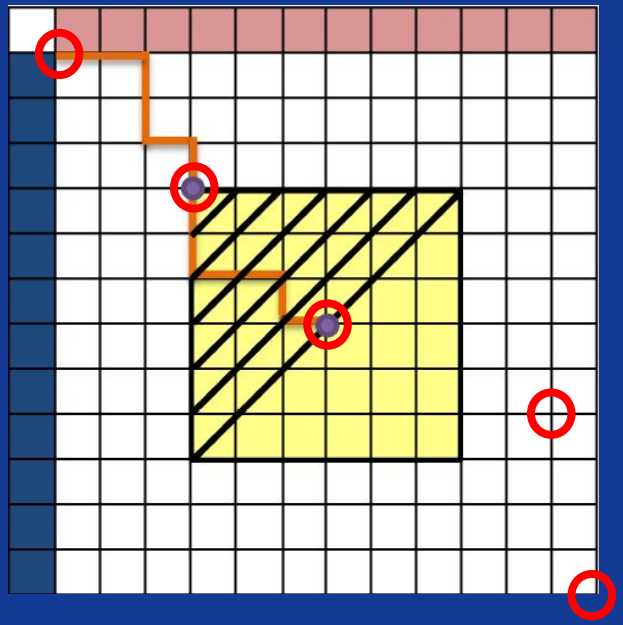
Pour toutes tailles |A|+|B| = d, plus petite que la taille de mémoire globale, écrivez deux kernel qui fusionnent et trient A et B en utilisant plusieurs blocks: le 1er kernel pathBig_k qui trouve le chemin et le 2nd mergeBig_k qui fusionne A et B. .

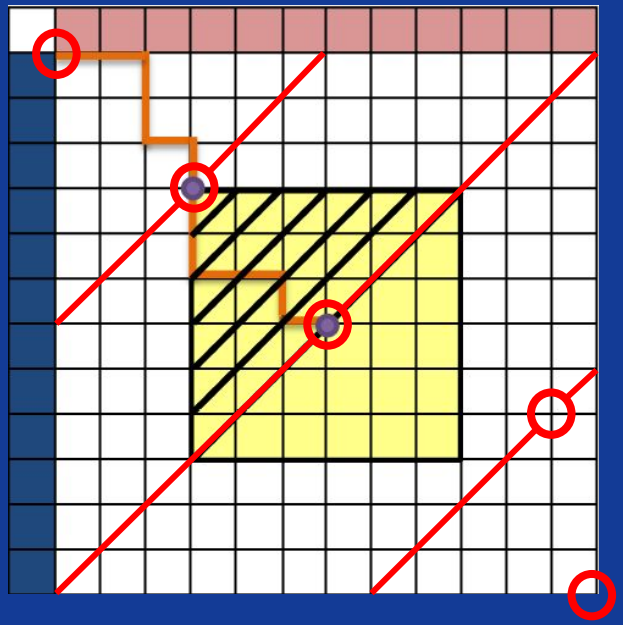
Utilisation des fenêtres glissantes





Utilisation des fenêtres glissantes





Logique du code

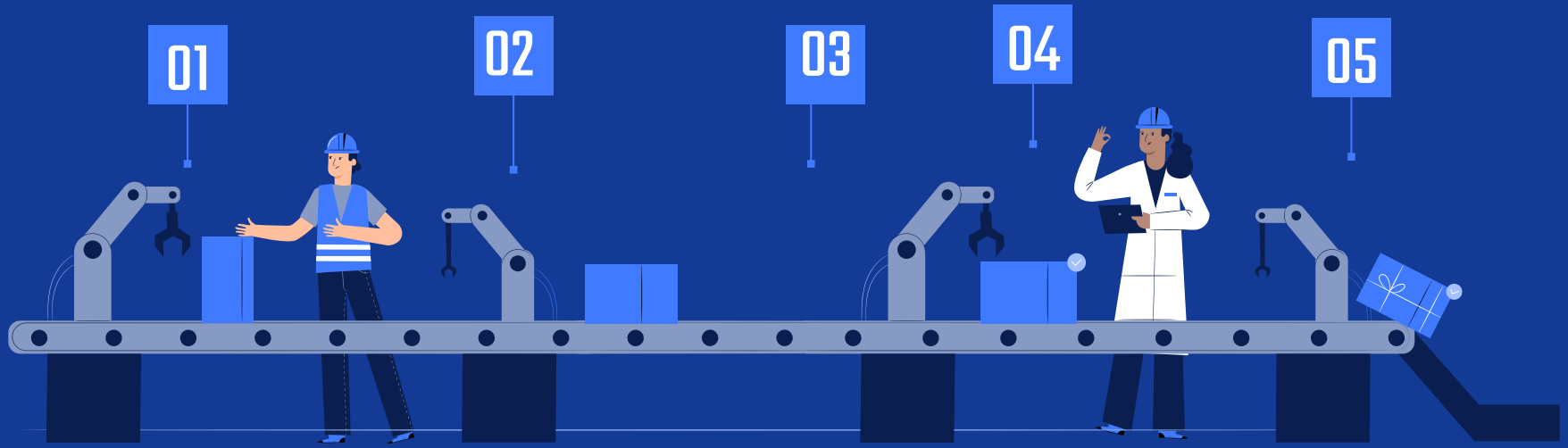
Reprises des fonctions utiles de la question 1

