

Laboratoriniai darbai

Domantas Keturakis

Lapkritis 2024

TURINYS

UŽDUOTIS #1

Pirma dalis

Lygiagretinimo galimybių analizė	2
Sprendimo lygiagretinimas	4
Rezultatai	5

Antra dalis

Rezultatai	6
------------------	---

UŽDUOTIS #2

Pirma dalis

Pirmas bandymas	7
<i>Rezultatai</i>	
<i>Antras bandymas</i>	

Antra dalis

Užduotis	12
Lygiagretinimas	12
Rezultatai	14

PRIEDAI

UŽDUOTIS #1

Pirma dalis

Lygiagretinimo galimybių analizė

Pirmasis skaičiavimo ciklas (Pav. 1) palyginus užtrunka nedaug laiko (apie 0.002 sekundės). Praktiniams tikslams, jį galima ignoruoti.

```
1 for (int i=0; i<numX; i++) {  
2     X[i] = i;  
3     bestX[i] = i;  
4 }  
5 u = evaluateSolution(X);  
6 bestU = u;
```

Pav. 1: Pradinės naujo ir geriausio sprendinių reikšmių apskaičiavimas

Didžiąją dalį laiko užima šis ciklas:

```
1 while (increaseX(X, numX-1, numCL) == true) {  
2     u = evaluateSolution(X);  
3     if (u > bestU) {  
4         bestU = u;  
5         for (int i=0; i<numX; i++) bestX[i] = X[i];  
6     }  
7 }
```

Pav. 2: Visų galimų sprendinių perrinkimas

Šio ciklo iš esmės “aklai” parelizuoti negalima, nes reikia atsižvelgti į tai kad:

- `increaseX` - keičia masyvo `X` reikšmę (Pav. 3), ir ne tik `index`-ąjį elementą, rekursyviai kviesdamas save sumažina `index` reikšmę vienu, t.y. iškvietus `increaseX` visos masyvos reikšmės yra keičiamos. To pasekmė, kad `index`-ojo elemento skaičiavimo negalima paskirstyti skirtingoms gijoms, kitaip vėlesnėms gijoms reiktų laukti, kol praeita gija baigs savo darbą, visiškai nustelbiant paralelizavimo naudą.
- `increaseX` skaičiavimai priklauso vienas nuo kito, t.y. norint apskaičiuoti `X` reikšmę n -ame ciklo vykdyme, reikia pirma apskaičiuoti `X` reikšmę $(n - 1)$ -ame ciklo vykdyme. Analogiškai negalima lygiagretinti nes kitos gijos lauktų, kol praeita gija baigs savo darbą.
- `if (u > bestU) { ... }` irgi gali tik vienas ciklas vienu metu, nes `bestU` ir `X` pakeitimas turi būti atliekamas “žingsniu” - t.y. atomiškai.

Iš esmės neperrašius `increaseX`, `X` skaičiavimų lygiagretinti nėra praktiška.

```

1  int increaseX(int *X, int index, int maxindex) {
2      if (X[index]+1 < maxindex-(numX-index-1)) {
3          X[index]++;
4      }
5      else {
6          if ((index == 0) && (X[index]+1 == maxindex-(numX-index-1))) {
7              return 0;
8          }
9          else {
10             if (increaseX(X, index-1, maxindex)) X[index] = X[index-1]+1;
11             else return 0;
12          }
13      }
14      return 1;
15 }

```

Pav. 3: Funkcija increaseX

Tuo tarpu funkcija evaluateSolution (Pav. 4) nekeičia jokių globalių kintamųjų ar savo argumentų. Analitiškai žiūrint galima spėti, kad čia ir didžioji dalis skaičiavimo laiko yra sugaištama. Teorinis šios funkcijos *big O* yra $O(\text{numDP} \cdot \text{numX})$, tuo tarpu increaseX rekursyviai save gali iškviesti daugiausiai numX kartų.

```

1  double evaluateSolution(int *X) {
2      double U = 0; double totalU = 0;
3      int bestPF, bestX;
4      double d;
5
6      for (int i=0; i<numDP; i++) {
7          totalU += demandPoints[i][2];
8          bestPF = 1e5;
9          for (int j=0; j<numPF; j++) {
10             d = HaversineDistance(i, j);
11             if (d < bestPF) bestPF = d;
12          }
13          bestX = 1e5;
14          for (int j=0; j<numX; j++) {
15             d = HaversineDistance(i, X[j]);
16             if (d < bestX) bestX = d;
17          }
18
19          if (bestX < bestPF) U += demandPoints[i][2];
20          else if (bestX == bestPF) U += 0.3*demandPoints[i][2];
21      }
22      return U/totalU*100;
23 }

