

VILNIAUS UNIVERSITETAS
MATEMATIKOS IR INFORMATIKOS FAKULTETAS
PROGRAMŲ SISTEMŲ KATEDRA

Blokų grandinių duomenų bazės finansinių duomenų kaupimui

Blockchain Databases for Financial Data

Kursinis darbas

Atliko: 4 kurso 3 grupės studentas
Matas Savickis (parašas)

Darbo vadovas: dr. Vytautas Valaitis (parašas)

Vilnius – 2019

TURINYS

PADEKA	2
IVADAS	1
TIKSLAS	1
UŽDAVINIAI	1
1. DUOMENŲ BAZIŲ GREIČIŲ TYRIMŲ LITERATŪROS ANALIZĖ	3
1.1. Metrikos	3
1.2. Hyperledger Fabric	3
1.2.1. Architektūra	3
1.2.2. Tranzakcijų valdymas	3
1.2.3. Parametrai	4
1.2.4. Testavimo metodologija	5
1.2.5. Tyrimų rezultatai	5
1.2.6. IBM tyrimo rezultatai	6
1.2.7. Tailando „Kompiuterių technologijos centro“ tyrimo rezultatai	6
1.2.8. Sharjah universiteto tyrimo rezultatai	6
1.2.9. Tyrimų rezultatų apibendrinimas	7
1.3. Apache Cassandra	7
1.3.1. Architektūra	7
1.3.2. Testavimo metodologija	9
1.3.3. Blekinge technologijų instituto tyrimo rezultatai	10
1.3.4. Coimbra politechnikos instituto tyrimo rezultatai	11
1.4. Greičių tyrimų apibendrinimas	11
2. DUOMENŲ BAZIŲ PALYGINIMAS PAGAL CAP TEOREMĄ	13
2.1. CAP teorema	13
2.2. Hyperledger Fabric pagal CAP teoremą	13
2.3. Hyperledger Fabric neprieštarinumas	14
2.4. Hyperledger Fabric pasiekiamumas	14
2.5. Hyperledger Fabric skaidinių toleravimas	14
2.6. Cassandra neprieštarinumas	14
2.7. Cassandra pasiekiamumas	14
2.8. Cassandra skaidinių toleravimas	15
2.9. Apibendrinimas	15
3. REZULTATAI IR IŠVADOS	16
3.1. Rezultatai	16
3.2. Išvados	16
3.3. Rekomendacijos ateities darbams	16
4. PRIEDAI	17
4.1. Žodynas	17
LITERATŪRA	17

Padėka

Noriu padėkoti savo darbo vadovui daktarui Vytautui Valaičiui už pagalbą, vadovavimą ir skirtą savo laisvą vasaros laiką padedant man rašyti šį darbą.

Įvadas

Per pastaruosius keletą metų blokų grandinių technologija susilaukė didelio žmonių susidomėjimo. Šis susidomėjimas daugiausiai kilo dėl išpopulerėjusių kriptovaliutų, tokių kaip Bitcoin, Ethereum, Litecoin ir daugeliu kitų kurios ir yra paremtos blokų grandinių technologija. Šią technologiją 2008 metais sukūrė Satošis Nakamoto [Nak08]. 2009 metais Nakamoto realizavo blokų grandinių technologiją sukurdamas Bitcoin kriptovaliutą [Nak09]. Nors, šiuo metu, žmonių susidomėjimas kripto valiutomis ir yra sumažėjęs [Goo19], tačiau informacinių technologijų industrija mato daugiau blokų grandinių panaudojimo atveju negu tik kriptovaliutos. Vienas iš blokų grandinių panaudojimo atvejų yra blokų grandinių duomenų bazės. Reliacinės ir dokumentų duomenų bazės ilgą laiką buvo pagrindinis duomenų saugojimo būdas. Tačiau šios duomenų bazės turi ir savo trūkumų, saugant duomenis reliacinėse duomenų bazėse kyla duomenų integralumo problemos [FSW81]. Naudojant duomenų bazes finansinėms transakcijoms sekti kyla dvigubo pinigų išleidimo problema[Hoe08]. Naudojantis tradicinėmis duomenų bazėmis taip pat kyla pasitikėjimo problema, visa duomenų prieiga yra trečiosios šalies valdžioje, ir vartotojas turi pasitikėti, kad duomenys nebus pakeisti be jo žinios. Per pastaruosius kelis metus šias problemas buvo stengtasi išspręsti kuriant duomenų bazes paremtas blokų grandinių technologija. Privačios blokų grandinių duomenų bazės užtikrina pasitikėjimą, nes kiekvienas vartotojas turi visą duomenų bazės kopiją. Darant pakeitimus tokioje duomenų bazėje kiekvienas vartotojas turi sutikti su daromais pakeitimais ir saugo visų pakeitimų istoriją. Blokų grandinių duomenų bazės išsprendžia duomenų integralumo ir dvigubo pinigų išleidimo problemą, nes kiekvienas mazgas blokų grandinės tinkle gali palyginti savo turimus duomenis su kitais mazgais. Nors blokų grandinės išsprendžia išvardytas problemas, tačiau finansiniams duomenims taip pat svarbu greitis, duomenų pasiekiamumas ir sistemos galimybė išlikti funkcionaliai po tam tikrų trikdžių. Yra nemaža aibė blokinių grandinių duomenų bazių(R3 Corda, BigChain DB), bei didelis pasirinkimas NoSQL duomenų bazių(Mongo DB, Redis, Neo4j) tačiau šiam darbui buvo pasirinktos Hyperledger Fabric ir Apache Cassandra, nes šios duomenų bazės yra naudojamos finansiniams duomenims saugoti [AF16; Dat17].

