

VILNIAUS UNIVERSITETAS
Duomenų mokslo ir skaitmeninių technologijų institutas
INFORMATIKOS INŽINERIJOS STUDIJŲ PROGRAMOS
DOKTORANTAS ALBERTAS JURGELEVIČIUS

VIEŠŪJŲ PASKIRSTYTŲ SKAIČIAVIMŲ PLATFORMA DIDELIŲ DUOMENŲ TYRYBAI

4 METŲ DOKTORANTŪROS STUDIJŲ IR MOKSLINIŲ TYRIMŲ
ATASKAITA

Tema

Disertacijos tema:

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

Darbo vadovas:

- **PROF. HABIL. DR. LEONIDAS SAKALAUSKAS**

Konsultantas:

- **DR. VIRGINIJUS MARCINKEVIČIUS**

- Įstojimo į doktorantūrą metai: 2016 m. spalio mėn. 1 d.
- Doktorantūros baigimo metai: 2020 m. rugsėjo mėn. 30 d.

Tyrimas

Tyrimo objektas:

- Viešas paskirstytas skaičiavimas;
- Didelių duomenų tyrybos metodai.

Tyrimo tikslas:

- Sukurti viešo paskirstyto skaičiavimo platformą didelių duomenų tyrybai, atitinkančią įmonių bei organizacijų keliamus reikalavimus tokio pobūdžio platformoms.

Pagrindiniai tyrimo metodai taikomi disertacijoje – apžvalga bei eksperimentas. Tyrimo metu buvo dirbama naudojant sugeneruotus duomenų rinkinius, skaičiavimams naudojami kompiuteriai ir jų veiksmi buvo stebimi naudojant sistemų stebėjimo metodus.

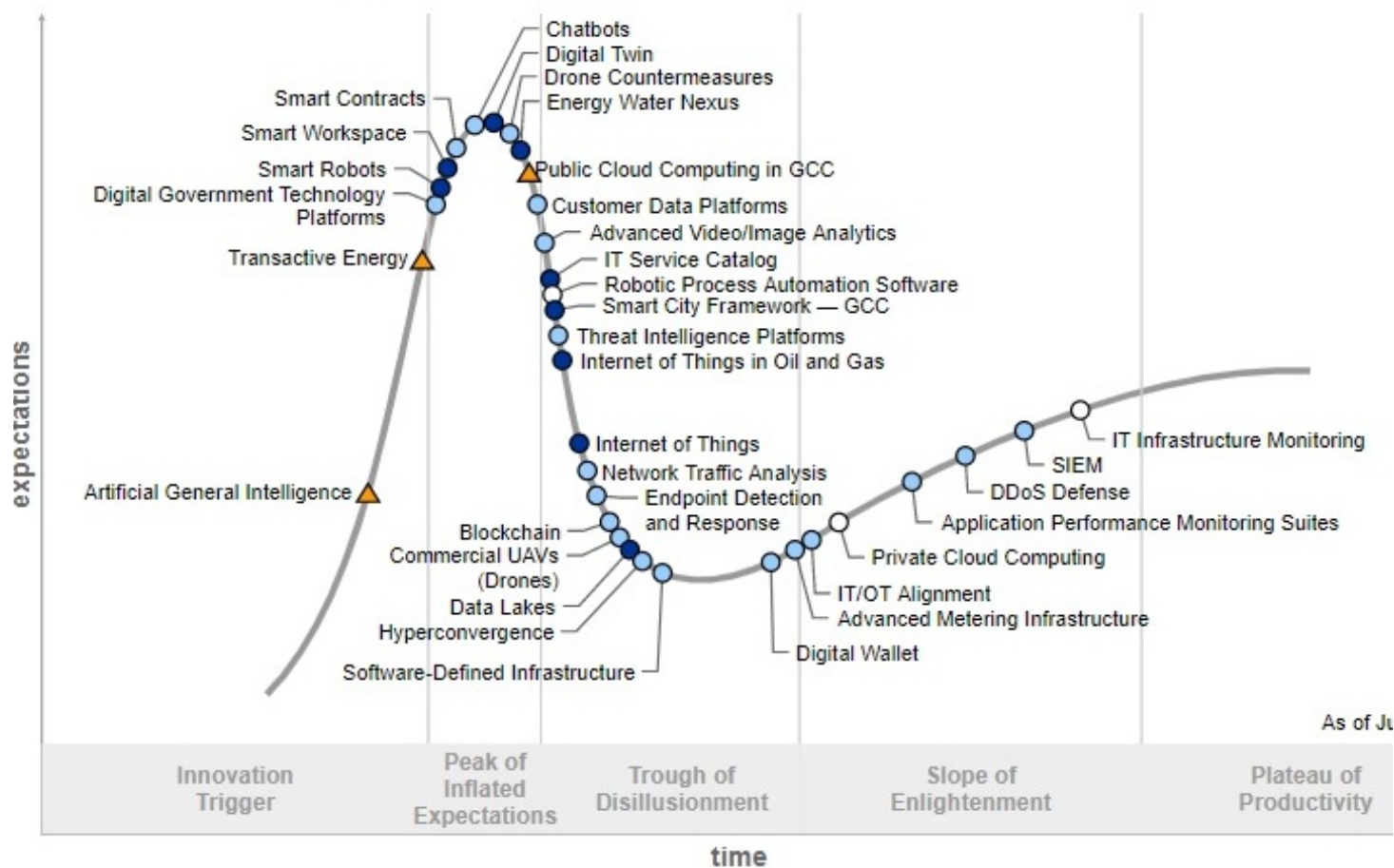
Tyrimas

Tyrimo uždaviniai:

- Įvertinti viešo paskirstyto skaičiavimo modelio taikymo galimybes didelių duomenų tyrybai;
- Įvertinti esamas viešo paskirstyto skaičiavimo 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 viešo paskirstyto skaičiavimo metodu grįstą platformos teorinį modelį, atitinkantį įmonių bei organizacijų keliamus poreikius ir reikalavimus.

Planuojami rezultatai:

- Viešo paskirstyto skaičiavimo platforma didelių duomenų tyrybai;
- Platformos taikymo praktiniams uždaviniams spręsti rezultatų apibendrinimas.



Tyrimų sritis ir problemos aktualumas

Viešo paskirstyto skaičiavimo platformos suteikia papildomus skaičiavimų resursus. Pagrindinis tokių platformų privalumas lyginant su tradiciniais sprendimais yra maži infrastruktūros diegimo bei išlaikymo kaštai.

Nors ir yra sprendimų, leidžiančių lengvai sujungti vidinius IT išteklius į paskirstytą skaičiavimų platformą, organizacijos linkusios rinktis kitas alternatyvas dėl kylančių **saugumo** bei **patikimumo** problemų.

Žinomi viešo paskirstyto skaičiavimo modelis bei taikomi užduočių paskirstymo algoritmai neišsprendžia skaičiavimams naudojamų išteklių heterogeniškumo problemos.

Plateau will be reached:

- less than 2 years
- 2 to 5 years
- 5 to 10 years
- ▲ more than 10 years
- ⊗ obsolete before plateau

© 2019 Gartner

Sistemiškai išnagrinėtas užduočių paskirstymo algoritmų taikymas paskirstyto skaičiavimo tinkluose;

Sukurta nauja hibridinė paskirstyto skaičiavimo platformos architektūra;

Pritaikytas bei integruotas užduočių sulaikymo buferis, naudojamas skirstant užduotis tarp dviejų skirtingo našumo klasterių.

