

VILNIAUS UNIVERSITETAS
Duomenų Mokslo ir Skaitmeninių Technologijų institutas
INFORMATIKOS INŽINERIJOS STUDIJŲ PROGRAMOS
DOKTORANTAS ALBERTAS JURGELEVIČIUS

Doktorantūros metinė ataskaita

2017 M. SPALIO MĖN. 1 D. – 2018 M. RUGSĖJO
MĖN. 30 D.

Tema

Disertacijos tema:

- **„VIEŠŲJŲ PASKIRSTYTŲ SKAIČIAVIMŲ PLATFORMA DIDELIŲ DUOMENŲ TYRYBAI“**

Darbo vadovas:

- **PROF. HABIL. DR. LEONIDAS SAKALAUSKAS**
- **Įstojimo į doktorantūrą metai: 2016 m. spalio mėn. 1 d.**
- **Doktorantūros baigimo metai: 2020 m. rugsėjo mėn. 30 d.**

Tyrimas

Tyrimo objektas:

- Viešieji paskirstyti skaičiavimai;
- Didelių duomenų tyrybos metodai.

Tyrimo tikslas:

- Sukurti viešųjų paskirstytų skaičiavimų platformą didelių duomenų tyrybai, atitinkančią įmonių bei organizacijų keliamus reikalavimus tokio pobūdžio platformoms.

Tyrimas

Tyrimo uždaviniai:

- Įvertinti viešųjų paskirstytų skaičiavimų modelio taikymo galimybes didelių duomenų tyrybai;
- Įvertinti esamas viešųjų paskirstytų skaičiavimų modeliu grįstas didelių duomenų tyrybos platformas:
 - alternatyvių didelių duomenų tyrybos metodų atžvilgiu;
 - įmonių bei organizacijų keliamų poreikių ir reikalavimų joms atžvilgiu;
- Parengti ir empiriniu būdu įvertinti viešųjų paskirstytų skaičiavimų metodu grįstą, didelių duomenų tyrybai skirtą platformos teorinį modelį, atitinkantį įmonių bei organizacijų keliamus poreikius ir reikalavimus.

Planuojami rezultatai:

- Viešųjų paskirstytų skaičiavimų platforma didelių duomenų tyrybai;
- Platformos taikymo praktiniams uždaviniams spręsti rezultatų apibendrinimas.

Ataskaitinių metų darbo planas

Išlaikyti egzaminus:

- Daugiamačių duomenų vizualizavimo metodai;
- Lygiagretieji ir paskirstytieji skaičiavimai;

Atlikti mokslinį tyrimą:

- Tyrimo metodikos sudarymas;
- Teorinis tyrimas;

Dalyvauti mokslinėje konferencijoje;

Parengti 1 publikaciją.

Išlaikyti egzaminai

Daugiamačių duomenų vizualizavimo metodai. Gautas įvertinimas 8 (gerai). Egzamino komisijos sudėtis: prof. habil. dr. Gintautas Dzemyda (pirmininkas), prof. dr. Olga Kurasova, dr. Viktor Medvedev;

Lygiagretieji ir paskirstytieji skaičiavimai. Gautas įvertinimas 10 (puikiai). Egzamino komisijos sudėtis: prof. dr. Julius Žilinskas (pirmininkas), prof. habil. dr. Gintautas Dzemyda, dr. Viktor Medvedev.