Tikslas

CAP teorema - teorema sakanti, kad paskirstytoje duomenų sistemoje neįmanoma užtikrinti daugiau dviejų punktų: Duomenų neprieštaringumas, duomenų pasiekiamumas, dalivijusi toleravimas.

Šio darbo tikslas yra:

- Palyginti Apache Cassandra ir Hyperledger Fabric pagal išskirtas metrikas.
- Palyginti Apache Cassandra ir Hyperledger Fabric pagal CAP teoremą.

Uždaviniai

1. Išskirti bendras metrikas, pagal kurias būtų galima palyginti Apache Cassandra ir Hyperle-

dger Fabric duomenų bazes.

2. Palyginti Apache Cassandra ir Hyperledger Fabric pagal išskirtas metrikas.
3. Išskirti esamų tyrimų trūkumus ir galimas sritis ateities darbams.
4. Palyginti Apache Cassandra su Hyperledger Fabric pagal CAP teoremos punktus.

1. Duomenų bazių greičių tyrimų literatūros analizė

1.1. Metrikos

Šiame skyriuje bus apžvelgiami tyrimai, kuriuose buvo analizuotas ir vertintas Hyperledger Fabric [NQT⁺18; PST17; TNV18] ir Apache Cassandra [AB13; SLC19] veikimas. Šiuose darbuose buvo vertinamas: transakcijų praeinamumas, transakcijų vėlavimas, procesoriaus resursų sunaudojimas, mazgų skaičius paskirstytoje duomenų bazėje.

Šiame darbe Hyperledger Fabric ir Apache Cassandra bus lyginama pagal:

- Transakcijų praeinamumas - kiek transakcijų duomenų bazė gali apdoroti per sekundę.
- Vėlavimas - kiek laiko užtrunka nuo transakcijos išsiuntimo iki jos patvirtinimo duomenų bazėje.

1.2. Hyperledger Fabric

Hyperledger yra atviro kodo blokų grandinės pradėtos kurti Linux fondo. Šio projekto tikslas yra gerinti tarp industrinį bendradarbiavimą kuriant blokų grandines kurios užtikrintų patikimą, greitą ir saugų finansinių duomenų perdavimą pagrindinėse technologijų, finansų ir produktų tiekimo kompanijose [Hyp16]. Šiame skirsnyje bus apžvelgti Hyperledger Fabric, populiariausios Hyperledger realizacijos, architektūra, greičio tyrimai, kaip šie tyrimai buvo atlikti, kokie parametrai įtakoja Hyperledger greitį ir pateikta tolimesnės analizės pasiūlymai.