```

Pav. 4: Funkcija evaluateSolution

Sprendimo lygiagretinimas

Dėl anksčiau išvardintų priežasčių `increaseX` apskaičiavimas išskiriamas į `critical` bloką, tam, kad tik viena gija galėtų modifikuoti `X` reikšmę vienu metu. Apskaičiavus ir atnaujinus `X`, kiekviena gija susikuria savo `X` kopiją - `localX`. Šią kopiją galima naudoti `evaluateSolution` nes jiniai ne bus keičiama. Kiekviena gija taip pat gauna `u` kopiją į kurią įrašo `evaluateSolution` apskaičiuotą reikšmę. Ciklo gale vėl naudojamas `critical`, tam kad tik viena gija vienu metu galėtų įvertinti `u > bestU` ir pakeisti `bestU` ir `bestX` reikšmes.

```
1  bool increased = true;
2  int *manyXs = new int[NUM_THREADS * numX];
3
4  #pragma omp parallel private(u)
5  {
6      while (increased) {
7          int thread_id = omp_get_thread_num();
8          int *localX = manyXs + (thread_id * numX);
9
10         #pragma omp critical(increaseX)
11         {
12             increased = increaseX(X, numX-1, numCL);
13             memcpy(localX, X, sizeof(int) * numX);
14         }
15
16         u = evaluateSolution(localX);
17
18         #pragma omp critical(best)
19         {
20             if (u > bestU) {
21                 bestU = u;
22                 memcpy(bestX, localX, sizeof(int) * numX);
23             }
24         }
25     }
26 }
```

Pav. 5: Sulygiagretintas visų galimų sprendinių perrinkimas

Rezultatai

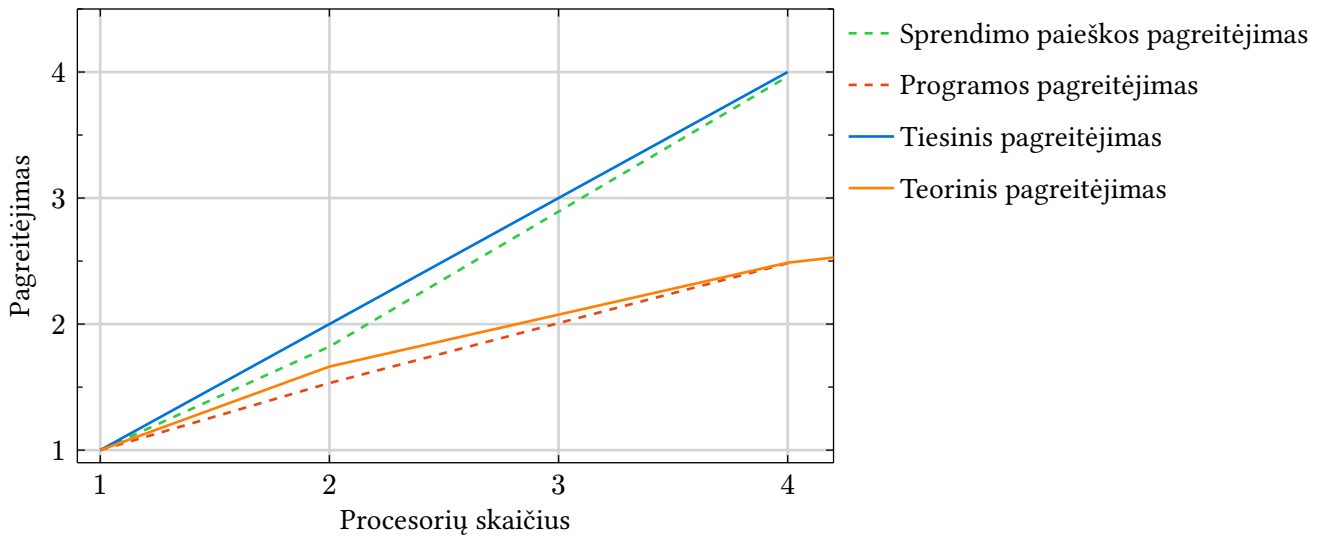


Fig. 1: Pagreitėjimo ir Procesorių skaičiaus santykis, kai matricos skaičiavimas nuoseklus

Iš Fig. 1 matoma, kad lygiagretinamos dalies pagreitėjimas seka tiesinį pagreitėjimą. O visos programos pagreitėjimas seka teorinį pagreitėjimą nusakytą pagal Amdalo dėsnį.