Skaityti pranešimai, bei publikuoti straipsniai

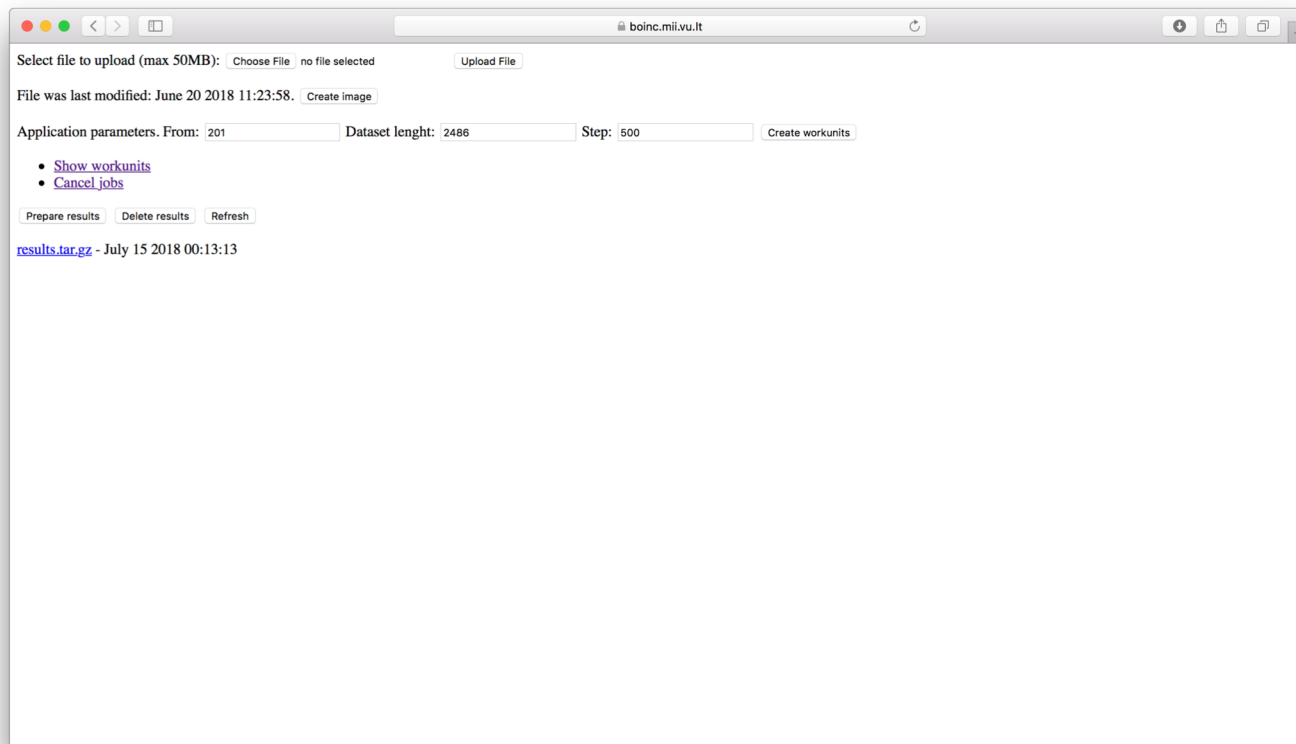
2017-11-30 dieną dalyvauta devintojoje tarptautinėje konferencijoje „Duomenų analizės metodai programų sistemoms“ (Druskininkai). Stendinio pranešimo tema: „**BOINC Based Enterprise Desktop GRID**“.

„**Big Data mining using public distributed computing**“. Straipsnis publikuotas tarptautiniame leidinyje „Information Technology and Control“ (DOI: <http://dx.doi.org/10.5755/j01.itc.47.2.19738>).

Kiti atlikti darbai

Viešųjų skaičiavimų sistemos duomenų ir užduočių paskirstymo bei kontrolės sistemos teorinio modelio tyrimo tikslais sukurtas BOINC karkaso bei Docker platformos pagrindu veikianti išskirstytų Octave užduočių vykdymo sistema duomenų analizei. Taip pat buvo paruošta sistemos techninė dokumentacija. Sistema pasiekama adresu <https://boinc.mii.vu.lt/octave-hmm/>.

Užduočių pateikimo ir rezultatų atsissiuntimo puslapis



Rezultatų peržiūros puslapis

boinc.mii.vu.lt

https://boinc.mii.vu.lt/octave-hmm/ Results

REPLACE WITH PROJECT NAME: Results

[log in](#)