1.2.1. Architektūra

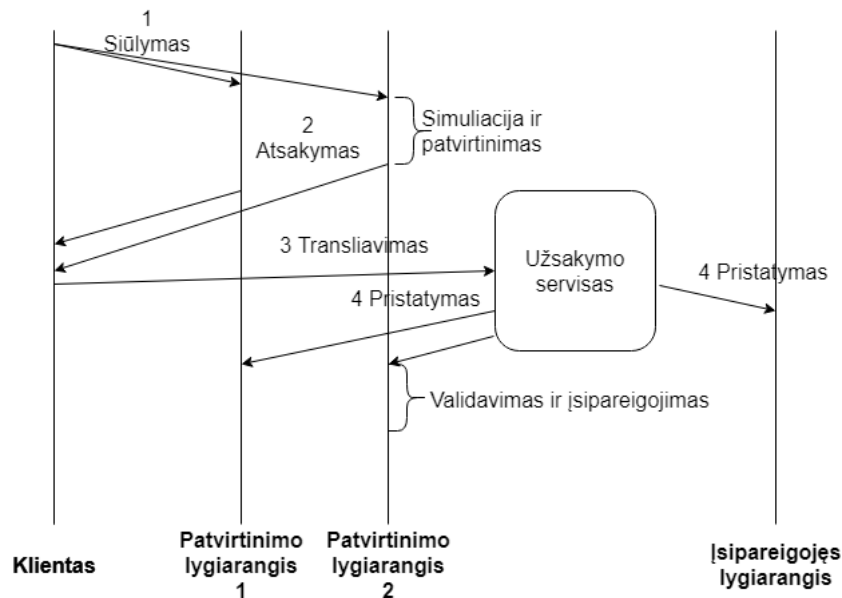
Hyperledger Fabric(toliau - Fabric) yra privati blokų grandinė skirta įmonių lygio aplikacijoms. Ši blokų grandinė gali vykdyti arbitrišką išmanųjų kontraktą parašyta Java, Go arba NodeJS kalbomis (grandinės kodai). Fabric sudaro šios esybės:

- Lygiarangis - šis mazgas yra atsakingas už grandinės kodą kurį vykdančias yra įgyvendinamas išmanusis kontraktas. Lygiarangis savyje turi visą tinklo informaciją(angl. Ledger).
- Rykiavimo servisas (angl. Ordering Service) - Rikiavimo serviso mazgas dalyvauja susitarimo(angl. consensus) protokole ir padalina bloką taip, kad jį būtų galima naudoti transakcijoms. Padalintas blokas būna persiunčiamas kitiems lygiarangiams.
- Klientas - atsakingas už transakcijos siūlymo sukūrimą, ir išsiuntimą tinklo lygiarangiams. Gavus patvirtinimą iš lygiarangių vartotojas siunčia prašymą tvarkytojui, kad jis informaciją įtrauktų į bloką ir išsiųtu ją visiems tinklo perams.

1.2.2. Tranzakcijų valdymas

Tranzakcijos vyksta trejomis fazėmis(1 pav.):

1. Patvirtinimo fazė - simuliuojama transakcija su pasirinktais vienaarangiais ir renkami būsenos pokyčiai.
2. Užsakymo fazė - užsakomos transakcijos per susitarimo protokolą.
3. Patvirtinimo fazė - patvirtinimas ir informacijos įdėjimas į didžiąją knygą.



1 pav. Transakcijų sekų diagrama

1.2.3. Parametrai

Yra grupė parametų [TNV18] kurie įtakoja Fabric greitį

- Mazgų skaičius - vartotojų skaičius privačioje blokų grandinėje.
- Transakcijų skaičius - keičiant transakcijų skaičių keičiasi vėlavimas ir pralaidumas.
- Bloko dydis - transakcijos yra sugrupuojamos į blokus. Blokai yra siunčiami visiems tinklo vienaarangiams. Užsakymo parašas būna verifikuojamas kiekvienam blokui, o perdavimo patvirtinimo parašas verifikuojamas kiekvienai transakcijai, todėl keičiant bloko dydį atsiranda kompromisas tarp pralaidumo ir vėlavimo.
- Patvirtinimo politika - diktuoja kiek transakcijų ir pasirašymų turi būti įvykdyta prieš siunčiant transakcijas užsakytoją. Didinant politikos sudėtingumą didės resursų sunaudojimas ir įvertinimo laikas.
- Kanalai - izoliuoja transakcijas viena nuo kitos ir norint persiųsti transakcijas iš vieno kanalo į kitą turi būti patvirtintos, surikiuotos ir apdorotos nepriklausomai viena nuo kitos.
- Resursų paskirstymas - kiekvienas lygiarangis vykdo grandinės kodą skirtą parašo skaičiavimams ir verifikavimo rutinoms. Keičiant procesoriaus branduolių skaičių keičiasi vykdymo greitis.

