TP 1 :

Points abordés :

- Outils de mesure de temps calcul
- Localités spatiale et temporelle
- Utilisation des registres
- Pipeline d'évaluation des expressions

Contexte

Le code proposé dans ce TP calcule une solution approchée en 3 dimensions d'espace (x, y, z) de l'équation de la chaleur dans un cube $[0, 1]^3$:

$$\frac{\partial u}{\partial t} = \Delta u$$

avec conditions aux limites de Dirichlet au bord.

Le calcul se fait à l'aide d'une méthode explicite en temps (Euler) et discrétisation par différences finies en espace.

Dans la dernière partie du TP, le code permet d'ajouter un terme d'advection (linéaire) $\frac{\partial u}{\partial x}$ au second membre.

Préparation

A chaque étape, regarder les messages affichés pour voir si tout s'est bien passé !

1. Récupérer l'archive `TP1.tar.gz` et extraire les fichiers.
2. Ouvrir un terminal et se placer dans le répertoire `I03_TP1` qui vient d'être créé
3. préparer la compilation du code du TP avec les commandes :

   mkdir -p build
   cd build
   cmake ../src
   cd ..

4. Se remettre dans le répertoire `I03_TP1` et compiler:

   make -C build

5. Exécuter le code avec la commande:
   ./build/PoissonSeq

Si on modifie un ou plusieurs fichiers sources (dans le sous-répertoire src), il faut recompiler (point 4).

Si on ajoute un nouveau fichier ou on enlève un fichier existant (dans le sous-répertoire src), il faut adapter les fichiers CMakeLists.txt et refaire les points 3 et 4.

Les commandes ci-dessus génerent une version optimisée par la compilateur ("Release"). Si nécessaire, on peut compiler une version "Debug" (non optimisée) qui permet d'utiliser un outil de débug (exécution pas à pas, afficher des valeurs en cours de calcul, etc).

Certains des outils donnent aussi plus de renseignements sur une version "Debug" que sur une version "Release".

Remplacer les commandes de 3. et 4. par:

```
mkdir -p build_debug
cd build_debug
cmake -DCMAKE_BUILD_TYPE=Debug ../src
cd ..
make -C build_debug
```

Mesure du temps de calcul global

Afficher le temps de calcul global avec `time` :

time ./build/PoissonSeq

A l'écran (exemple) :

```
real    0m30,283s
user    0m30,186s
sys     0m0,096s
```

- Avantage : n'est pas intrusif

  pas besoin de modifier le code, ni de le compiler avec des options spécifiques.

- Désavantage : donne une information globale

  on ne sait pas dans quelle partie du code, on passe peu/beaucoup de temps, ni pourquoi.

Mesures plus précises : outil de “profiling” `gprof`

Fait partie de la famille gcc/g++/gfortran.

- Avantage : calcule le nombre d’appels de chaque fonction et le (pourcentage du) temps qui y est passé
- Avantage : n’est pas très intrusif (pas besoin de modifier le code, mais il faut le recompiler avec une option spécifique: -pg).
- Désavantage : le temps passé dans une fonction est peu précis dans une fonction “courte”
- Désavantage : ne mesure pas le temps dans les différentes parties d’une fonction.
- Désavantage : ne rentre pas dans les librairies dynamiques.

Mode de fonctionnement:

Ajoute dans chaque fonction, un comptage du nombre d'appels de cette fonction et évalue statistiquement le temps passé dans cette fonction (tous les 0.01 secondes on enregistre dans quelle fonction on se trouve).

Utilisation de `gprof`

Recompiler en utilisant l'option -pg:

```
mkdir -p build_gprof
cd build_gprof
cmake -DCMAKE_CXX_FLAGS="-pg" ../src
cd ..
make -C build_gprof
```

Exécuter le code:

```
./build_gprof/PoissonSeq
```

Collecter les mesures

```
gprof ./build_gprof/PoissonSeq
```

Mesures plus précises : outil de “profiling” `perf`

Outil spécifique linux

- Avantage : très puissant (mesure le temps passé dans une fonction, une instruction C/C++/fortran, une instruction binaire, multiples indicateurs de performance)
- Avantage : non intrusif
- Désavantage : plus compliqué à utiliser
- Désavantage : ne rentre pas toujours dans les librairies dynamiques.
- Désavantage : nécessite que la machine soit configurée correctement.