Darbo
naujumas ir
jo
aktualumas

Darbo rezultatų praktinė reikšmė

- Pritaikytas užduočių sulaikymo buferis, sprendžiant užduočių paskirstymo problemą tarp skirtingo našumo įrenginių;
- Patobulintas FIFO algoritmas integruotas į hibridinį paskirstyto skaičiavimo architektūrą. Tai leidžia sumažinti bendrą užduočių vykdymo trukmę. Remiantis šiuo metodu sukurta platforma, apjungianti viešo bei privataus skaičiavimo klasterius į hibridinį skaičiavimo tinklą.

Ginamieji teiginiai

1. Įmonių bei organizacijų darbuotojų kompiuterių ištekliai gali būti naudojami didžiųjų duomenų tyrybai netrikdant darbuotojų darbo;
2. Esamos viešo paskirstyto skaičiavimo platformos neatitinka įmonių bei organizacijų keliamų reikalavimų didžiųjų duomenų tyrybos užduotims vykdyti naudojant įmonės kompiuterius;
3. Viešo paskirstyto skaičiavimo platformų skaičiavimų išteklių heterogeniškumo problema yra išsprendžiama taikant hibridinio paskirstyto skaičiavimo architektūrą;
4. Užduočių sulaikymo buferis gali būti pritaikytas hibridinėse paskirstyto skaičiavimo platformose, kadangi tai leidžia sutrumpinti užduočių atlikimo laiką, palyginus su FIFO algoritmu.

Darbo rezultatų aprobavimas

Disertacijos tyrimų rezultatai publikuoti 2 moksliniuose straipsniuose, vienas straipsnis yra įteiktas:

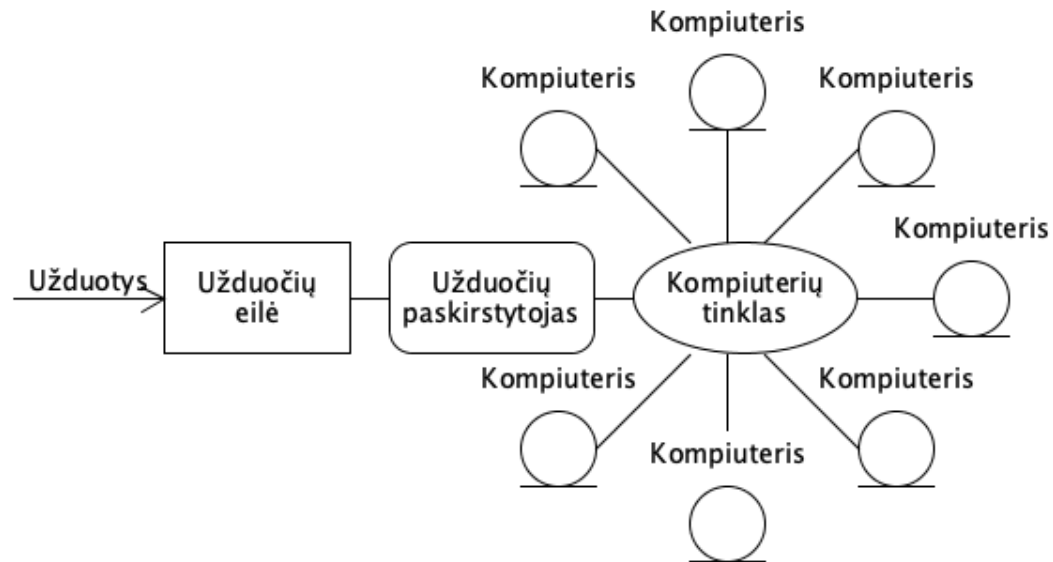
- „BOINC from the view point of Cloud computing“. Straipsnis publikuotas „CEUR Workshop Proceedings“ (referuojamas CEUR-WS.org, ISSN 1613-0073, <http://ceur-ws.org/Vol-1973/paper08.pdf>);
- „Big Data mining using public distributed computing“. Straipsnis publikuotas tarptautiniame leidinyje „Information Technology and Control“ (DOI: <http://dx.doi.org/10.5755/j01.itc.47.2.19738>).
- „Application of Task Stalling Buffer in Distributed Hybrid Cloud Computing“. Straipsnis įteiktas tarptautiniam žurnalui „Simulation Modelling Practice and Theory“ (Special Issue on „Modeling and Simulation of Hybrid Clouds“).

Darbo rezultatų aprobavimas

Rezultatai pristatyti 6-iose respublikinėse ir tarptautinėse mokslinėse konferencijose:

- 2017-04-03 respublikinėje konferencijoje „Informacinių technologijų iššūkiai kūrybos ekonomikoje“ (Lietuva). Pranešimo tema: „BOINC karkaso taikymas didelių duomenų tyrybai“;
- **2017-08-30 tarptautinėje konferencijoje „BOINC: Fundamental and Applied Science and Technology (BOINC:FAST 2017)“ (Petrozavodskas, Rusija). Pranešimo tema: „BOINC from the view point of Cloud computing“;**
- 2017-09-22 XVIII tarptautinėje mokslinėje kompiuterininkų konferencijoje („Kompiuterininkų dienos – 2017“, Kaunas). Pranešimo tema: „Big Data mining using public distributed computing“.
- 2017-11-30 tarptautinėje konferencijoje „Duomenų analizės metodai programų sistemoms“ (Druskininkai). Stendinio pranešimo tema: „BOINC Based Enterprise Desktop GRID“;
- **2020-04-30 septintojoje tarptautinėje konferencijoje „Open International Conference on Electrical, Electronic and Information Sciences“ (Lietuva). Skaityto pranešimo tema: „Distributed Hybrid Cloud Controller for Runbook Operations“;**
- **2020-06-05 ketvirtojoje tarptautinėje konferencijoje „4th International Conference on Innovations and Creativity“ (Latvija). Skaityto pranešimo tema: „Task Stalling Buffer Application in Grid Computing“.**