Antra dalis

Šioje vietoje for direktyva atrodo lengvai pritaikoma, kadangi atstumų matricos kiekvieną eilutę galima apskaičiuoti nepriklausomai nuo to ar praeitos eilutės yra apskaičiuotos. Tačiau, pirmos eilutės reikia žymiai mažiau laiko negu paskutines, todėl pritaikyta *guided* paskirstymo (angl. *scheduling*) direktyva. Tai leidžia efektyviau paskirstyti ciklo darbą per skirtingas gijas.

```
1 distanceMatrix = new double*[numDP];
2 #pragma omp parallel
3 {
4     #pragma omp for schedule(guided)
5     for (int i=0; i<numDP; i++) {
6         distanceMatrix[i] = new double[i+1];
7         for (int j=0; j<=i; j++) {
8             distanceMatrix[i][j] = HaversineDistance(
9                 demandPoints[i][0],
10                demandPoints[i][1],
11                demandPoints[j][0],
12                demandPoints[j][1]
13            );
14        }
15    }
16 }
```

Pav. 6: Sulygiagretintas atstumų matricos skaičiavimas

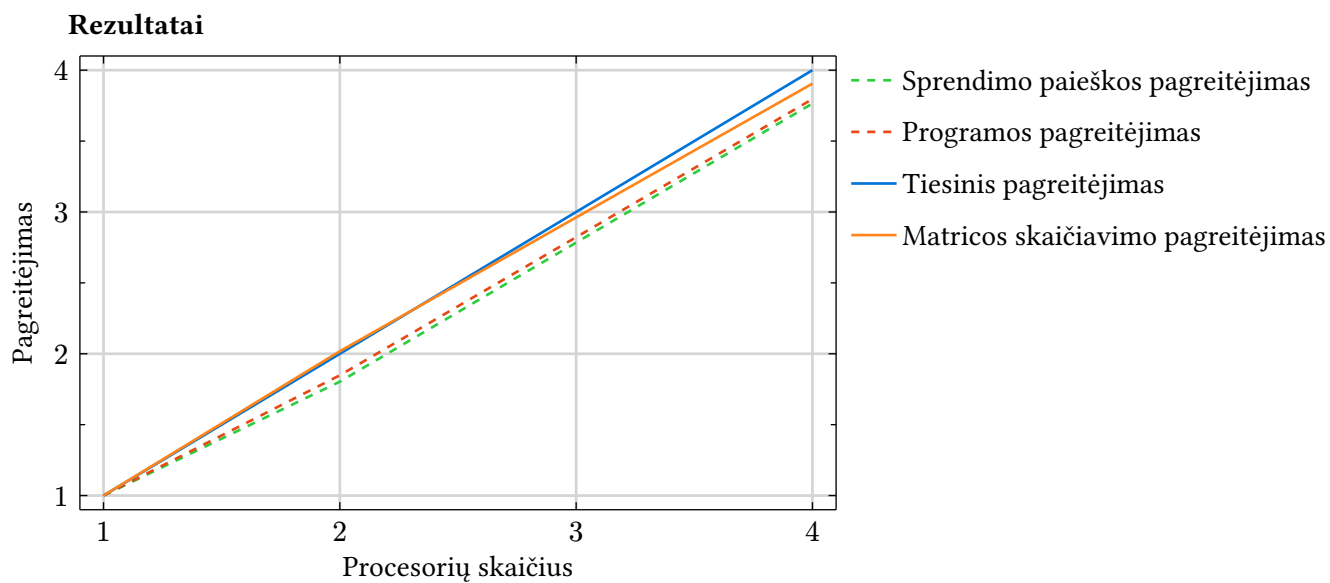


Fig. 2: Pagreitėjimo ir Procesorių skaičiaus santykis, kai matricos skaičiavimas sulygiagretintas

Iš Fig. 2 matoma, kad matricos skaičiavimo ir sprendimo ieškojimo pagreitėjimas seka tiesinę kreivę, kas matosi ir visos programos pagreitėjime, kuri irgi seka tiesinę kreivę.