Code qui trouve les points d'intersection entre les diagonales et le chemin

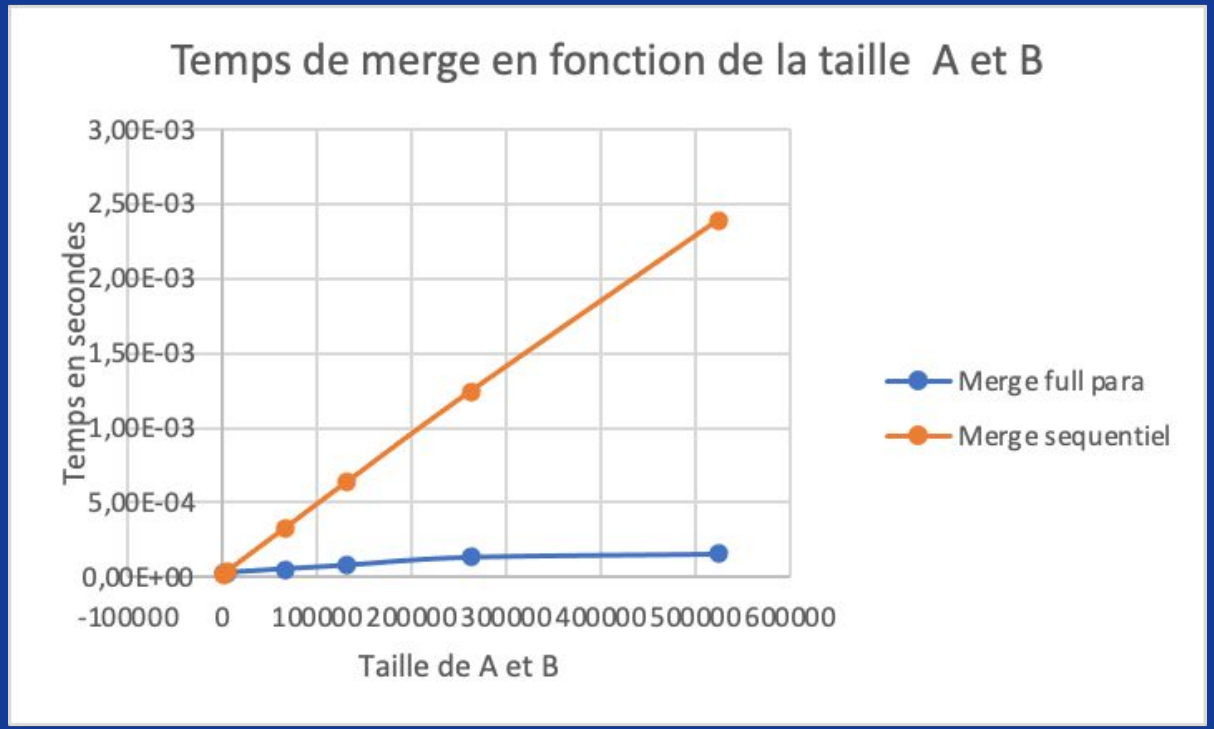
Code qui lance les fenêtres de tri 1 par 1