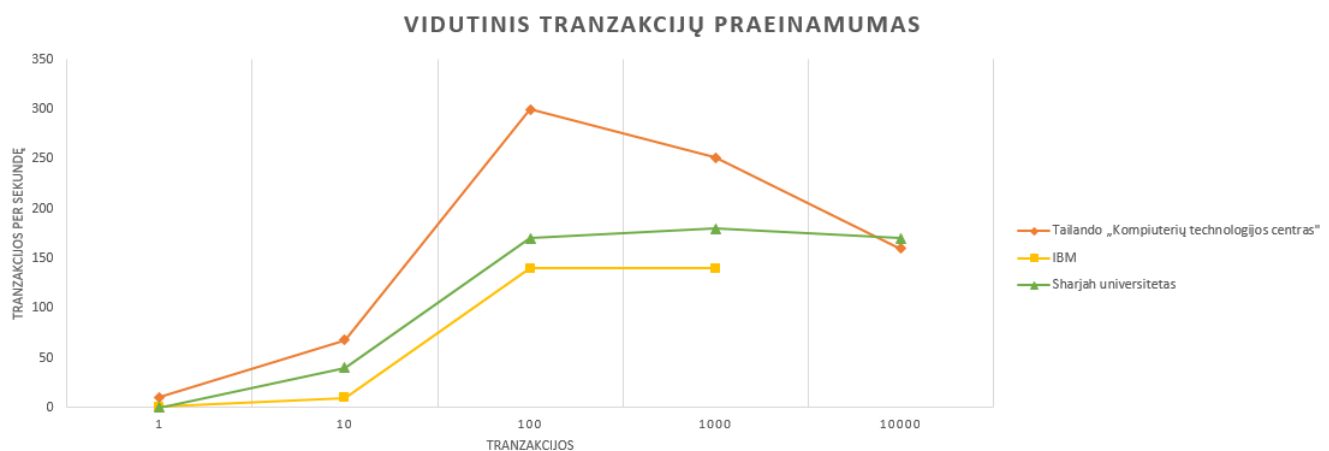
- Būsenos duomenų bazė - Fabric naudoja dvi duomenų bazines, CouchDB ir GoLevelDB, kuriuose galima saugoti didžiosios knygos būseną.

1.2.4. Testavimo metodologija

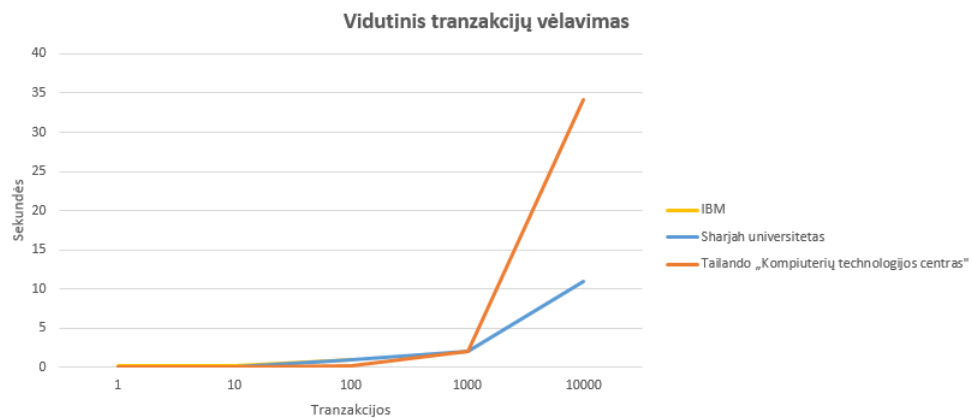
Šiame darbe apžvelgta trijų straipsnių duomenys, kuriuose naudota tokia kompiuterinė įranga:

[TNV18]	x86 64 virtuali mašina IBM SoftLayer duomenų centre. Kiekvienai virtualiai mašinai yra suteikta 32 vCPU's Intel(R) Xeon(R) CPU E5-2683 v3 @ 2.00GHz ir 32 GB atminties. Trys kliento mašinos skirtos generuoti apkrovą buvo paskirta 56 vCPU ir 128 GB RAM. Mazgai prijungti prie 3 Gbps duomenų centro tinklo
[PST17]	Amazon AWS EC2 su Intel E5-1650 8 branduolių CPU, 15GB RAM, 128GB SSD
[NQT ⁺ 18]	HPC serveris su Intel(R) Xeon(R) CPU E5-2690, 2.60 GHz, 24 branduoliai CPU, 64 GB RAM, Ubuntu 16.04