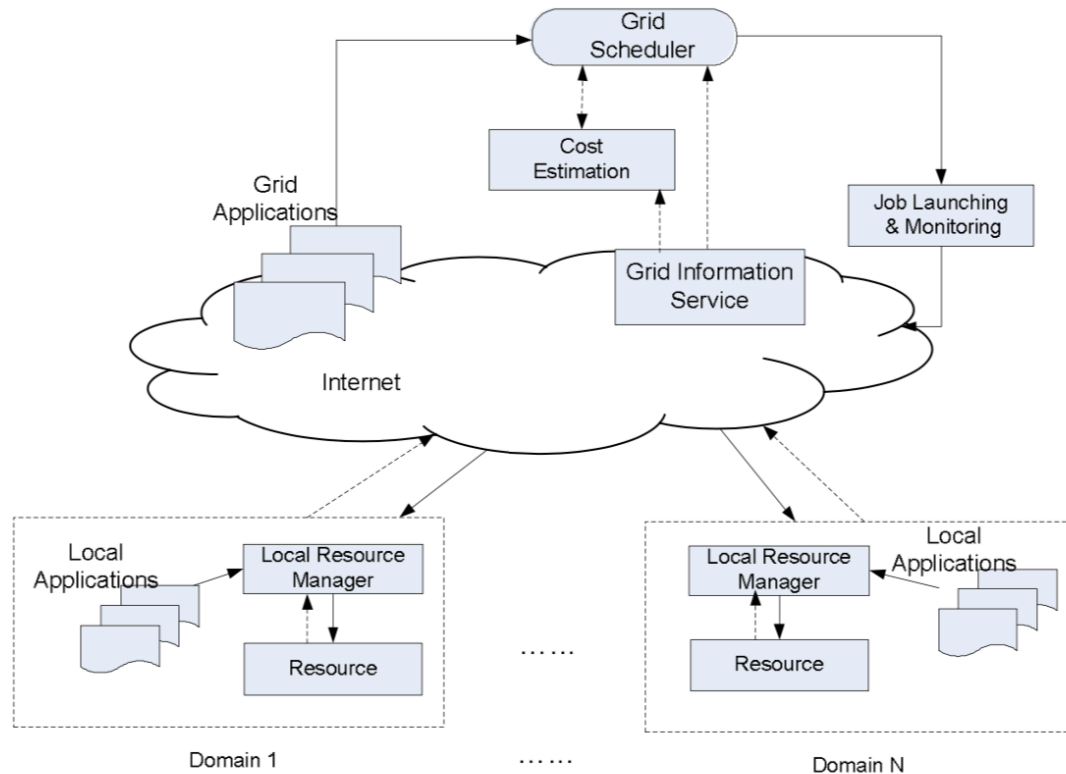
Viešas paskirstytas skaičiavimas



Viešas paskirstytas skaičiavimas yra skaičiavimo metodas, kai lygiagrečiai naudojami savanorių kompiuteriai skaičiavimams atlikti. Komunikacija atliekama tinklu, naudojant kliento-serverio architektūrą, kur klientai pateikia savo turimus išteklius naujoms užduotims iš serverio gauti.

Viešo paskirstyto skaičiavimo platforma

Viešo paskirstyto skaičiavimo platforma (angl. „public distributed computing platform“) – atviraime tinkle esančių kompiuterių sistema, skirta užduočių vykdymui panaudojant tinkle esančių kompiuterių išteklius.



Viešo paskirstyto skaičiavimo platformų problemos

Heterogeniškumo problemos:

- kompiuterių ištekliai gali bet kuriuo metu tapti neprieinami;
- gali skirtis kompiuterinė architektūra bei našumas.

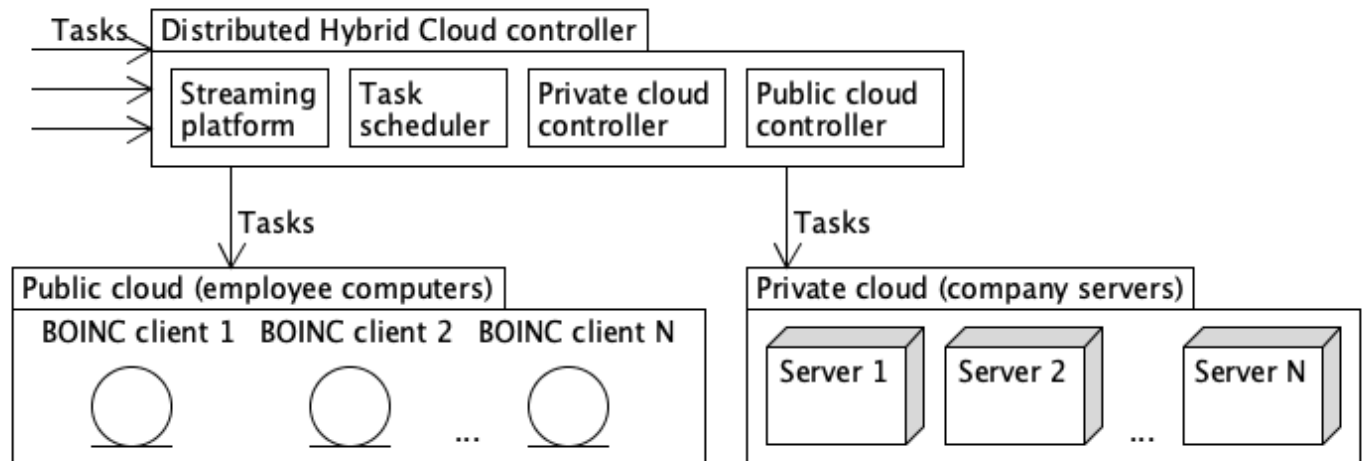
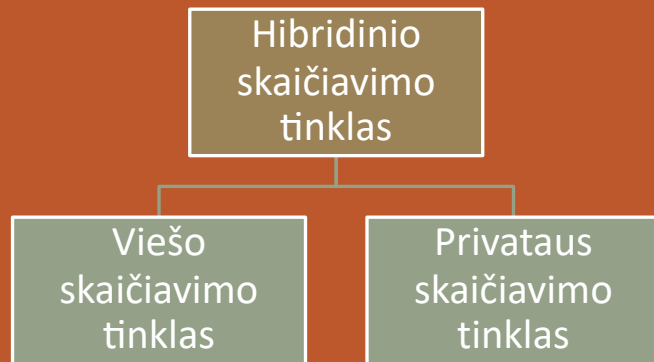
Užduočių įstrigimo problema:

- užduotys mažai išteklių turinčiuose kompiuteriuose vykdomos žymiai ilgiau;
- kompiuteriai arba jų ištekliai gali tapti neprieinami.

Saugumo problemos:

- turi būti užtikrintas privačių duomenų konfidencialumas.

Hibridinio skaičiavimo tinklas



Skyriaus išvados

1. BOINC platforma yra plačiai naudojama didelių duomenų tyrybai naudojant viešuosius paskirstytus skaičiavimus;
2. Nėra sukurta viešo paskirstyto skaičiavimo platforma didelių duomenų tyrybai, atitinkanti įmonių bei organizacijų keliamus reikalavimus;
3. Viešo paskirstyto skaičiavimo platformų skaičiavimų išteklių heterogeniškumo problema yra išsprendžiama taikant hibridinio paskirstyto skaičiavimo architektūrą;

Egzistuojančios platformos

Viešųjų paskirstytų skaičiavimų platformos didelių duomenų tyrybai: BOINC, BOINC-MR, V-BOINC, GridBot ir kt.