UŽDUOTIS #2

Sulygiagretinti skaičiavimus naudojantis MPI biblioteka.

Pirma dalis

Sulygiagretinti sprendinio skaičiavimus.

Pirmas bandymas

Procesus galima paskirstyti taip, kad yra vienas pagrindinis (*main*) procesas ir likę - darbuotojai (*workers*). Kur pagrindinis procesas skaičiuoja atnaujintas X reikšmes ir jas išsiunčia kitiems procesams.

```
1  int world_size; MPI_Comm_size(MPI_COMM_WORLD, &world_size);
2  int world_rank; MPI_Comm_rank(MPI_COMM_WORLD, &world_rank);
3
4  // Iškirpta: loadDemandPoints(), distanceMatrix skaičiavimas, MPI_Buffer_attach,
   t.t.
5  MPI_Barrier(MPI_COMM_WORLD);
6
7  bool increased = true;
8  while (increased) {
9      int sends = 1;
10     if (world_rank == 0) {
11         for(int ix = 1; ix < world_size; ++ix) {
12             increased = increaseX(X, numX - 1, numCL);
13
14             MPI_Send(X, numX, MPI_INT, ix, DATA_X, MPI_COMM_WORLD);
15             sends = ix + 1;
16
17             if (!increased) {
18                 for (int ix = 1; ix < world_size; ++ix) {
19                     MPI_Bsend(X, 0, MPI_BYTE, ix, SIGNAL_DONE, MPI_COMM_WORLD);
20                 }
21                 break;
22             }
23         }
24     }
25     // ...
26 }
```

Radus galutinę X reikšmę pagrindinis procesas išsiučia žinutę su žyme SIGNAL_DONE, kad pranešti procesams darbuotojams, kad jie gali baigti savo darbą.

Tuo tarpu procesai darbuotojai laukia naujos X reikšmės, jos sulaukę, apskaičiuoja u reikšmę ir išsiunčia ją, kartu su X, pagrindiniam procesui. Taip pat jie patikrina ar negavo žinutės su SIGNAL_DONE žyma, kurios pagalba jie sužino, kad daugiau negaus X reikšmių ir todėl gali baigti savo darbą.

```
1  while(increased) {
2      // ...
3      if (world_rank != 0) {
4          MPI_Status status;
5          MPI_Probe(0, MPI_ANY_TAG, MPI_COMM_WORLD, &status);
6
7          if (status.MPI_TAG == SIGNAL_DONE) {
8              MPI_Recv(NULL, 0, MPI_BYTE, 0, SIGNAL_DONE, MPI_COMM_WORLD, MPI_STATUS_IGNORE);
9              break;
10         }
11
12         MPI_Recv(X, numX, MPI_INT, 0, DATA_X, MPI_COMM_WORLD, MPI_STATUS_IGNORE);
13         u = evaluateSolution(X);
14
15         MPI_Send(&u, 1, MPI_DOUBLE, 0, DATA_U, MPI_COMM_WORLD);
16         MPI_Send(X, numX, MPI_INT, 0, DATA_X, MPI_COMM_WORLD);
17     }
18     // ...
19 }
```

Tuo tarpu pagrindinis laukia tiek u ir X reikšmių kiek išsiuntė procesams darbuotojams, gavęs reikšmę su didesne u reikšme atnauja savo bestU ir bestX kintamuosius.

```
1  while(increased) {
2      // ...
3      if (world_rank == 0) {
4          for (int ix = 1; ix < sends; ++ix) {
5              MPI_Recv(&u, 1, MPI_DOUBLE, ix, DATA_U, MPI_COMM_WORLD, MPI_STATUS_IGNORE);
6              MPI_Recv(X, numX, MPI_INT, ix, DATA_X, MPI_COMM_WORLD, MPI_STATUS_IGNORE);
7
8              if (u > bestU) {
9                  bestU = u;
10                 memcpy(bestX, X, sizeof(int) * numX);
11             }
12         }
13     }
14 }
```