Outil à privilégier quand c’est possible

Mode de fonctionnement:

Enregistre les événements dans le noyau Linux, évalue statistiquement le temps passé dans les fonctions et les instructions.

Utilisation simple de `perf`

Il n'est pas toujours nécessaire de recompiler. Sur certaines machines, une version "Debug" est préférable.

Instrumenter le code (générer les mesures):

```
perf record ./build/PoissonSeq
(ou perf record ./build_debug/PoissonSeq)
perf report
```

Voir la documentation de perf pour les (nombreuses) autres options.

Mesures plus précises : outil de “profiling” `valgrind-callgrind`

- Avantage : intermédiaire (mesure le temps passé dans une fonction, une instruction C/C++/fortran)
- Avantage : non intrusif
- Avantage : gère mieux les librairies dynamiques.
- Avantage : l'outil `kcachegrind` offre une interface graphique pratique à utiliser.
- Désavantage : temps d'exécution multiplié par ≈ 40

Mode de fonctionnement:

> *Le code s'exécute dans une machine virtuelle (simule une machine "idéale"). Permet de mesurer précisement un grand nombre d'indicateurs. Mais explique le facteur de ralentissement.*

Utilisation: Utiliser une version "Debug".

Instrumenter le code

```
valgrind --tool=callgrind ./build_debug/PoissonSeq
```

Crée un fichier `callgrind.out.XXXX` où `XXXX` est le numéro de processus qu'on vient d'exécuter.

Afficher et explorer les résultats de mesure:

```
kcachegrind callgrind.out.XXXX
```

Mesures “manuelles” des temps de calcul.

Principe:

- encadrer le bloc d'instructions que l'on veut mesurer par des appels à des fonctions qui renvoient la valeur de l'horloge interne de la machine,
- calculer le temps d'exécution dans ce bloc en faisant la différence des valeurs ci-dessus.

Fonctions système disponibles:

- `clock()` (fonction système C)
- `gettimeofday(...)` (fonction système C)
- `system_clock(...)` (fonction système fortran 90)
- `high_resolution_clock` (classe C++ 11)
- ...

Librairies externes (timers haute résolution) :

- PAPI : `http://icl.cs.utk.edu/papi`
- BoostTimers : `http://www.boost.org/doc/libs/1_65_1/libs/timer/doc/cpu_timers.html`

Remarques

Chaque fonction système est utilisable ou non suivant le langage de programmation utilisé.
La précision n'est pas la même et dépend de la machine : consulter la documentation, faire des tests.

Exemple avec la fonction C système `clock`:

```
#include<stdio.h>
#include<math.h>
#include<time.h>
int main()
{
   clock_t t1, t2;
   t1 = clock();

   int n=100000;
   double s = 0.0;
   for(int i = 0; i < n; i++) s += sin((0.2*i)/n);

   t2 = clock();
   float diff = (float)(t2 - t1)/CLOCKS_PER_SEC;
   printf("temps calcul : %f s",diff);
   return 0;
}
```

Dans le cadre de ce TP, on fournit une classe C++ Timer (fichiers `timer.cxx` et `timer.hxx`) avec les fonctionnalités suivantes:

```
 //definit une variable "chronometre"
Timer T;
//demarrer le chronometre T
T.start();
//arreter le chronometre T
T.stop();
//remettre a zero
T.reset();

// retourne le temps mesure
// entre un appel start() et un stop()
// (utiliser juste apres le stop())
double dt = T.elapsed();
```

Ajout de mesures manuelles de temps calcul :

Se mettre dans le répertoire `I03_TP1`, faire une copie du répertoire `src` dans Q1/src et compiler:

```
mkdir -p Q1/build
cp -rf src Q1/src
cd Q1/build && cmake ../src && cd ..
make -C build
```

Q1. Modifier le programme principal `src/main.cxx` en utilisant la classe Timer pour mesurer séparément

- **le temps d'initialisation (lignes 31-32 du fichier `src/main.cxx`)**
- **la somme du temps d'exécution des parties calcul des itérations (lignes 43-45 du fichier `src/main.cxx`)**

Exécuter le code avec et sans l'option `out=10` (la seconde exécution sauvegarde les résultats sur fichier):

```
./build/PoissonSeq
./build/PoissonSeq out=10
```

L'exécution avec paramètre `out=10` sert à sauvegarder les résultats pour pouvoir les comparer avec les autres versions.