1.2.5. Tyrimų rezultatai



2 pav. Transakcijų praeinamumas



3 pav. Transakcijų vėlavimas

1.2.6. IBM tyrimo rezultatai

IBM atliktame tyrime [TNV18] buvo tiriama kaip keičiant Fabric parametrus keičiasi transakcijų praeinamumas ir vėlavimas. Didinant transakcijų atvykimo kiekį nuo 20 transakcijų per sekundę iki 100 praeinamumas didėja tiesiškai, o nuo 100 transakcijų per sekundę praeinamumas sustoja ir nebekilo. Bloko dydis praeinamumui įtakos neturėjo.

Mažiausias vėlavimas būna pasirinkus mažiausią bloko dydį. Keičiant transakcijų atvykimo kiekį nuo 25 iki 125 vėlavimas išlieka apie 0,3 sekundės ir tuomet smarkiai kyla iki 10 sekundžių transakcijų atvykimo kiekį pakėlus iki 150 transakcijų.

Iš šio tyrimo imsime geriausius rezultatus (2 pav. ir 3 pav.) ir vėliau lyginsime juos su Apache Cassandra duomenų baze.

1.2.7. Tailando „Kompiuterių technologijos centro“ tyrimo rezultatai

Tailando „Kompiuterių technologijos centro“ atliktame darbe [PST17] buvo simuliuojamos pinigų siuntimas, pinigų išdavimas ir vartotojų sukūrimas. Keliant transakcijų skaičių iki 100 praeinamumas kilo iki 299.85 transakcijų per sekundę. Didinant transakcijų skaičių iki 1000 praeinamumas pakito nežymiai, kaip ir IBM [TNV18] atliktame darbe, ir didinant transakcijų skaičių iki 10000 praeinamumas krito iki 159.76 transakcijų per sekundę. Didinant transakcijas nuo 1 iki 100 vėlavimas kilo nežymiai, nuo 0.09 iki 0.17. Padidinus transakcijas nuo 100 iki 10000 staiga pakilo vėlavimas net iki 34.08 sekundžių.

1.2.8. Sharjah universiteto tyrimo rezultatai

Sharjah universiteto mokslininkų tyrime [NQT⁺18] Fabric 0.6 ir Fabric 1.0 transakcijų praeinamumas ir vėlavimas. Nuo 10 iki 100 transakcijų praeinamumas kilo nuo 40 transakcijų per sekundę iki 165 transakcijų per sekundę, toliau didinant transakcijų skaičių praeinamumas nebe-kito. Toks rezultatas gautas naudojant Fabric 1.0 versiją ir praeinamumas geresnis negu Fabric 0.6 versijos. Didinant transakcijų skaičių nuo 10 iki 1000 vėlavimas pakilo nežymiai, nuo 0.1 sekundės iki 1 sekundės, tačiau padidinus transakcijų skaičių iki 10000 pastebėtas didelis vėlavimo pašoki-mas iki 10 sekundžių. Šio darbo metodologija buvo paremta jau minėtu Tailando mokslininkų

darbu [PST17]

1.2.9. Tyrimų rezultatų apibendrinimas

Iš aukščiau aptartų darbų([NQT⁺18; PST17; TNV18]) pastebima tendencija, kad didinant transakcijų skaičių iki 100 praeinamumas kilo tiesiškai, o didinant transakcijų skaičių toliau praeinamumas nebekilo arba net krito([PST17]).

Didinant transakcijas nuo 1 iki 1000 vėlavimas praktiškai nekyla, ir tada nuo 1000 iki 10000 vėlavimas smarkiai šoktelėja.