Code qui lance les fenêtre de tri en parallèle

Tests et mesures

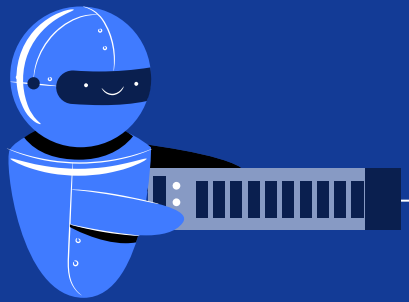


Résultats des mesures : Durée globale



04 Etape 3 :

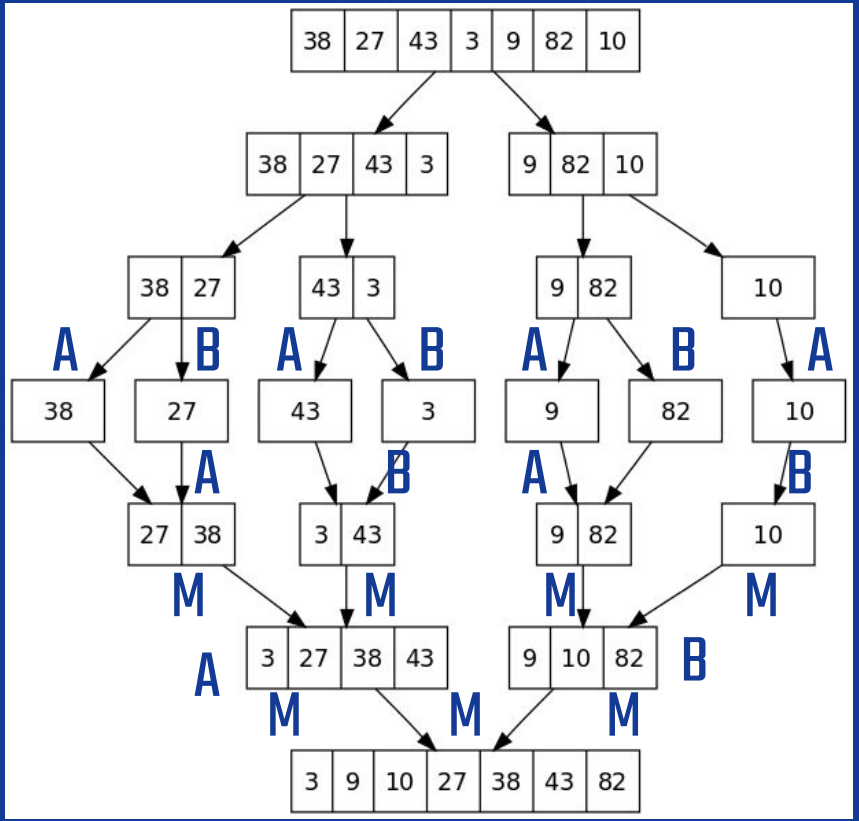
merge_sort



ETAPE 3

En bouclant sur les appels appropriés de pathBig_k et de mergeBig_k, écrivez une fonction qui trie tout tableau M de taille d suffisamment plus petite que la mémoire globale. Donnez le temps d'exécution par rapport à d.

Merge Sort


38 27 43 3 9 82 10
38 27 43 3
9 82 10
38 27
43 3
9 82
10
A
B
A
B
A
B
A
38
27
43
3
9
82
10
A
B
A
B
27 38
3 43
9 82
10
M
M
M
M
A
3 27 38 43
9 10 82
B
M
M
M
3 9 10 27 38 43 82