Rezultatai

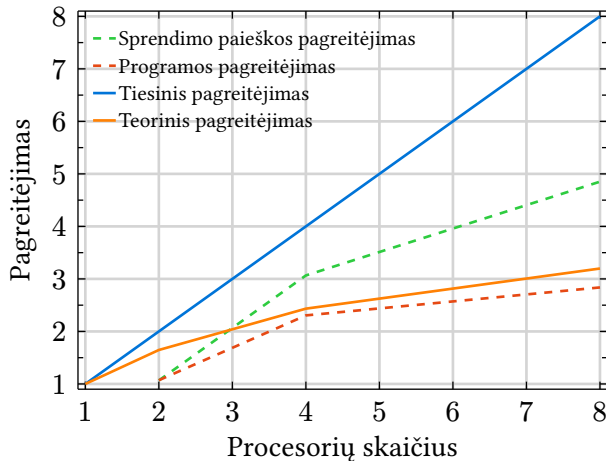


Fig. 3: Pagreitėjimo ir Procesorių skaičiaus santykis

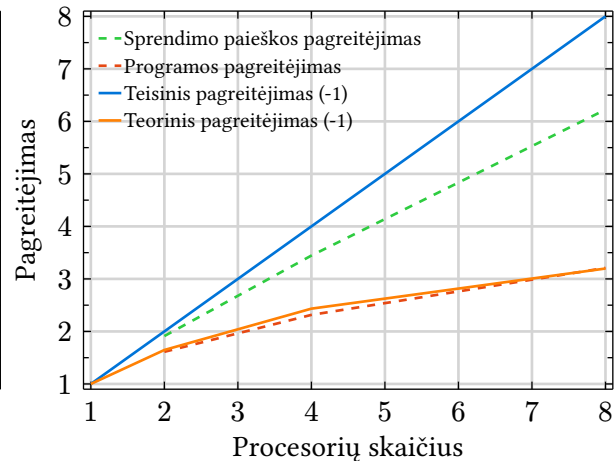


Fig. 4: Pagreitėjimo ir Procesorių skaičiaus santykis, neskaičiuojant pagrindinio proceso

Šis sprendimas nėra prastas (Fig. 3), bet pilnai neišnaudoja pagrindinio proceso, galima daryti išvadas, kad jis dažnai neturi darbo ir laukio kol galės kitiem procesam išsiųsti atnaujintas X reikšmes. Neskaičiuojant pagrindinio proceso (Fig. 4), t.y. skaičiuojant pagreitėjimą su 8 procesais ištikrųjų paleidžiami 9 procesai, praktiškai pasiekiamas teorinis pagreitėjimas.

Antras bandymas

Visgi galima padaryti taip, kad pagrindinis procesas irgi skaičiuotu X reikšmes. Kad procesai darbuotojai nelauktų, kol pagrindinis procesas baigs skaičiuoti savo dalį, nebelaukiama atsakymo iš procesų darbuotojų, X siuntimui pasitelkiamas `MPI_Bsend`.

```
1  while (true) {
2      if (world_rank == 0 && increased) {
3          for(int ix = 1; ix < world_size; ++ix) {
4              increased = increaseX(X, numX - 1, numCL);
5              MPI_Bsend(X, numX, MPI_INT, ix, DATA_X, MPI_COMM_WORLD);
6              sends += 1;
7
8              if (!increased) {
9                  for (int ix = 1; ix < world_size; ++ix) {
10                     MPI_Bsend(&dummy_load, 1, MPI_INT, ix, SIGNAL_DONE, MPI_COMM_WORLD);
11                 }
12                 break;
13             }
14         }
15     }
16     // ...
17 }
```

Toliau while cikle pridedamas paskaičiavimas pagrindiniam procesui ir `SIGNAL_DONE` išsiuntimas, jeigu šiame cikle baigtusi X skaičiavimas:

```

1  while(true) {
2      // ...
3      if (world_rank == 0 && increased) {
4          increased = increaseX(X, numX - 1, numCL);
5
6          u = evaluateSolution(X);
7          if (u > bestU) {
8              bestU = u;
9              memcpy(bestX, X, sizeof(int) * numX);
10         }
11
12         if (!increased) {
13             for (int ix = 1; ix < world_size; ++ix) {
14                 MPI_Bsend(&dummy_load, 1, MPI_INT, ix, SIGNAL_DONE, MPI_COMM_WORLD);
15             }
16         }
17     }
18     // ...
19 }