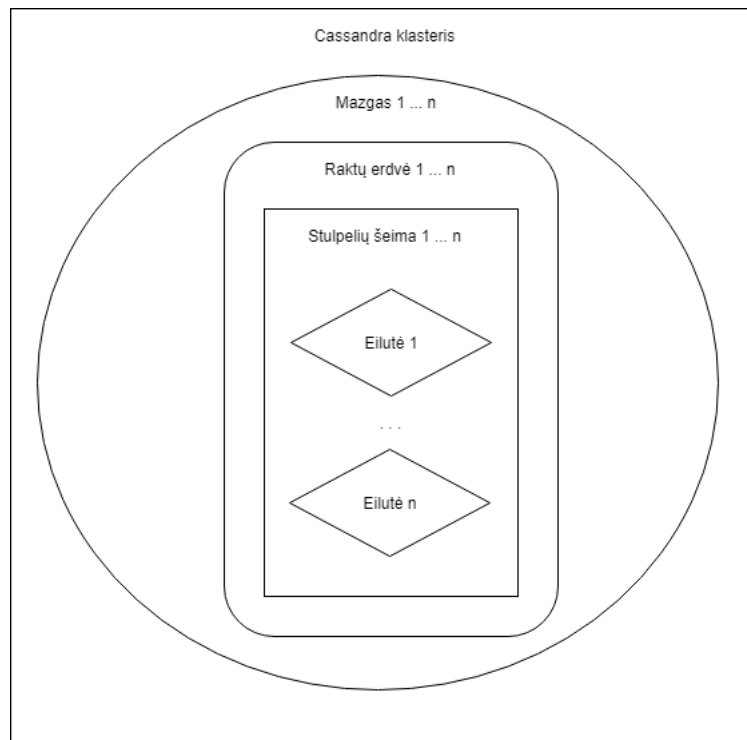
1.3. Apache Cassandra

Apache Cassandra yra atviro kodo, paskirstyta NoSQL duomenų bazė kuri palaiko klasterizavimą bei asinchroninį be valdančiojo kompiuterio duomenų replikavimą. Šios Cassandra galimybės yra panašios į blokinių grandinių duomenų bazių galimybes, todėl šiame skyriuje apžvelgsime jau padarytus Cassandra duomenų bazės našumo tyrimus.

1.3.1. Architektūra

Apache Cassandra yra atviro kodo tarpusavio paskirstytoji duomenų bazė. Cassandra laikosi tarpusavio paskirstymo modelio. Laikantis šio modelio kiekvienas tinklo mazgas turi tą pačią struktūrą kas palengvina Cassandra praplėtimą. Cassandra duomenų struktūrą sudaro:

- Stulpelis - jį sudaro pavadinimas ir reikšmė.
- Laiko žyma - rodanti kada paskutinį kartą stulpelis buvo pakeistas.
- Eilės raktas - unikalus eilės identifikatorius.
- Eilių šeima - vieta kurioje yra panašių stulpelių aibės.
- Raktų erdvė - atitinkmuo duomenų bazei reliacinėse duomenų bazėse.

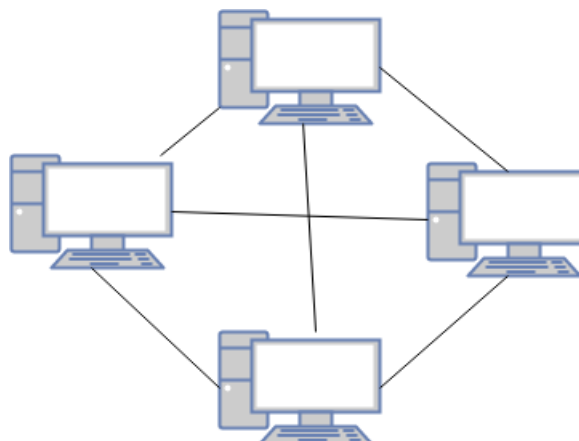


4 pav. Tranzakcijų vėlavimas

‘ Replikavimas - darant užklausas galima pasirinkti replikavimo faktorių. Replikavimo faktorius nurodo per kiek klasterio mazgų užklausos duomenys bus kopijuojami. Didesnis replikavimo reiškia, kad duomenys klasteryje bus lengviau pasiekiami. Turint duomenis tik viename mazge ir praradus ryšį su tuo mazgu duomenys tinkle dings. Tačiau nustačius aukštą replikavimo faktorių Cassandra užtrunka daugiau laiko koordinuoti duomenis tinkle.

Skaidymas - nusako kaip duomenys bus paskirstyti mazguose. Šis funkcionalumas leidžia nurodyti kaip eilių raktai bus saugomi kas įtakoja kaip bus atliekamos užklausos į duomenų bazę.