Logique du code

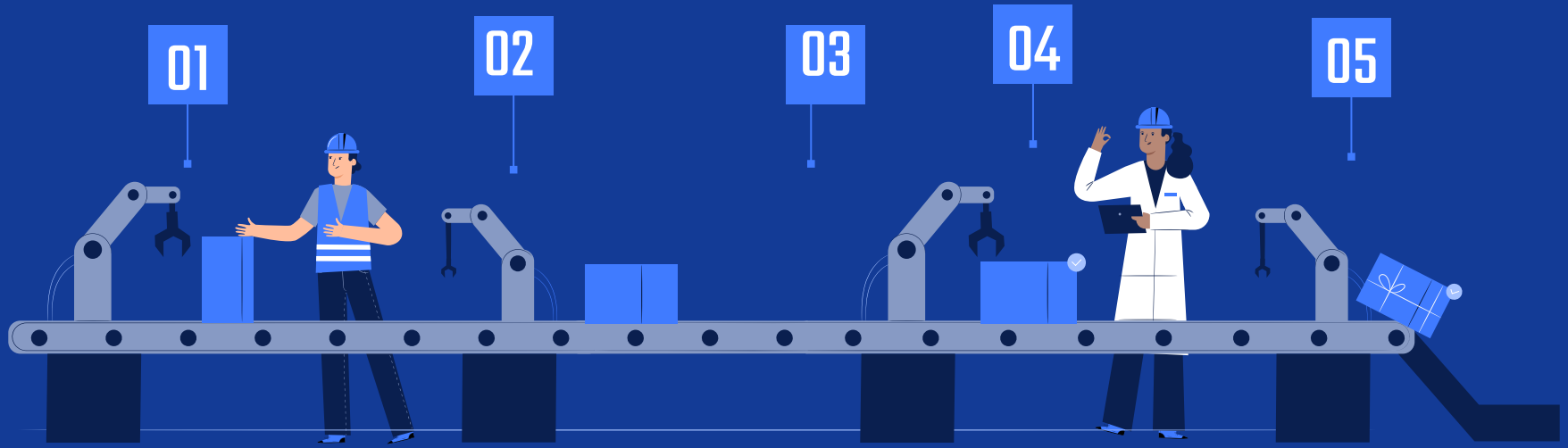
01 Code en séquentiel en C

02 Reprises des fonctions utiles de la question 1 et 2

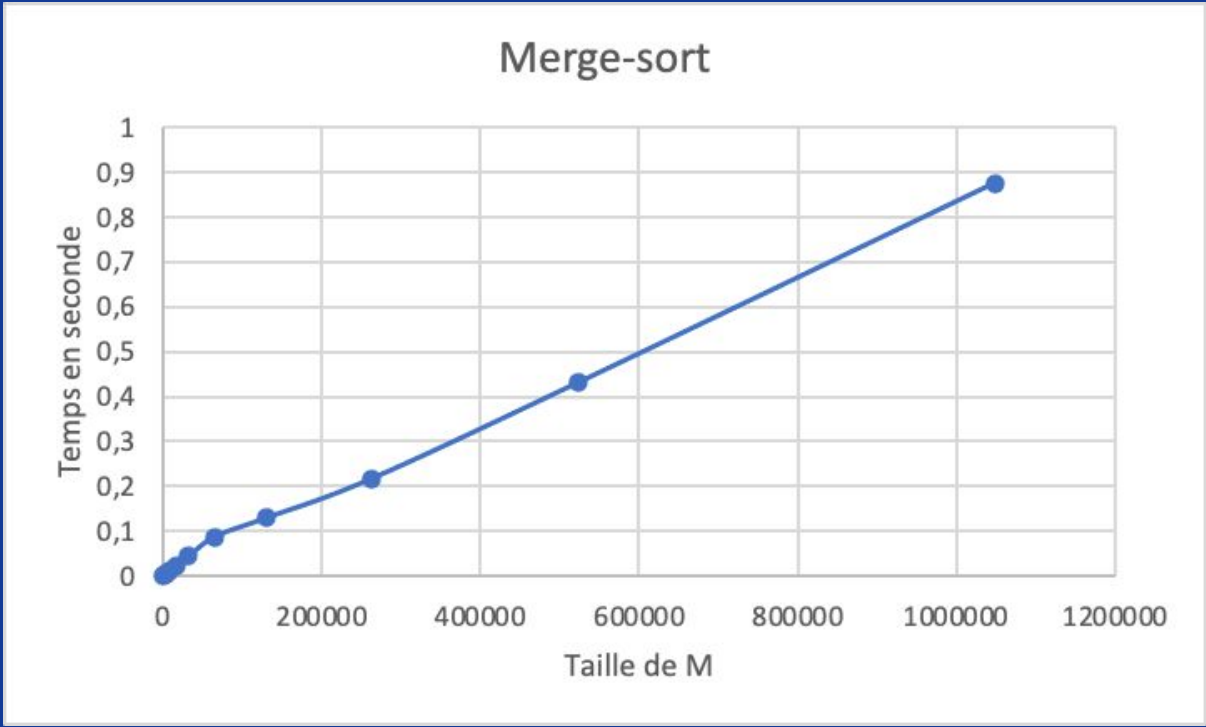
03 “ Conversion ” du code séquentiel en CUDA

04 Tests et mesures

05 Amélioration du code

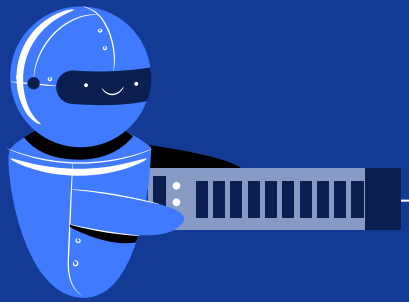


Résultats des mesures



05 Etape 5 :

mergeSmallBatch_k



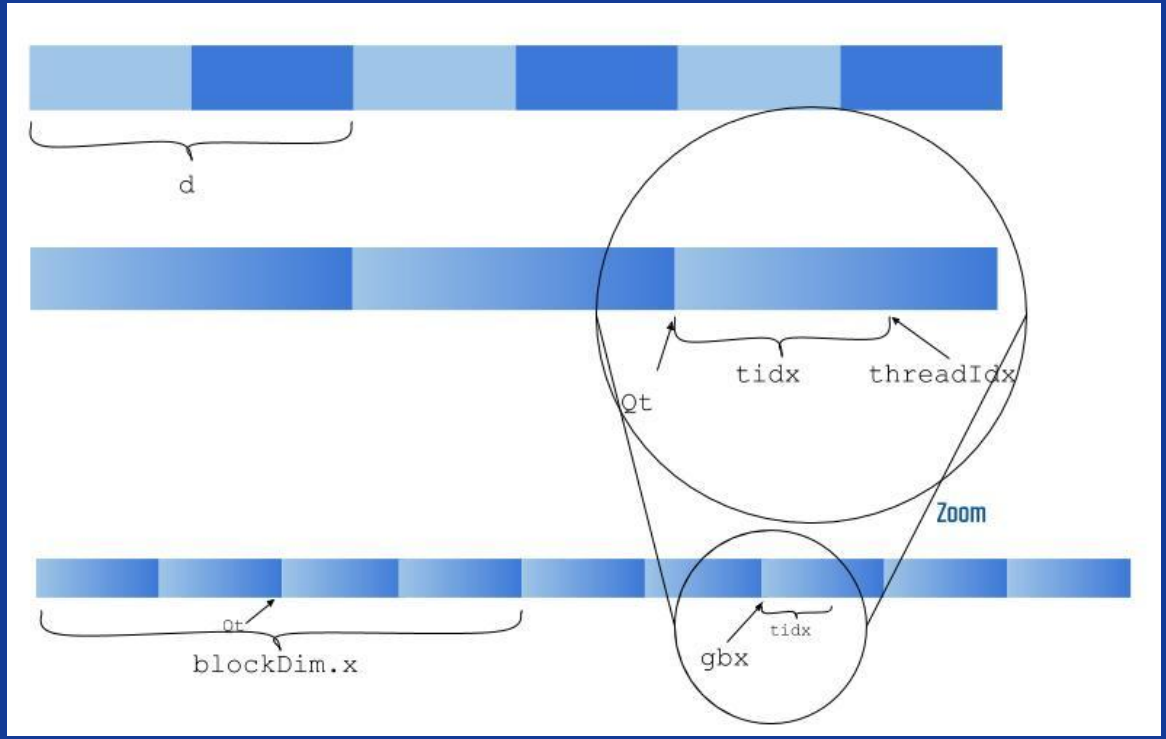
ETAPE 4

Expliquer pourquoi les indices

- `int tidx = threadIdx.x%d;`
- `int Qt = (threadIdx.x-tidx)/d;`
- `int gbx = Qt + blockIdx.x*(blockDim.x/d);`

Sont importants dans la définition de `mergeSmallBatch k`.