```

Procesams darbuotojams pagrįsde ne daug kas keičiasi. Jie irgi pakeičia MPI_Send į MPI_Bsend.

```

1  if (world_rank != 0) {
2      int master_done = WRP_Check_for(0, SIGNAL_DONE, MPI_COMM_WORLD);
3      int master_sent_X = WRP_Check_for(0, DATA_X, MPI_COMM_WORLD);
4
5      if (master_sent_X) {
6          MPI_Recv(X, numX, MPI_INT, 0, DATA_X, MPI_COMM_WORLD, MPI_STATUS_IGNORE)
7
8          u = evaluateSolution(X);
9          if (bestU < u) { bestU = u; bestX = memcpy(bestX, X, sizeof(int) * numX)
10
11          MPI_Bsend(&u, 1, MPI_DOUBLE, 0, DATA_U, MPI_COMM_WORLD);
12          MPI_Bsend(X, numX, MPI_INT, 0, DATA_X, MPI_COMM_WORLD);
13      }
14
15      if (master_done && !master_sent_X) { break; }
16  }

```

Čia WRP_Check_for(int source, int tag, MPI_Comm comm) (apibrėžimas Pav. 7) viduje naudoja MPI_Iprobe(int source, int tag, MPI_Comm communicator, int* flag, MPI_Status* status) ir gražina flag dalį.

```

1  if (world_rank == 0) {
2      if (!increased && receives == sends) { break; }
3
4      MPI_Status status;
5      int worker_sent_X;
6      MPI_Iprobe(MPI_ANY_SOURCE, DATA_X, MPI_COMM_WORLD, &worker_sent_X, &status)
7      while(worker_sent_X) {
8          MPI_Recv(&copy_u, 1, MPI_DOUBLE, status.MPI_STATUS, DATA_U, MPI_COMM_WORLD,
9          MPI_STATUS_IGNORE);
10         MPI_Recv(copy_X, numX, MPI_INT, status.MPI_STATUS, DATA_X, MPI_COMM_WORLD,
11         MPI_STATUS_IGNORE);
12         receives += 1;
13
14         if (copy_u > bestU) {
15             bestU = copy_u;
16             memcpy(bestX, copy_X, sizeof(int) * numX);
17         }
18         worker_sent_X = WRP_Check_for(MPI_ANY_SOURCE, DATA_X, MPI_COMM_WORLD);
19     }
20 }

```

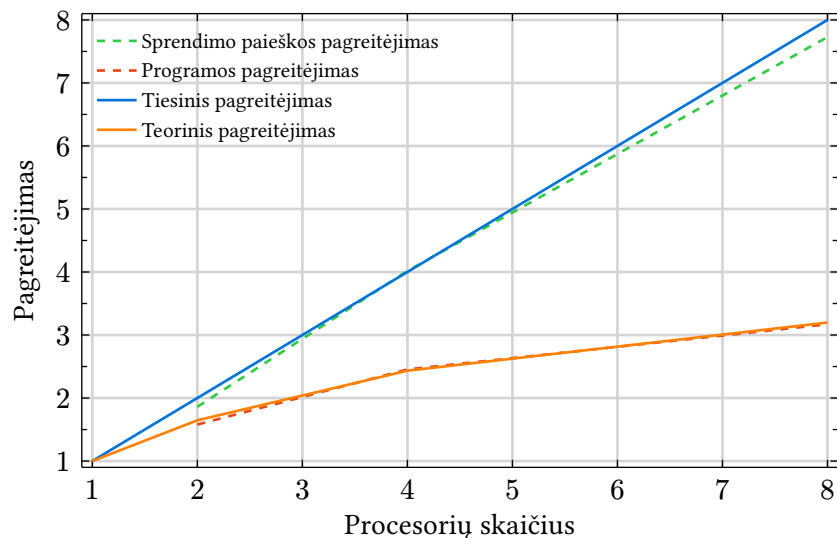


Fig. 5: Pagreitėjimo ir Procesorių skaičiaus santykis

Šis metodas pasiekia teorinį pagreitėjimą be papildomo proceso.

Antra dalis

Užduotis

Sulygiagretinti matricos skaičiavimus.

