

### L3. Lygiagretusis programavimas grafiniams procesoriams

Duomenų failai naudojami to paties formato, kaip L1 ir L2 metu, tačiau minimalus elementų kiekis faile – 250 (galima ir dubliuoti duomenis pridedant vienodų elementų). Sumažinto duomenų failo pavyzdys:

```
{
  "students": [
    {"name": "Antanas", "year": 1, "grade": 6.95},
    {"name": "Kazys", "year": 2, "grade": 8.65},
    {"name": "Petras", "year": 2, "grade": 7.01},
    {"name": "Sonata", "year": 3, "grade": 9.13},
    {"name": "Jonas", "year": 1, "grade": 6.95},
    {"name": "Martynas", "year": 3, "grade": 9.13},
    {"name": "Artūras", "year": 2, "grade": 7.01},
    {"name": "Vacys", "year": 2, "grade": 8.65},
    {"name": "Robertas", "year": 3, "grade": 6.43},
    {"name": "Mykolas", "year": 1, "grade": 6.95},
    {"name": "Aldona", "year": 3, "grade": 9.13},
    {"name": "Asta", "year": 2, "grade": 7.01},
    {"name": "Viktoras", "year": 2, "grade": 8.65},
    {"name": "Artūras", "year": 5, "grade": 8.32},
    {"name": "Vytautas", "year": 3, "grade": 7.85},
    {"name": "Jonas", "year": 1, "grade": 6.95},
    {"name": "Zigmas", "year": 3, "grade": 9.13},
    {"name": "Artūras", "year": 2, "grade": 7.01},
    {"name": "Simas", "year": 3, "grade": 6.43}
  ]
}
```

Duomenis, nuskaitytus iš failo, padalinti pasirinktam GPU gijų kiekiui (duomenų kiekis turi **nesidalinti** iš bendro gijų kiekio; reikia naudoti bent du blokus, gijų kiekis bloke turi būti 32 kartotinis). Kiekvienam įrašui iš savo duomenų rinkinio GPU gija suskaičiuoja pasirinktą **tekstinio tipo** rezultatą ir jei jis tenkina pasirinktą sąlygą, įrašo į bendrą visoms gijoms rezultatų masyvą (galima nustatyti, kiek simbolių išskirti vienam rezultatų įrašui, ir jei liko neužpildytos vietos, ją užpildyti tarpais). Jei ne visi elementai atitinka filtro sąlygą, rezultatuose tarpų likti neturi — visi rezultatai turi būti rašomi į rašymo metu esančią pirmą laisvą masyvo vietą.

Programą realizuoti naudojant CUDA gijas ir CUDA atminties valdymo funkcijas.

Rezultatų failo pavyzdys (vardas pakeistas į didžiąsias raides, prijungti metai bei pagal pažymį nustatytas įvertinimas nuo A iki F; atrinkti įrašai, kurių pirma raidė abėcėlėje yra toliau, nei P):

```
SONATA-3A
VACYS-2B
ROBERTAS-3D
VIKTORAS-2B
VYTAS-3C
ZIGMAS-3A
SIMAS-3D
```

#### Testas

Testas atliekamas Moodle aplinkoje laboratorinių darbų metu. Testo klausimai sudaryti iš teorijos temų „C ir C++ atminties valdymas“, „Lygiagretusis programavimas CUDA“ ir „Thrust“. Testui skiriama 20 minučių.

### L3 programų vertinimas

- L3a — 6 taškai (15 paskaita, 16 sav. pagal AIS)
- Kontrolinis — 4 taškai (16 paskaita, 17 sav. pagal AIS)

---

LD programų veikimą demonstruoti užsiėmimų laiku pagal tvarkaraštį, programų, duomenų ir rezultatų failus pateikti iki atsiskaitymo savaitės pabaigos.

---

Norint atlikti darbą reikalinga NVidia vaizdo plokštė.

### **CUDA diegimas savo kompiuteryje su NVidia GPU**

CUDA galima atsisiųsti iš NVidia puslapio. Svarbu pasitikrinti, kad pasirinkta CUDA versija palaikytų kompiuteryje esantį GPU.

### **CUDA serveris**

CUDA serveris pasiekiamas per SSH adresu `culab.int.ktu.lt`, prisijungimo duomenys tokie patys, kokie buvo sugeneruoti prisijungimui prie MPI serverio. Neturintys prisijungimo duomenų gali kreiptis į savo laboratorinių darbų dėstytoją. Serveryje įdiegta CUDA 11.6 versija, programos kompiliavimui žr. „CUDA kompiliavimas Linux aplinkoje“. Serveryje yra įdiegta `nlohmann-json` biblioteka.

### **CUDA projekto kūrimas Visual Studio**

Visual Studio kuriant naują projektą reikia pasirinkti CUDA projektą (pasirinkti programavimo kalbų laukelyje arba susirasti paieškoje).

### **CUDA kompiliavimas Linux aplinkoje**

Naudojamas kompiliatorius NVCC. Naudojimas analogiškas gcc, pvz:

```
nvcc program.cu -o program
```

Naudojant CUDA serverį reikia kompiliuoti naudojant g++-10 kompiliatorių:

```
nvcc -ccbin=/usr/bin/g++-10 --std=c++17 main.cu
```