Question 4



ETAPE 5

Écrivez le noyau mergeSmallBatch k qui fusionne deux par deux $\{Ai\}1 \leq i \leq N$ et $\{Bi\}1 \leq i \leq N$

Donnez le temps d'exécution par rapport à d = 4, 8,..., 1024.

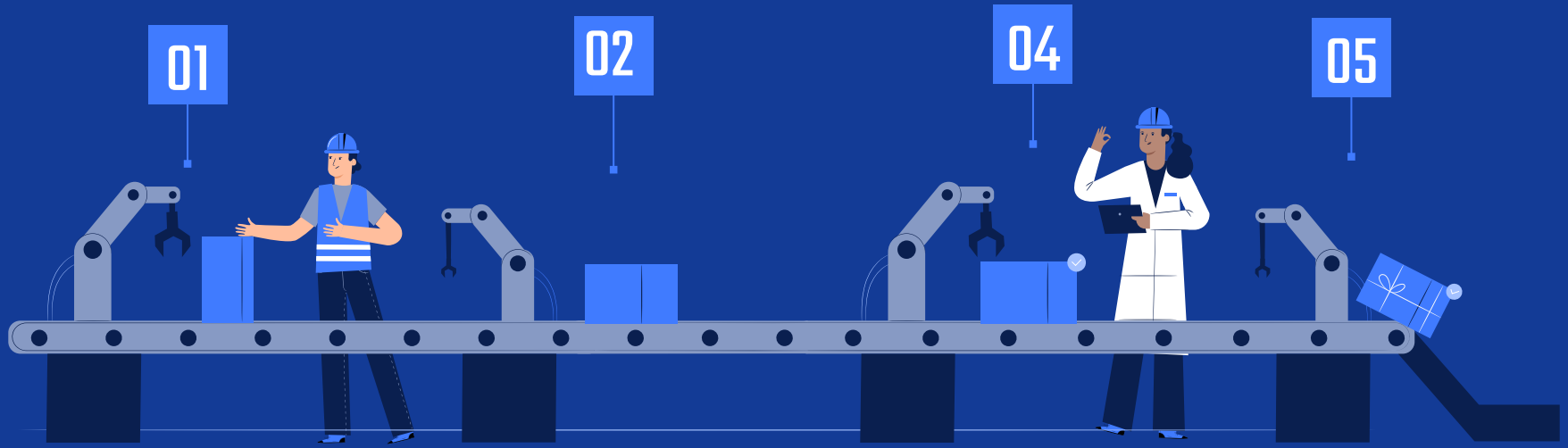
Logique

Reprise de la fonction MergeSmall

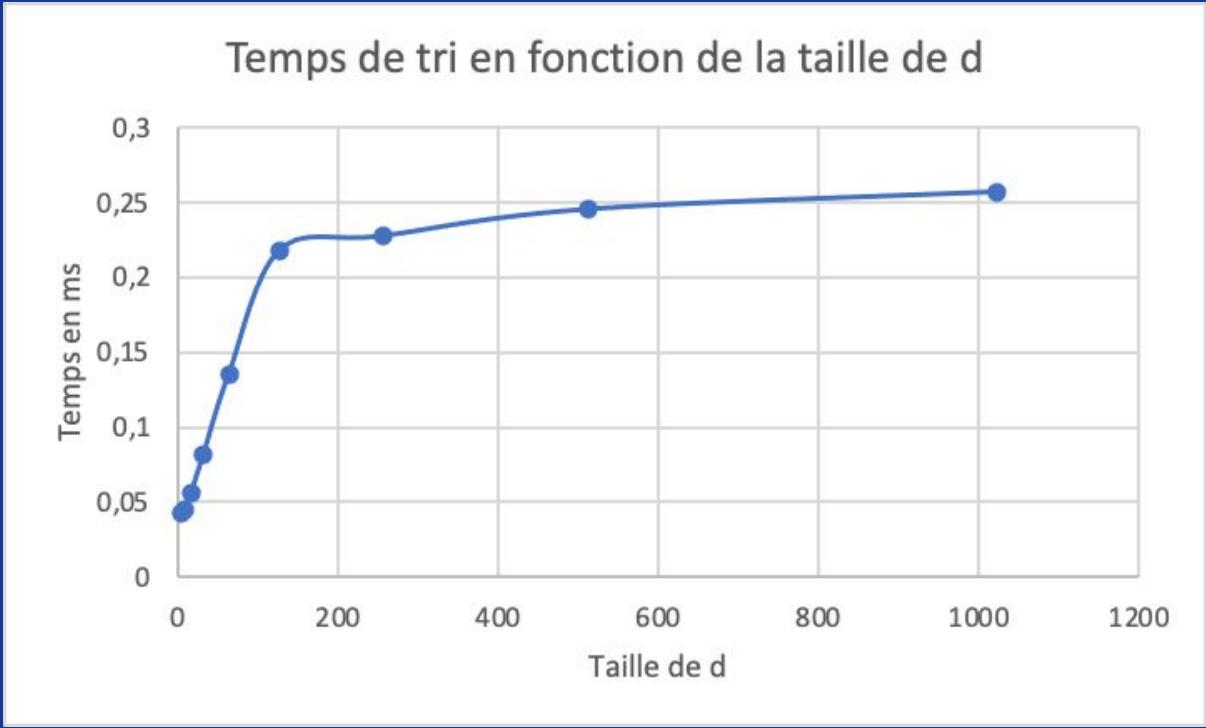
Utilisation des indices de la question 4