Lygiagretinimas

Tam, kad gerai sulygiagretinti, reikia paskirstyti skaičiavimus taip, kad kiekvienas procesas turėtų po tiek pat darbo. Tam panaudojama funkcija `lengths` (Jos apibrėžimas - Pav. 8), kuri parenka tinkamus intervalus, taip, kad kiekvienas procesas paskaičiuotu apytiksliai tiek pat matricos elementų.

```
1  int main() {
2      // ...
3      int *lens = lengths(numDP, world_size);
4      int *counts = calloc(sizeof(int), world_size);
5      for (int ix = 0; ix < world_size; ++ix) {
6          counts[ix] = (lens[ix + 1] - lens[ix]) * numDP;
7      }
8
9      int *disps = calloc(sizeof(int), world_size);
10     for (int ix = 0; ix < world_size; ++ix) {
11         disps[ix] = lens[ix] * numDP;
12     }
13
14     distanceMatrix = calloc(sizeof(double), numDP * numDP);
15     for (int i = lens[world_rank]; i < lens[world_rank + 1]; i++) {
16         for (int j = 0; j <= i; j++) {
17             distanceMatrix[numDP * i + j] =
18                 HaversineDistance4(demandPoints[i][0], demandPoints[i][1],
19 demandPoints[j][0], demandPoints[j][1]);
19         }
20     }
21     // ...
```

Čia svarbu paminėti, kad `distanceMatrix` tipas buvo pakeistas iš `double **` į `double *` ir atmintis visai matricai priskiriamia iškart `calloc(sizeof(double), numDP * numDP)`, prieiga prie matricos irgi atitinkamai pakeista iš `distanceMatrix[i][j] = v` į `distanceMatrix[i * numDP + j] = v`. Kadangi šis pakeitimas iš esmės pakeičia greitaveikos savybės (eksperimentiškai visos programos skaičiavimas sumažėja nuo 22s iki 17s), tai pradinė programa, su kuria lygininama lygiagretinta programa, irgi buvo pakeista, tam kad palyginimai būtų teisingi.

Kai procesas baigia savo dalį, jis nusiunčia nusiunčia kitiems procesams savo dalį ir laukia kitų dalių naudodant `MPI_Allgatherv`.

```

1  MPI_Allgatherv(
2      distanceMatrix + disps[world_rank],
3      counts[world_rank],
4      MPI_DOUBLE,
5      distanceMatrix,
6      counts,
7      disps,
8      MPI_DOUBLE,
9      MPI_COMM_WORLD
10 )

```

Galima pakeisti duomenų apsikeitimą naudojant `MPI_Iallgatherv`, pridedant prieš kiekvieną `evaluateSolution` (nes tik ten ir naudojami matricos duomenys) šią eilutę, tam, kad duomenys pirma būtų pilnai surinkti.

```

1  if (first_run) { MPI_Wait(&req, MPI_STATUS_IGNORE); first_run = false; }

```

Tiesa tada laiko matavimas tampa šiek tiek netikslus, nes iškart po `MPI_Iallgatherv` nuskaitomas laikas. Galima sakyti, kad matricos skaičiavimas baigiasi tada, kai pirmą kartą prireikia matricos duomenų, t.y.:

```

1  if (first_run) { MPI_Wait(&req, MPI_STATUS_IGNORE); first_run = false; t_matrix
    = getTime(); }