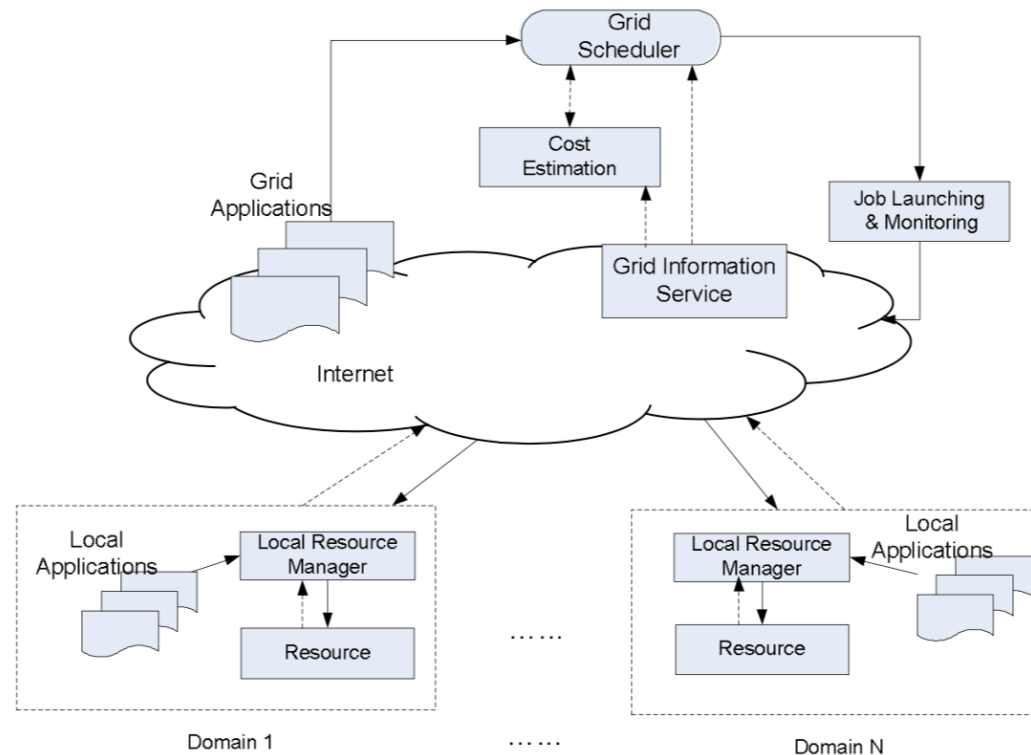
Platformose užduotys skirstomos:

1. vertinant užduočių vykdymo laikus;
2. sudarant užduočių vykdymo tvarkaraštį;
3. taikant užduočių replikaciją;

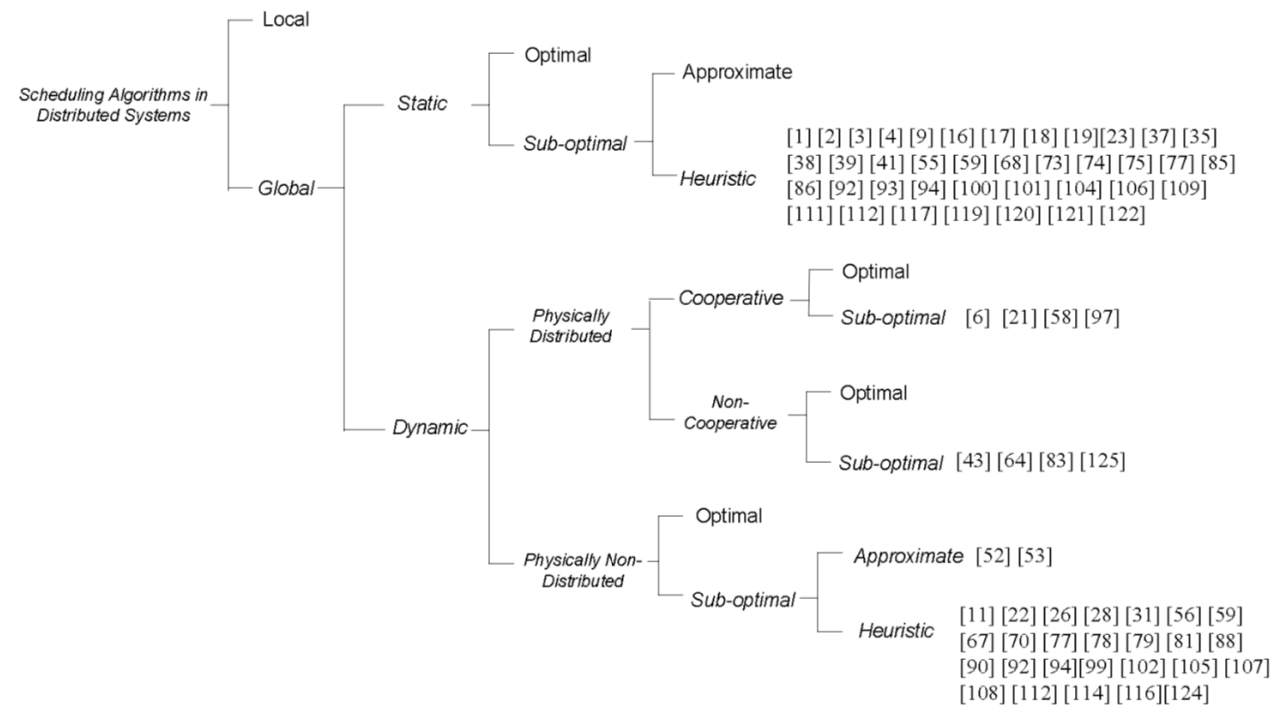
Trūkumai:

1. užduočių sąrašas bei užduočių vykdymo laikai ne visada yra iš anksto žinomi;
2. nesprenžiamos saugumo problemos;
3. pritaikytos konkrečioms uždaviniais spręsti.

Užduočių paskirstymo modeliai



Užduočių paskirstymo metodai



Dong, Fangpeng. (2006). Scheduling Algorithms for Grid Computing: State of the Art and Open Problems.

Užduočių paskirstymo algoritmai

1. Fair-share – kiekviena užduočių grupė gauna tam tikrą skaičiavimų resursų dalį;
2. FIFO – užduotys aptarnaujamos iš eilės pagal pirmumą;
3. Capacity – sudaromos užduočių eilės, kurioms aptarnauti priskiriamos skaičiavimų resursų dalys;
4. LATE – naudoja užduočių replikaciją;
5. Round-robin (RR) – kiekvienai užduočių grupei priskiriamas atskiras skaičiavimų resursas.

Heterogeniškumo problema: joks užduočių paskirstymo algoritmas lyginant su kitais užduočių paskirstymo algoritmais negali visais atvejais užduotis paskirstyti taip, jog bendras visų užduočių vykdymo laikas būtų trumpiausias. Iš to seka išvada, jog užduočių paskirstymo algoritmas turi gebėti prisitaikyti prie vykdomų užduočių bei skaičiavimų tinklo pajėgumų. [1]