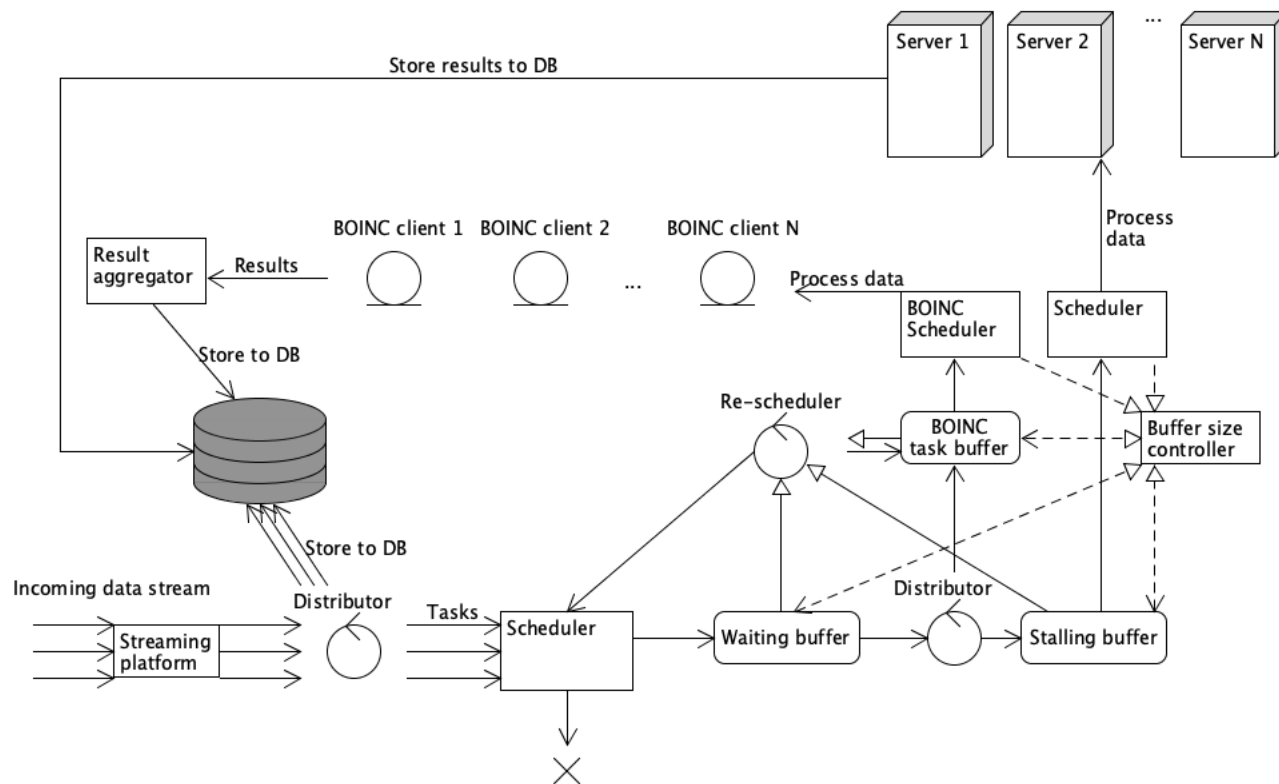
Query: **select * from result limit 20,20**

79 records match the query. Displaying 21 to 40.

[Summary](#) | [More detail](#) | [Return to main admin page](#)

result ID	WU ID	server state	outcome	client state	validate state	delete state	exit status	host (user)	app ver	received or deadline or created	CPU hours	granted credit
21	11	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 10:10:23 UTC	0.5	8.333
22	11	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 18:55:53 UTC	0.5	8.333
23	12	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 11:41:58 UTC	1.3	8.333
24	12	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 14:10:17 UTC	1.2	8.333
25	13	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 12:53:33 UTC	1.0	8.333
26	13	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 16:15:15 UTC	1.0	8.333
27	14	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 18:23:00 UTC	1.1	8.333
28	14	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt) windows_x86_64 (1)	4 Jul 2018, 19:56:30 UTC	1.0	8.333
29	15	Over [5]	Success [1]	Done [5]	Valid [1]	Deleted 0 (0x00000000)		3 (Albertas)	boinc2docker v1.07 (vbox64_mt)	4 Jul 2018, 17:18:37 UTC	1.0	8.333

Sukurtas teorinis modelis



Sudaryta tyrimo metodika

Mokslinis tyrimas bus atliekamas penkiais etapais. Pirmajame etape bus:

1. apžvelgiami didelių duomenų tyrybos metodai;
2. vertinamos viešųjų paskirstytų skaičiavimų modelio taikymo didelių duomenų tyrybai galimybės;
3. vertinamos esamos viešųjų paskirstytų skaičiavimų modeliu grįstos didelių duomenų tyrybos platformos.

Antrajame etape bus parengtas viešųjų paskirstytų skaičiavimų metodu grįstas, didelių duomenų tyrybai skirtas platformos teorinis modelis, atitinkantis pirmojo etapo metu sudarytų specifikacijų reikalavimus.

Sukurto teorinio modelio vertinimas bus išskaidytas į tris etapus atliekant empirinius tyrimus dinamiškoje darbo aplinkoje.

Paskutinio etapo metu bus atlikta statistinė duomenų analizė naudojantis statistiniu duomenų paketu SPSS, R Statistical Package ir Microsoft Office programomis. Remiantis šia analize bus atliktas sukurto teorinio modelio kiekybinis bei kokybinis vertinimas pagal minėtus rodiklius.

Kitų metų darbo planas

Remiantis sukurtu teoriniu modeliu realizuoti viešųjų paskirstytų skaičiavimų platformą didelių duomenų tyrybai;

Atlikti sukurtos viešųjų paskirstytų skaičiavimų platformos didelių duomenų tyrybai eksperimentinį tyrimą analizuojant:

- kompiuterių įjungtų į viešo skaičiavimo potinklį darbą;
- atliekamų skaičiavimų efektyvumą;

Nustatyti praktines sritis, kuriose būtų tinkama sukurtoji platforma;

Atlikti gautų duomenų analizę, bei parengti išvadas;

Parengti 1 publikaciją.