```

Rezultatai

TODO: actual data

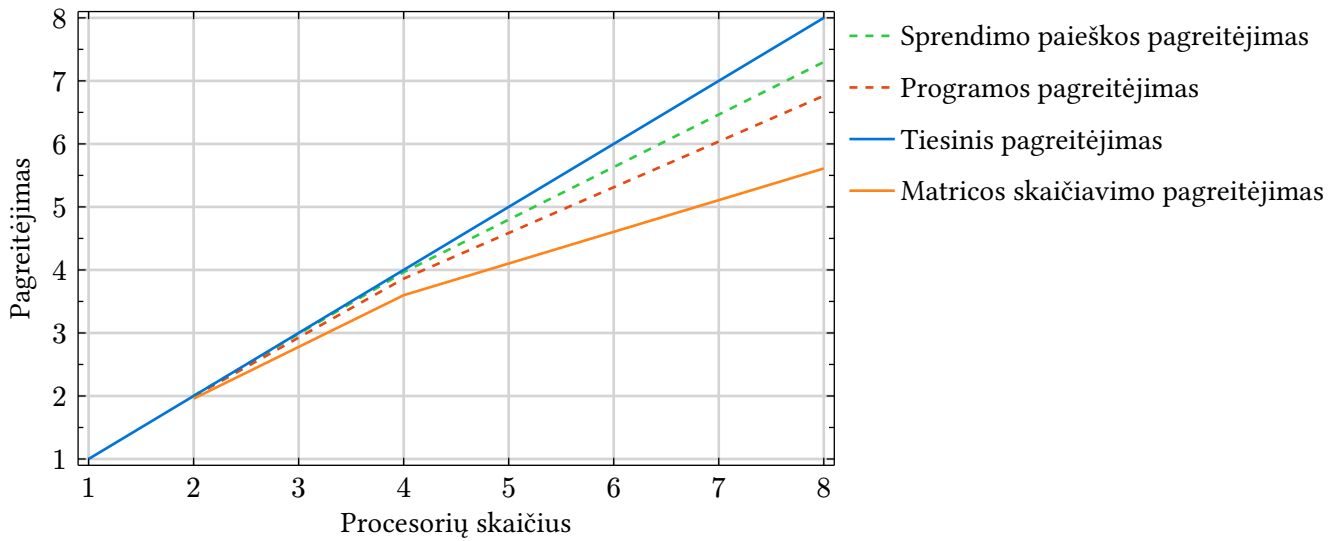


Fig. 6: Pagreitėjimo ir Procesorių skaičiaus santykis, kai matricos skaičiavimas sulygiagretintas (MPI_Allgather)

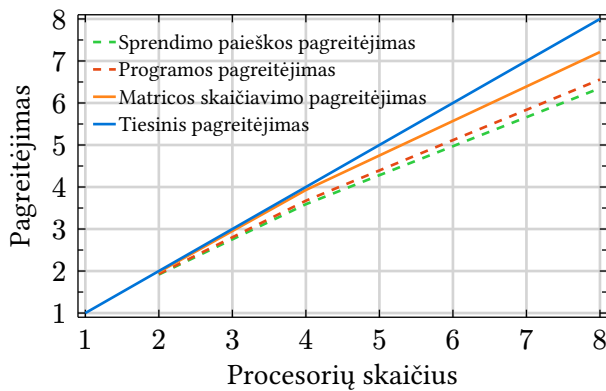


Fig. 7: Pagreitėjimo ir Procesorių skaičiaus santykis, kai matricos skaičiavimas sulygiagretintas (MPI_Iallgather), kai laikas nuskaitomas iškart po MPI_Iallgather

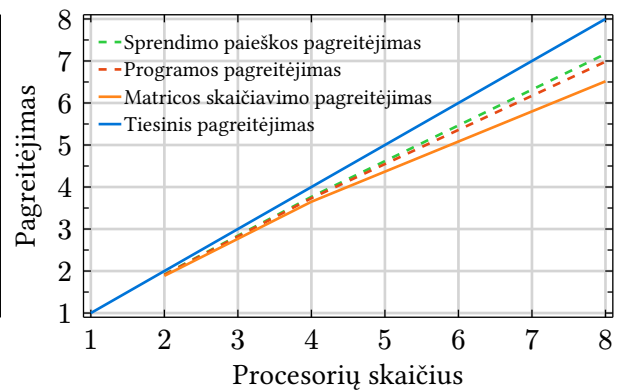


Fig. 8: Pagreitėjimo ir Procesorių skaičiaus santykis, kai matricos skaičiavimas sulygiagretintas (MPI_Iallgather), kai laikas nuskaitomas kai matricos praeikia pirmą kartą

PRIEDAI

```
1 int WRP_Check_for(int source, int tag, MPI_Comm comm) {  
2     MPI_Status status;  
3     int flag;  
4     MPI_Iprobe(source, tag, comm, &flag, &status);  
5  
6     return flag;  
7 }
```

Pav. 7: Optimalus intervalus parinkimas

```
1 int *lengths(int leg_length, int process_count) {  
2     int area = leg_length * leg_length / 2;  
3  
4     int small_area = area / process_count;  
5     int *lengths = calloc(sizeof(int), process_count + 1);  
6     for (int ix = 1; ix < process_count; ++ix) {  
7         lengths[ix] = (int) sqrt(2 * (ix) * small_area);  
8     }  
9     lengths[process_count] = leg_length;  
10  
11     return lengths;  
12 }
```

Pav. 8: Optimalus intervalus parinkimas