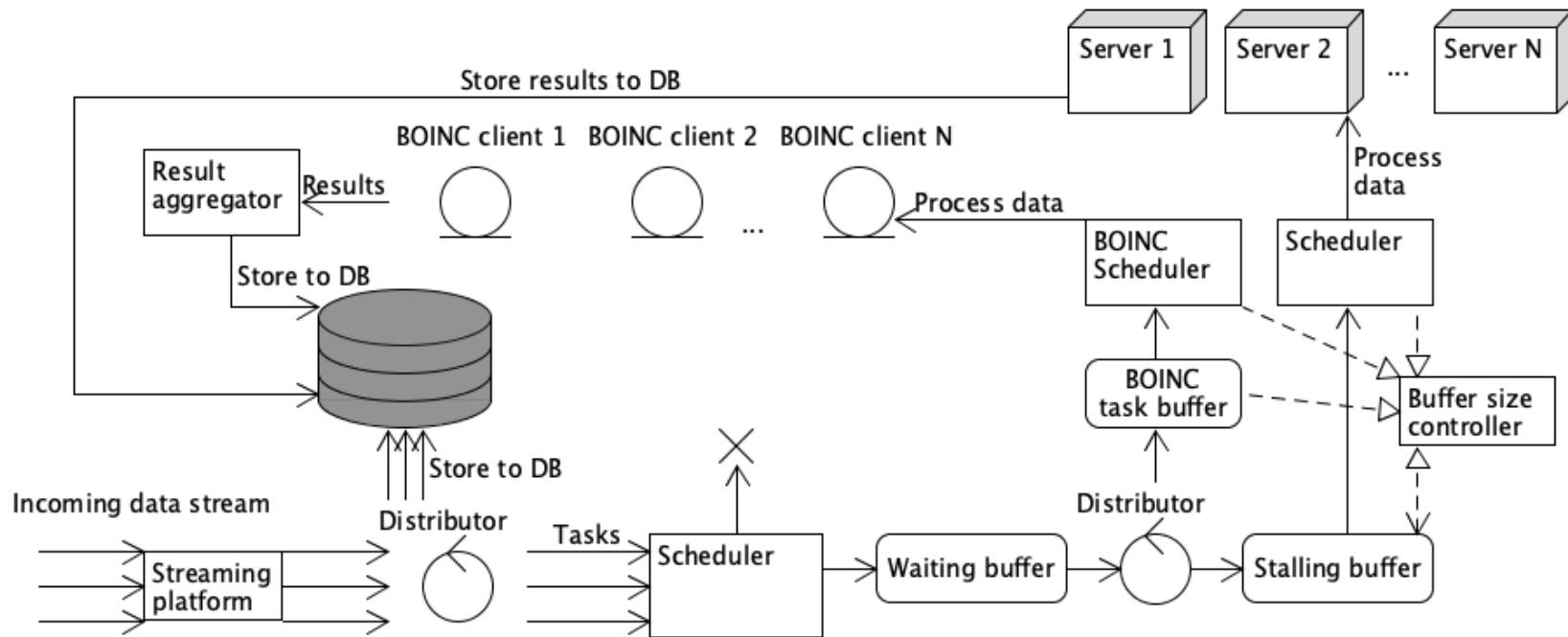
[1] M. Maheswaran, S. Ali, H. J. Siegel, D. Hensgen and R. F. Freund, Dynamic Matching and Scheduling of a Class of Independent Tasks onto Heterogeneous Computing Systems, in J. of Parallel and Distributed Computing, Vol. 59, No. 2, pp.107--131, November 1999.

Skyriaus išvados

1. Viešo paskirstyto skaičiavimo platformos kūrimas yra susijęs su šiomis problemomis:
 1. skaičiavimuose dalyvaujantys kompiuteriai turi skirtingas architektūras;
 2. skaičiavimuose dalyvaujantys kompiuteriai yra skirtingų našumų;
 3. skaičiavimuose dalyvaujantys kompiuteriai bet kada gali tapti nepasiekiami;
 4. užduočių sąrašas bei užduočių vykdymo laikai ne visada yra iš anksto yra žinomi.

Hibridinės paskirstyto skaičiavimo sistemas architektūra

Užduočių paskirstymo teorinis modelis



Apibrėžimai

Berkeley Open Infrastructure for Network Computing (BOINC) – viešo paskirstyto skaičiavimo platforma

Apache Mesos – paskirstyti skaičiavimo bei klasterių administravimo platforma

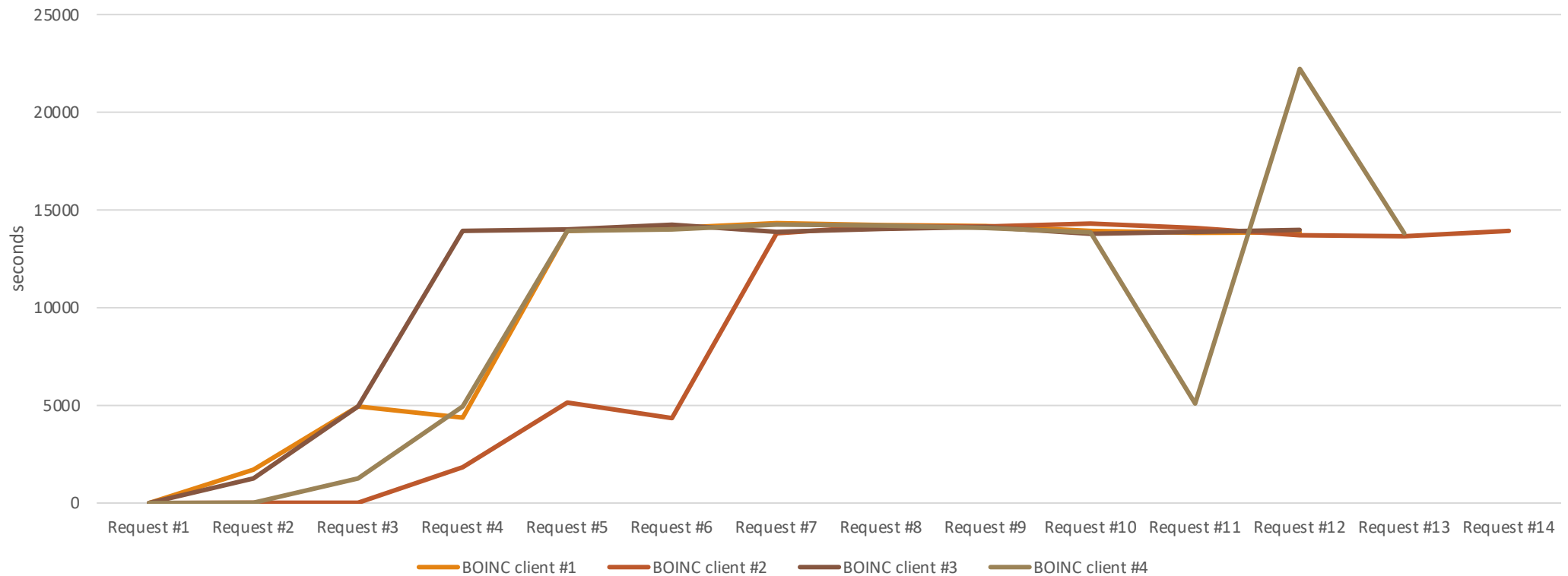
Apache Chronos – užduočių paskirstymo sistema, skirta MESOS platformai