Tests et mesures

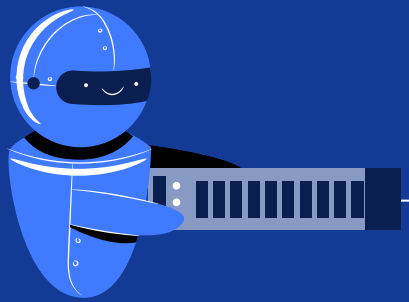
Optimisation du code



Résultats des mesures

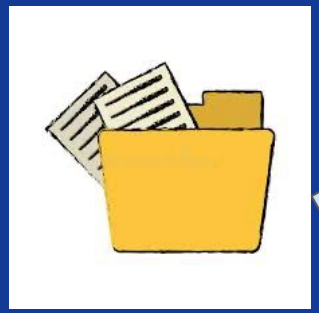


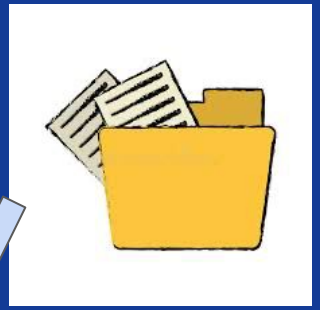
05 Applications



D'un point de vue "pratique"

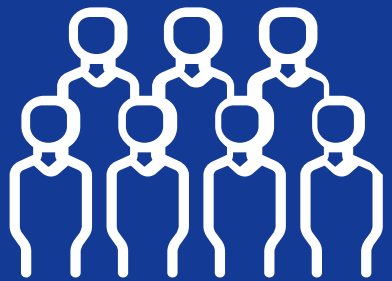
Fusion d'entreprise





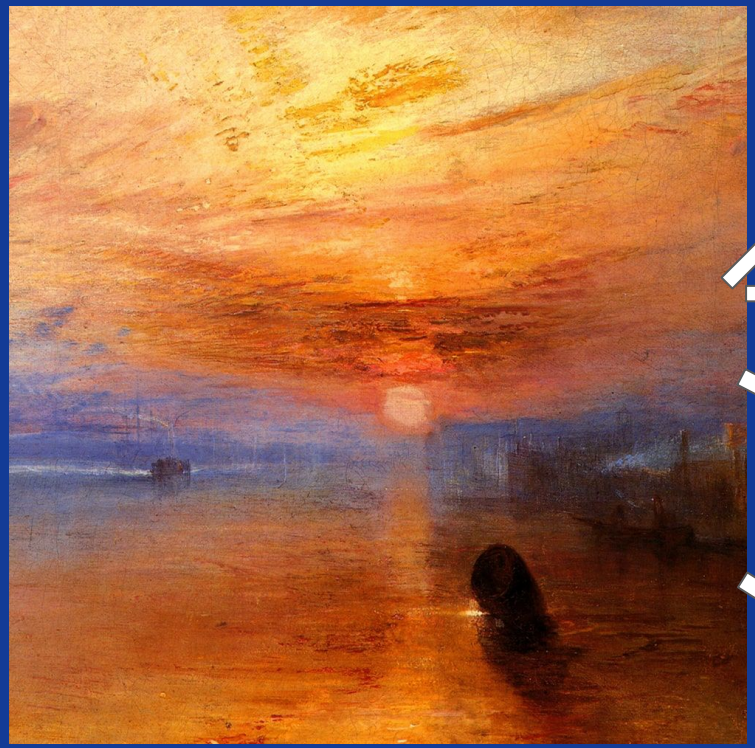
Entreprise A

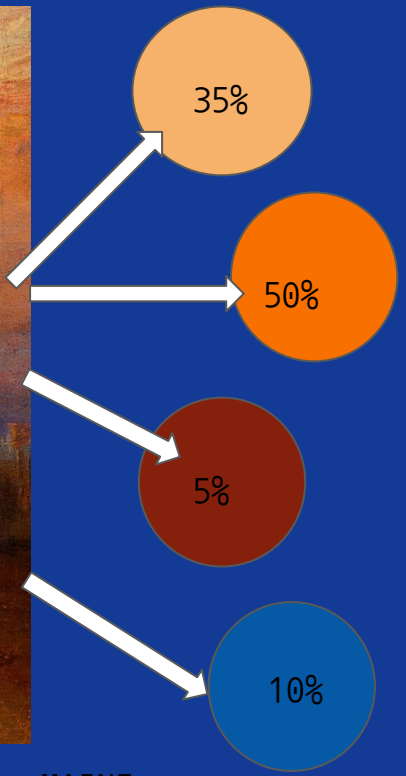
Entreprise B



Entreprise M