Localité spatiale

Se mettre dans le répertoire `I03_TP1`, faire une copie du répertoire `Q1/src` dans `Q2/src` et compiler:

```
mkdir -p Q2/build
cp -rf Q1/src Q2/src
cd Q2/build && cmake ../src && cd ..
make -C build
```

Q2. La classe Values repésente un ensemble de $n_0 \times n_1 \times n_2$ valeurs et possède des fonctions (`operator()`) qui retournent la composante $v_{i,j,k}$ où v est de type `Values`.

- Décrire comment sont rangées en mémoire les composantes $v_{i,j,k}$ (examiner les fichiers `src/values.cxx` et `src/values.hxx`).
- Changer la façon de ranger les composantes en mémoire (il y a plusieurs possibilités)
- Comparer les résultats et les temps de calcul avec la version de référence dans `Q1`.
- Expliquer les différences éventuelles de temps calcul (examiner aussi le fichier `scheme.cxx`).

Autre équation

On a modifié le code pour calculer la solution approchée de

$$\frac{\partial u}{\partial t} = \Delta u + \frac{\partial u}{\partial x}$$

Version de référence:

Se mettre dans le répertoire `I03_TP1`, faire une copie du répertoire `Q2/src` dans `Q3a/src`, remplacer le fichier `Q3a/src/scheme.cxx` par le fichier `scheme_Q3a.cxx` et compiler:

```
mkdir -p Q3a/build
cp -r Q2/src Q3a
cp scheme_Q3a.cxx Q3a/src/scheme.cxx
cd Q3a/build && cmake ../src && cd ..
make -C build
```

Q3a. Implémentation de référence : la convection et la diffusion sont traitées dans des boucles séparées

- Examiner la nouvelle version du fichier `scheme.cxx`
- Exécuter le code

Localité temporelle

Première variante : calcul des termes des équations dans une seule boucle

Se mettre dans le répertoire `I03_TP1`, faire une copie du répertoire `Q3a/src` dans Q3b/src et compiler:

```
mkdir Q3b
cp -rf Q3a/src Q3b/src
mkdir -p Q3b/build
cd Q3b/build && cmake ../src && cd ..
make -C build
```

Q3b. Fusion des boucles de calcul

- Examiner le fichier `Q3b/src/scheme.cxx`: il contient 4 boucles du type

```
for (i=imin; i<imax; i++)
  for (j=jmin; j<jmax; j++)
    for (k=kmin; k<kmax; k++) {
      ...
    }
```

  Rassembler ces 4 boucles en une seule.

- Faire les tests analogues au cas Q3a, comparer et commenter les résultats.

Variables scalaires locales

Seconde variante : utilisation d'une variable scalaire auxiliaire.

Se mettre dans le répertoire `I03_TP1`, faire une copie du répertoire `Q3b/src` dans Q3c/src et compiler:

```
mkdir Q3c
cp -rf Q3b/src Q3c/src
mkdir -p Q3c/build
cd Q3c/build && cmake ../src && cd ..
make -C build
```

Q3c. Remplacer la séquence d'instructions:

```
u2(j,i,k) = u1(i,j,k);
u2(i,j,k) -= f(u1)
u2(i,j,k) -= g(u1)
```

par

```
u_aux = u1(i,j,k);
u_aux -= f(u1);
u_aux -= g(u1);
u2(i,j,k) = u_aux;
```

- Faire les tests analogues aux cas Q3a et Q3b, comparer et commenter les résultats.

Expressions les “plus longues possibles”

Troisième variante : fusionner les expressions.

Se mettre dans le répertoire `I03_TP1`, faire une copie du répertoire `Q3c/src` dans Q3d/src et compiler:

```
mkdir Q3d
cp -rf Q3c/src Q3d/src
mkdir -p Q3d/build
cd Q3d/build && cmake ../src && cd ..
make -C build
```

Q3d. Fusionner les expressions dans le fichier `scheme.cxx`:

```
d = u1(i,j,k);
d -= f(u1);
d -= g(u1);
u2(i,j,k) = d
```

en

```
u2(i,j,k) = u1(i,j,k) - f(u1) - g(u1);
```

- Faire les tests analogues aux cas précédents, comparer et commenter les résultats.