Paskalos - Cassandra naudoja „paskalų” protokolą kurio pagalba visi tinklo mazgai žino informaciją apie kitus tinklo mazgus. Mazgai seka informaciją apie tai ar su kažkuriuo mazgu buvo užmegztas ryšys arba ar mazgas dingio iš tinklo. Jeigu mazgas dingsta iš tinklo kiti mazgai laiko įrašus savyje pasakančius kokius duomenis reiks nusiųsti dingusiam mazgui kai jis vėl atsiras tinkle. Paskalų algoritmas išsiunčia „širdies dūžio” žinutę kas sekundę, kad sužinotų visų mazgų būseną.

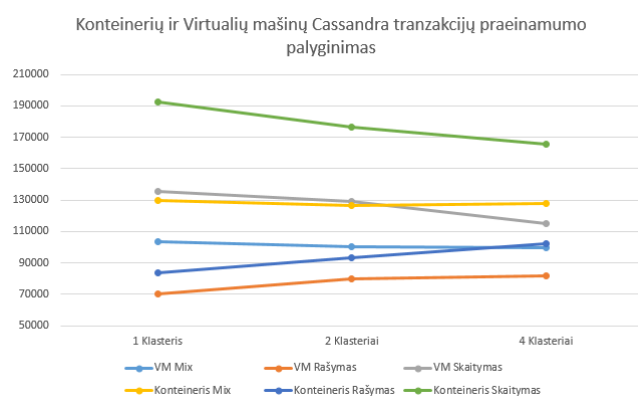


5 pav. Tranzakcijų vėlavimas

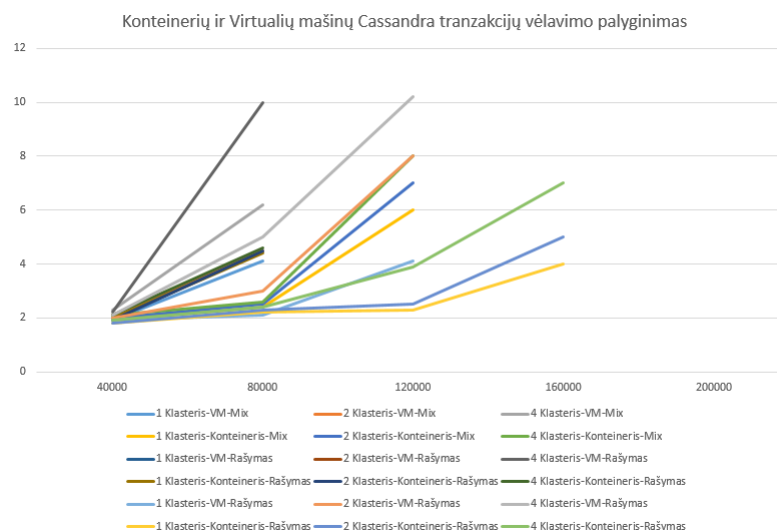
Patvarumas Cassandra duomenų bazėje yra užtikrinamas naudojant patvirtinimų žurnalą kaip sistemos atstatymo mechanizmą. Įrašai nebus pripažinti tinkle kol jie nebus įrašyti į patvirtinimų žurnalą. Po to kai įrašas atsiranda žurnale jis yra perduodamas į vietinę atmintį(memtable) ir kai tų įrašų kiekis pasiekia tam tikrą ribą jie yra perkeliama į diską SSTable formatu. Įrašai Cassandra duomenų bazėje yra labai greiti, nes atliekant rašymo užklauso nėra poreikio atlikti disko skaitymo ir paieškų. Šis funkcionalumas yra pasiekiamas naudojant memtable ir SSTable. Reliacinėse duomenų bazėse atlikti rašymą reikalauja daugiau resursų negu Cassandra duomenų bazėje.