Traitements images : dans l'art

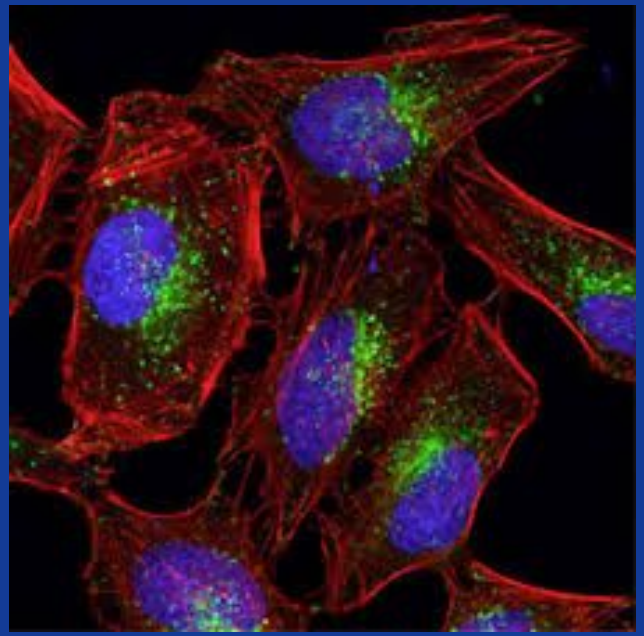




- Reconnaitre la palette de couleur
- Reconnaître le peintre
- Classer par couleurs
- Reconnaître le mouvement artistique
- Reconnaître la période de l'artiste

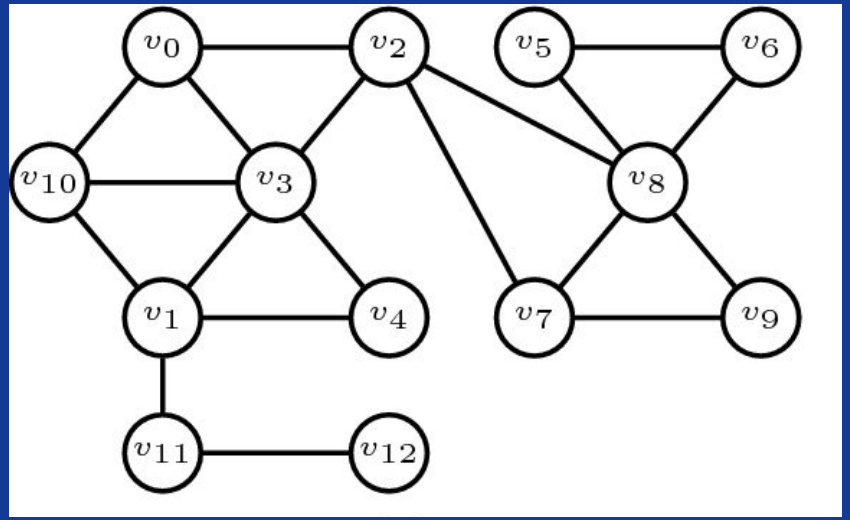
Traitements images : dans la médecine

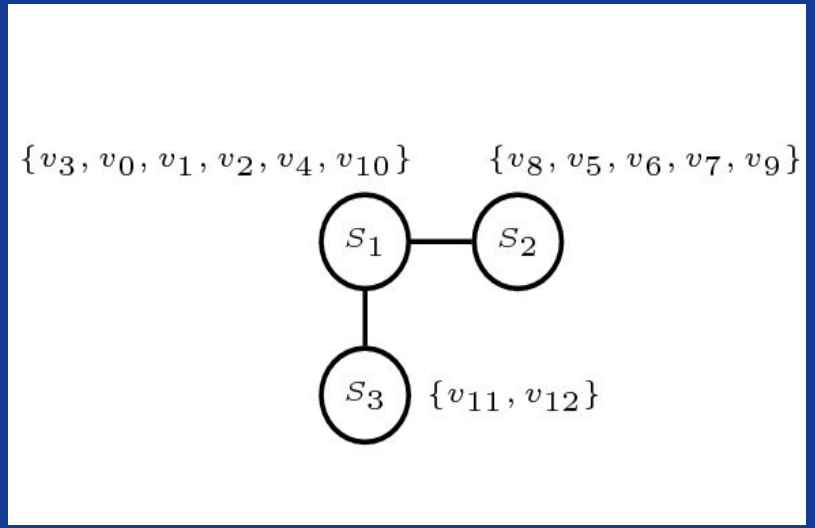
- Comptage de cellules cancéreuses
- Repérage de cellules cancéreuses



