

Atomické operace a redukce

PCG – Paralelní výpočty na GPU

Cvičení 4, 2022/2023

Jirka Jaroš

Vysoké učení technické v Brně, Fakulta informačních technologií
Božetěchova 1/2, 612 66 Brno - Královo Pole
jarosjir@fit.vutbr.cz



- **Export cest ke CUDA**

```
export PATH=$PATH:/usr/local/share/cuda-11.4/bin
```

```
export LD_LIBRARY_PATH=$LD_LIBRARY_PATH:/usr/local/share/cuda-  
11.4/bin:/usr/local/share/cuda-11.4/targets/x86_64-  
linux/lib:/usr/local/share/cuda-11.4/extras/CUPTI/lib64
```

- **Pro export cest použijte script**

```
cd Utilities  
chmod +x path_export.sh  
eval `./path_export.sh`
```

PŘIPOJENÍ NA KAROLÍNU

- **Připojte se na Karolínu**

```
ssh karolina
```

- **Vaše PC: Připojte si disk z Karolíny a nakopírujte tam obsah 1. cvičení**

```
mkdir /tmp/karolina  
sshfs karolina: /tmp/karolina
```

- **Nastartujte job (pozor, máme jen 8 uzlů). Kompilovat lze i na loginu**

- středa – 09.11.

```
qsub -q R1572270 -A DD-21-22 -I -X -l walltime=1:00:0 -l select=1
```

- pátek – 11.11.

```
qsub -q R1572272 -A DD-21-22 -I -X -l walltime=1:0:0 -l select=1
```

- Kdykoliv jindy

```
qsub -q qgpu -A DD-21-22 -I -X -l walltime=1:0:0 -l select=1
```

- **Natáhněte moduly**

```
m1 CUDA/11.1.1-GCC-10.2.0 Qt5/5.14.2-GCCcore-10.2.0
```

1. Zkopírujte zdrojové kódy na Barboru

2. Přeložte knihovnu libwb – bude se hodit pro další ukázky

```
cd libwb  
make -j  
make libwb.a
```

3. Implementaci a testování funkčnosti můžete dělat v CVT, benchmarkování a profilování na Karolíně pomocí skriptu run.pbs

HISTOGRAM

- **Cílem je spočítat četnost výskytu znaků v textovém řetězci s využitím atomických operací nad globální pamětí**
- **Otevřete soubor `template.cu`**
 - Doplněte kernel `cudaHistogramGlobal`
 - Vycházejte z předpokladu, že je pole větší než počet vláken v gridu.
 - Spočtete histogram pomocí atomických operací nad globální pamětí.
 - Doplněte funkce `histogramGlobal`
 - Funkce nejprve nuluje histogram v globální paměti.
 - Následně se volá kernel.

Překlad a spuštění

CVT (bez profil)

```
make  
make gen  
  
make run{0-5}  
make prof{0-5}
```

Karolína (profil)

```
qsub run.pbs
```

Upravte soubor:

```
make  
make prof{0-5}
```

- **Cílem je spočítat četnost výskytu znaků v textovém řetězci s využitím atomických operací nad globální pamětí**
- **Otevřete soubor `template.cu`**
 - Doplněte kernel `cudaHistogramShared`
 - Vycházejte z předpokladu, že je pole větší než počet vláken v gridu.
 - Definujte pole `sharedHistogram` jako externí proměnnou v SM.
 - Vynulujte histogram ve sdílené paměti.
 - Spočtete histogram ve sdílené paměti.
 - Pomocí atomických operací aktualizujte histogram v globální paměti.
 - Doplněte funkce `histogramShared`
 - Funkce nejprve nuluje histogram v globální paměti.
 - Následně se volá kernel. Zde je nutné specifikovat počet bytů v SM.

Překlad a spuštění CVT (bez profil)

```
make  
make gen  
  
make run{0-5}  
make prof{0-5}
```

Karolína (profil)

```
qsub run.pbs
```

Upravte soubor:

```
make  
make prof{0-5}
```


- **Cílem je spočítat součet hodnot vektoru pomocí redukce ve sdílené paměti**
- **Otevřete soubor `template.cu`**
 - Doplňte kernel `cudaReduction`
 - Vycházejte z předpokladu, že je pole větší než počet vláken v gridu.
 - Definujte pole `partialSum` jako externí proměnnou v SM.
 - Vynulujte dílčí sumy ve sdílené paměti.
 - Proveďte paralelní redukci nad sdílenou pamětí. Pozor na správně nastavený stride.
 - Pomocí atomických operací aktualizujte výslednou hodnotu v globální paměti. Toto provede vždy jedno vláko v blocku.

Překlad a spuštění

CVT (bez profil)

```
make  
make gen  
  
make run{0-5}  
make prof{0-5}
```

Karolína (profil)

```
qsub run.pbs
```

Upravte soubor:

```
make  
make prof{0-5}
```

- **Cílem je spočítat součet hodnot vektoru pomocí redukce přes registry a warp-synchronního programování**
- **Otevřete soubor `template.cu`**
 - Doplněte kernel `cudaReduction`
 - Vycházejte z předpokladu, že pole je větší než počet vláken v gridu.
 - Každé vlákno si spočte dílčí sumu v registru.
 - Proveďte paralelní redukci pomocí shufle operací nad registry. Pozor na správně nastavený stride.
 - Pomocí atomických operací aktualizujte výslednou hodnotu v globální paměti. Toto provede vždy jedno vlákno v blocku.

Překlad a spuštění CVT (bez profil)

```
make  
make gen  
  
make run{0-5}  
make prof{0-5}
```

Karolína (profil)

```
qsub run.pbs
```

Upravte soubor:

```
make  
make prof{0-5}
```

**Vzorové řešení zveřejníme v
pátek**