Docker – virtualizacijos platforma

Apache Kafka – srauto paskirstymo platforma

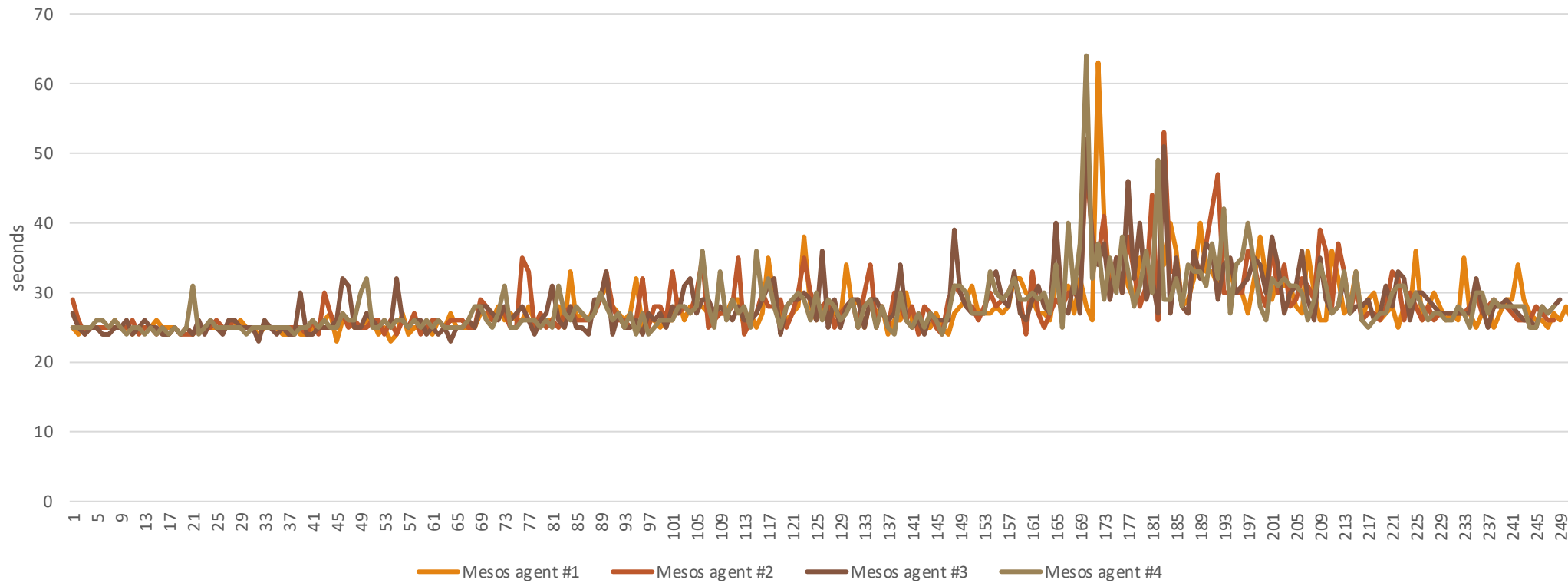
BOINC klientų darbo užklauso

Ekspperimentų metu gauta viešo paskirstyto skaičiavimo tinklo klientų darbo užklauso statistika.



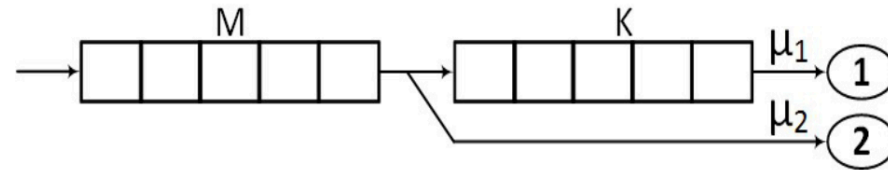
MESOS klientų darbo užklauso

Ekspertymetu gauta privataus paskirstyto skaičiavimo tinklo klientų darbo užklauso statistika.



Užduočių sulaikymo buferis

Užduočių sulaikymo buferis (angl. „task stalling buffer“) sprendžia užduočių įstrigimo problemą [1].



- M – laukimo buferio ilgis
- K – užduočių sulaikymo buferio ilgis
- μ_1 – greitojo kanalo našumas
- μ_2 – lėtojo kanalo našumas
- 1 – greitis kanalas
- 2 – lėtasis kanalas

[1] L. Kaklauskas, L. Sakalauskas, V. Denisovas. Stalling for solving slow server problem. RAIRO-Oper. Res., 53 4 (2019) 1097-1107

Užduočių sulaikymo buferis

Litlo formulė tvirtina, kad vidutinis paraiškų skaičius sistemoje yra lygus įėjimo srauto intensyvumo ir vidutinės buvimo sistemoje trukmės sandaugai. Laukimo eilės atveju Litlo formulė išreiškiama taip:

$$\bar{Q} = \lambda \cdot \bar{W}$$

t. y., vidutinis laukimo eilės ilgis yra lygus įėjimo srauto intensyvumo ir laukimo trukmės sandaugai. Remiantis [1], užduočių sulaikymo buferio ilgį K galima išreikšti taip:

$$K \approx r \cdot (1 - q)$$

r – greitojo ir lėtojo kanalų našumo santykis

q – užduočių vykdymo našumo koeficientas

$$q = \frac{c}{t \cdot m \cdot \mu_1}$$

c – įvykdytų užduočių kiekis

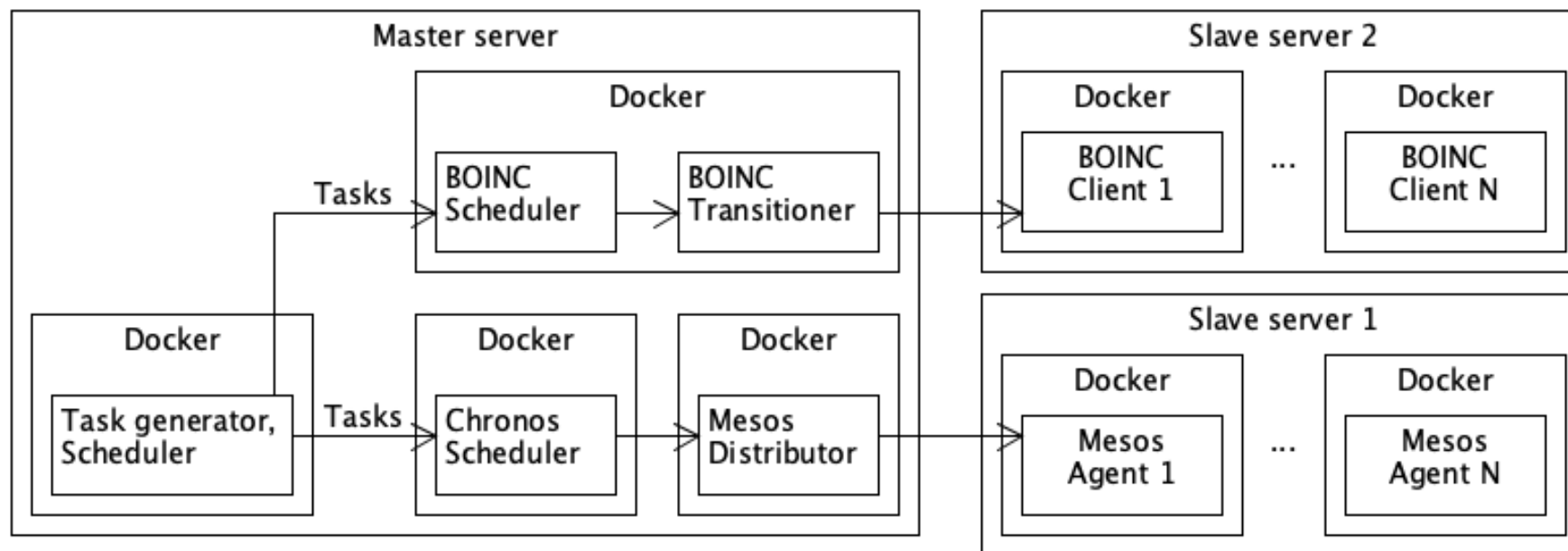
t – įvykdytų užduočių suminis laikas

m – greitą kanalą aptarnaujančių kompiuterių kiekis

μ_1 – greitojo kanalo našumas

[1] L. Kaklauskas, L. Sakalauskas, V. Denisovas. Stalling for solving slow server problem. RAIRO-Oper. Res., 53 4 (2019) 1097-1107

Siūlomos platformos architektūra



Skyriaus išvados

1. Sukurta hibridinė paskirstyto skaičiavimo architektūra;
2. Užduočių sulaikymo buferis gali būti taikomas hibridinėse paskirstyto skaičiavimo sistemose.