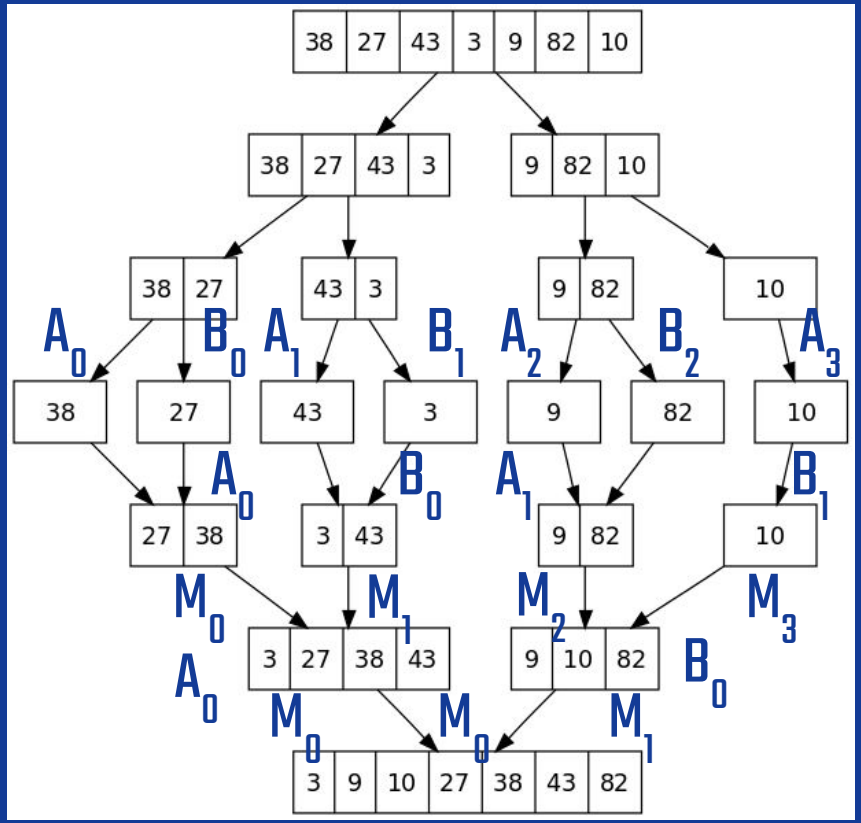
D'un point de vue "technique"

Compression de graphes





Merge Sort avec MergeSmallBatch



COMPARAISONS

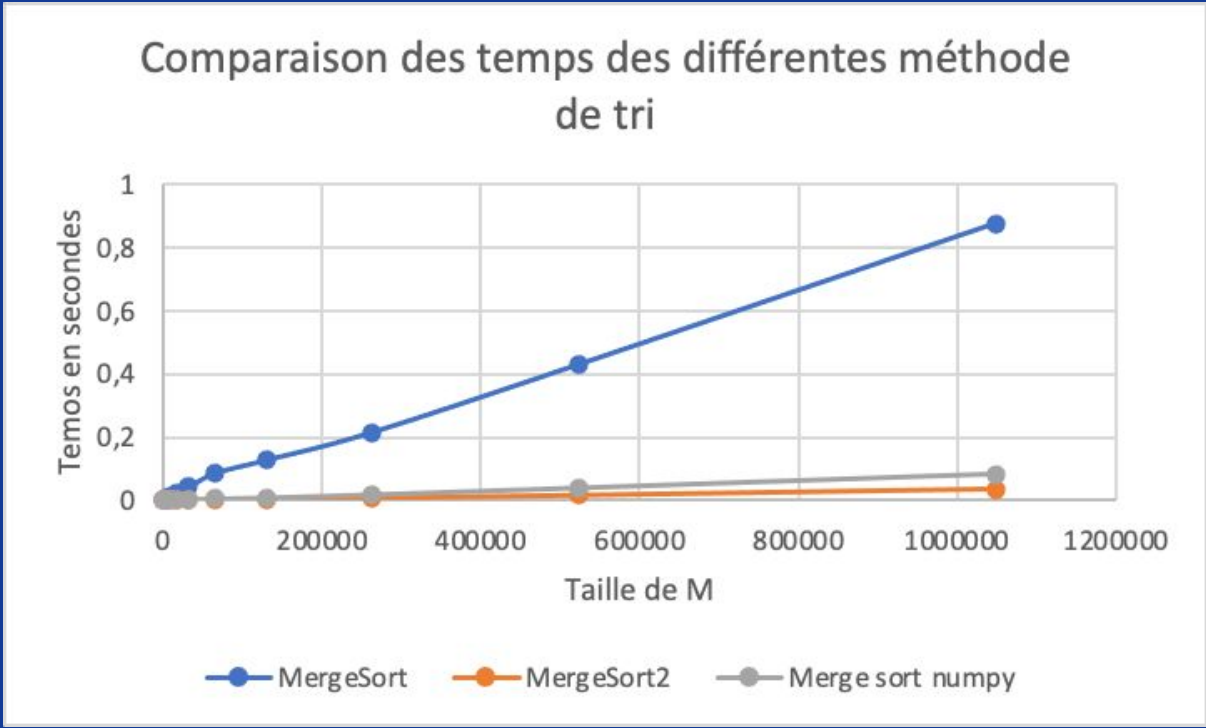
01.

Comparaison entre les temps de la question 3 et 5

02.

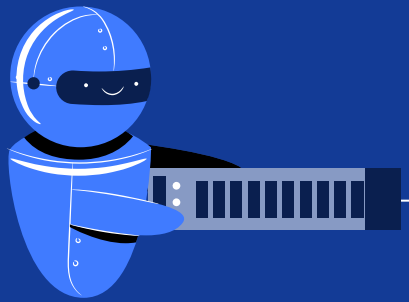
Comparaison avec les temps de mergesort numpy

Résultats des mesures



Fin

Conclusion



Conclusion

GPU

Découvrir le monde du GPU en pratique

Cuda - Parallélisme

Découvrir un autre moyen de mettre en place du parallélisme

Difficultées

- Milieu
- Boucle infinie
- Gestion des restes

Réfléchir à différentes applications possibles

Merci

Avez-vous des questions ?