Kompiuteriniai
eksperimentai
ir kompiuterinis
modeliavimas

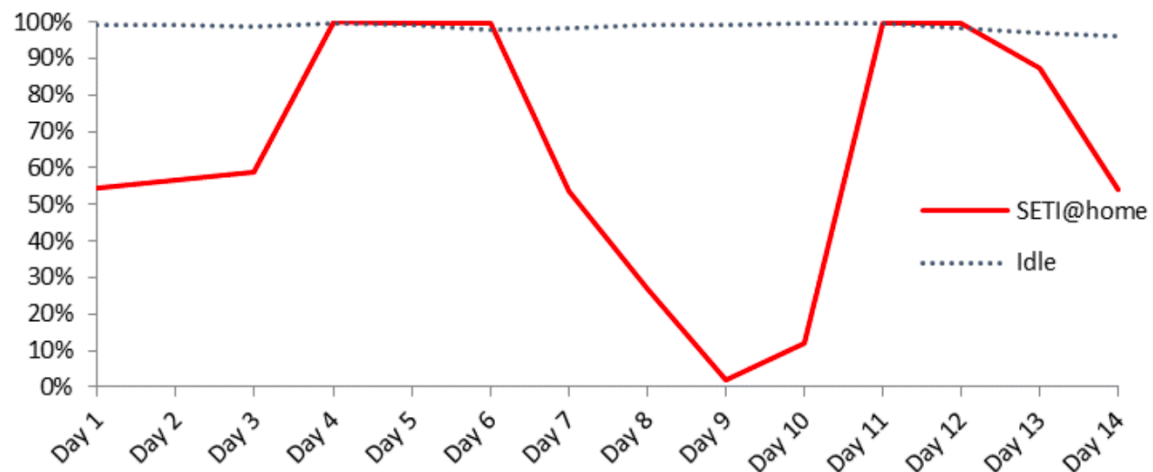
Darbo kompiuterių apkrovos ir energijos vartojimo tyrimas

Atliktas 28 dienas trukęs tyrimas naudojant du atsitiktinai parinktus kompiuterius dviejose įstaigose (A ir B) siekiant ištirti:

- kompiuterių procesorių prastovą;
- kompiuterių sunaudojamos elektros energijos kiekį.

Kompiuteriuose buvo įdiegta BOINC programa, dvi savaites vykdanti SETI@home projekto užduotis. Skaičiavimų metu buvo matuojami elektros energijos vartojimas, bei procesoriaus užimtumo laikas naudojant “Performance monitor” (programa prieinama Microsoft Windows sistemoje) bei elektroninį elektros energijos vartojimo matuoklį. Po to procesas buvo pakartotas nevykdant BOINC projektų. Tyrimo metu abu kompiuteriai buvo naudojami įstaigų darbuotojų darbo reikmėms.

Eksperimentinio tyrimo rezultatai



Organizacija	BOINC projektas	CPU prastova	Elektros sunaudojimas
A	-	98,77%	16,61 kWh
A	SETI@home	65,23%	22,03 kWh
B	-	83,49%	1 kWh
B	SETI@home	26,86%	2,09 kWh

Platformoje vykdomos užduotys

Užduotys pateikiamos JSON duomenų formatu:

```
{"container": "<užduotis>", "cluster": "<klasteris>"}
```

užduotis – Docker konteinerio pavadinimas bei parametrai (sprendimas paremtas [1])

klasteris – užduočių vykdymo potinklio arba užduočių paskirstymo metodo pavadinimas. Skaičiavimų nukreipimas į norimą klasterį leidžia apsaugoti konfidencialius duomenis.

Pavyzdžiai:

1. {"container": "ashael/pi 100000", "cluster": "first-available"}
2. {"container": "ashael/pi 100000", "cluster": "stalling-buffer"}
3. {"container": "ashael/pi 50000", "cluster": "boinc"}
4. {"container": "ashael/pi 200000", "cluster": "mesos"}
5. {"container": "-e \"INPUT_FILENAMES=2015-01-01-0.json.gz;2015-01-01-1.json.gz;2015-01-01-2.json.gz\" -v /opt/mii@home/tasks/docker-map-reduce-example/data/:/data -v /opt/mii@home/tasks/docker-map-reduce-example/out/:/out mrquad/analysis_test", "cluster": "mesos"}

[1] Schlitter, Nico & Laessig, Joerg. (2013). Distributed Data Analytics using RapidMiner and BOINC.

Eksperimentinis tyrimas

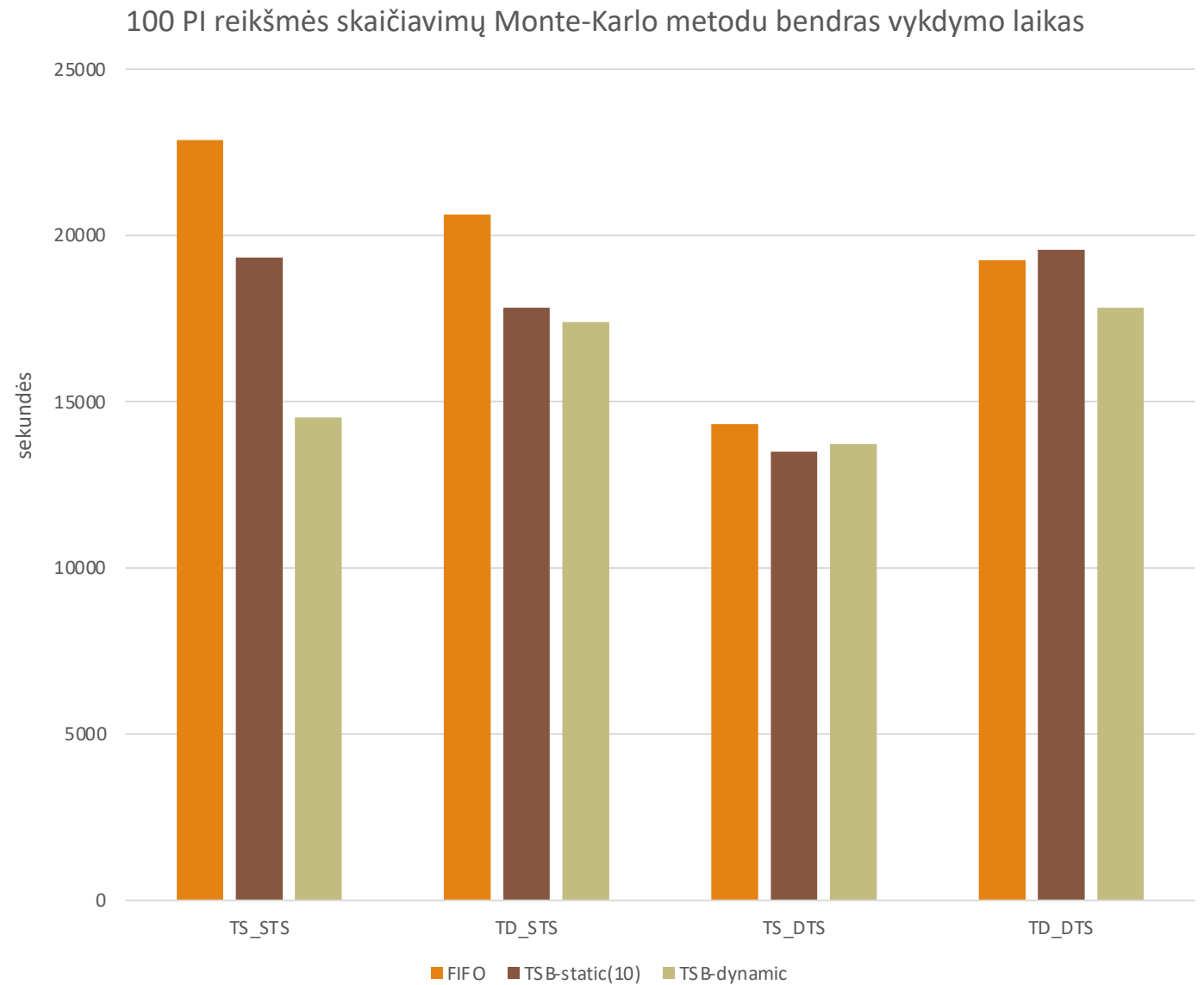
Tyrimo tikslas naudojant sukurtą platformą palyginti bendrą visų užduočių vykdymo trukmę užduočių paskirstymui taikant:

1. FIFO algoritmą;
2. statinio ilgio užduočių sulaikymo buferį (TSB-static);
3. dinaminio ilgio užduočių sulaikymo buferį (TSB-dynamic).

Užduočių vykdymo trukmei įvertinti sprendžiamas PI reikšmės skaičiavimo Monte-Karlo metodu uždavinys, išskaidytas į 100 užduočių. Kiekvienas algoritmas tiriamas taikant šiuos parametrus (viso 12 atvejų):

1. vienodo dydžio užduotys, statinis užduočių patekimo į sistemą srautas (TS_STS);
2. vienodo dydžio užduotys, dinaminis užduočių patekimo į sistemą srautas (TS_DTS);
3. kintamo dydžio užduotys, statinis užduočių patekimo į sistemą srautas (TD_STS);
4. kintamo dydžio užduotys, dinaminis užduočių patekimo į sistemą srautas (TD_DTS).

Eksperimentinio tyrimo rezultatai



Eksperimentinio tyrimo rezultatai

Taikant dinaminio dydžio buferį užduočių sulaikymui yra sutrumpinamas suminis visų užduočių atlikimo laikas:

1. Taikant statinio ilgio užduočių sulaikymo buferį bendras užduočių vykdymo laikas yra trumpesnis 75% atvejų. Gaunamas vidutinis **8,875%** pagreitinėjimas.
2. Taikant dinaminio ilgio užduočių sulaikymo buferį bendras užduočių vykdymo laikas visais atvejais yra trumpesnis. Gaunamas vidutinis **17,63%** pagreitinėjimas.

Vidutinis užduočių eilės įvykdymo laikas naudojant FIFO algoritmą.

Scenarijus	20 užduočių		40 užduočių		60 užduočių	
	Vidutinis vykdymo laikas (sek.)	Standartinis nuokrypis (sek.)	Vidutinis vykdymo laikas (sek.)	Standartinis nuokrypis (sek.)	Vidutinis vykdymo laikas (sek.)	Standartinis nuokrypis (sek.)
TS_STS	3598,4	189,2	7063,6	212,86	9995,4	223,01
TD_STS	5841,8	654,4	9622,2	1229,63	14905,4	1382,52
TS_DTS	3474	104,82	6576	145,25	9754	340,32
TD_DTS	5933,4	455,89	10038,4	1439,31	14620,4	1255,28

Vidutinis užduočių eilės įvykdymo laikas naudojant TSB-static(10) algoritimą.

Scenarijus	Užduočių kiekis	Vidutinis vykdymo laikas (sek.)	Standartinis nuokrypis (sek.)	Pagerėjimas lyginant su FIFO	p-reikšmė lyginant su FIFO
TS_STS	20	3503,8	394,22	2,63%	0,646
	40	6377,8	32,8	9,71%	0,002
	60	8978,6	297,52	10,17%	0,0005
TD_STS	20	3967,4	959,76	32,09%	0,009
	40	9068,4	172,92	5,76%	0,375
	60	13381,8	393,63	10,22%	0,064
TS_DTS	20	3241,6	203	6,69%	0,063
	40	6168,8	256,39	6,19%	0,021
	60	9254	240,38	5,13%	0,031
TD_DTS	20	3127	824,78	47,3%	0,001
	40	8690	301,94	13,43%	0,11
	60	12293	1566,73	15,92%	0,032

Vidutinis užduočių eilės įvykdymo laikas naudojant TSB-dynamic algoritmą.

Scenarijus	Užduočių kiekis	Vidutinis vykdymo laikas (sek.)	Standartinis nuokrypis (sek.)	Pagerėjimas lyginant su FIFO	p-reikšmė lyginant su FIFO
TS_STS	20	3493	133,24	2,93%	0,342
	40	6191,8	218,76	12,34%	0,0002
	60	9147,8	264,42	8,48%	0,001
TD_STS	20	5016,4	117,66	14,13%	0,05
	40	8759,8	257,83	8,96%	0,2
	60	13256,2	480,22	11,06%	0,053
TS_DTS	20	3497,4	133,06	-0,67%	0,765
	40	6342,2	88,81	3,56%	0,018
	60	9204	118,35	5,64%	0,019
TD_DTS	20	4697	607,02	20,84%	0,008
	40	9018,8	193,9	10,16%	0,192
	60	13262,2	461,7	9,29%	0,072

Skyriaus išvados

1. Pritaikius užduočių sulaikymo buferį, užduočių eilė gali būti įvykdoma greičiau nei taikant FIFO algoritmą;
2. Įmonių bei organizacijų savanorių kompiuterių ištekliai gali būti naudojami didžiųjų duomenų tyrybai netrikdant darbuotojų darbo.

Darbo išvados

1. Sukurta hibridinė paskirstyto skaičiavimo platforma didelių duomenų tyrybai atitinka įmonių bei organizacijų keliamus reikalavimus tokio pobūdžio platformoms;
2. Integruotas užduočių paskirstymo algoritmas, naudojantis užduočių sulaikymo buferį, užduočių eilę leidžia įvykdyti greičiau nei taikant FIFO algoritmą;

Praktinės sritys

Remiantis tuo jog ši platforma yra hibridinis skaičiavimų tinklas, galėtų būti skaičiavimų sritys, pasižyminčios tokiomis savybėmis:

- iš anksto nėra žinoma būsima skaičiavimų sistemos apkrova;
- galimi skaičiavimų sistemos apkrovos šuoliai;
- reikalingas aukštas skaičiavimų sistemos prieinamumas;
- yra duomenų apsaugos reikalavimai, ribojantys duomenų viešinimą.



Ačiū už dėmesį