1.3.2. Testavimo metodologija

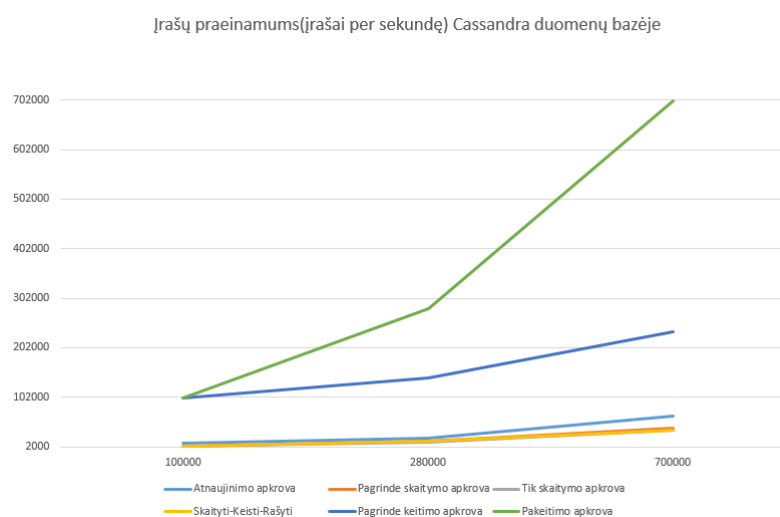
[SLC19]	Testai buvo atlikti su trimis HP serveriais DL380 G7 iš viso su 16 branduolių (naudojant HyperThreading) ir 64 GB RAM ir HDD 400 GB. Red Hat Enterprise Linux Server 7.3 (Maipo) (Kernel Linux 3.10.0-514.el7.x86_64) ir Cassandra 3.11.0 yra įdiegta į visus kompiuterius, taip pat ir virtualias mašinas. Ta pati Cassandra versija naudojama ir apkrovos generavimui. Konteinerių testavimui naudota, Docker versija 1.12.6. Virtualioms mašinoms naudota VMware ESXi 6.0.0.
[SLC19]	Testai buvo leidžiami naudojant Ubuntu Server 12.04 32bit Virtual Machine leidžiant ant VMware Player. Virtuali mašina turėjo 2GB RAM ir priimančioji mašina turi vieno mazgo Core 2 Quad 2.40 GHz su 4GB RAM ir Windows 7 operacine sistema. Duomenų bazių versijos: MongoDB version 2.4.3 ir Cassandra version 1.2.4.



6 pav. Transakcijų vėlavimas



7 pav. Transakcijų vėlavimas



8 pav. Transakcijų vėlavimas

1.3.3. Blekinge technologijų instituto tyrimo rezultatai

Blekinge technologijų instituto atliktame tyrime [SLC19] buvo atliktas Cassandra duomenų bazės našumo palyginimas lyginant duomenų bazę veikiančia konteineriuose ir virtualioje mašinoje. Buvo lyginama transakcijų praeinamumas ir vėlavimas keičiant klasterių skaičių. Įvairioje apkrovoje virtualios mašinos praeinamumas laikėsi apie 100000 transakcijų per sekundę, rašymo apkrovoje apie 75000 transakcijų per sekundę, skaitymo apkrovoje apie 120000 transakcijų per sekundę. Įvairioje apkrovoje konteinerių praeinamumas buvo apie 127000 transakcijų per sekundę, rašymo apkrovoje apie 95000 transakcijų per sekundę, o skaitymo apkrovoje apie 180000 transakcijų per sekundę. Virtualiose mašinose didinant klasterių skaičių padidėjo praeinamumas didėjo su rašymo apkrova ir mažėjo su skaitymo apkrova. Tas pats buvo pastebėta ir konteineriuose. Tyrime mažiausias vėlavimas buvo naudojant vieną Cassandra klasterį su 40000 tranzakcijų apkrova, o didžiausias vėlavimas buvo pastebėtas naudojant keturis Cassandra klasterius veikiančius ant

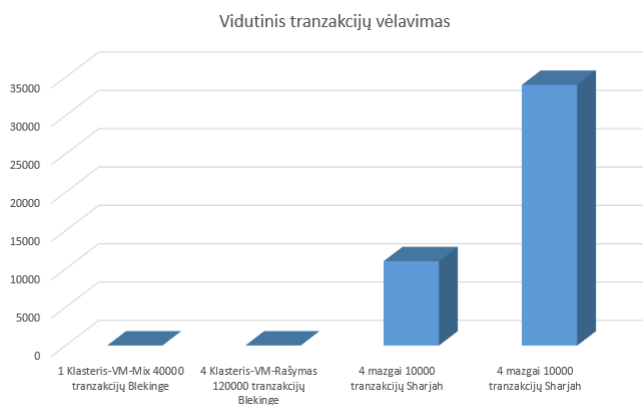
virtualių pašinių naudojant 120000 tranzakcijų apkrovą. Viso tyrimo metu buvo pastebėtas vėlavimas intervale nuo 1 iki 11 milisekundžių.

1.3.4. Coimbra politechnikos instituto tyrimo rezultatai.

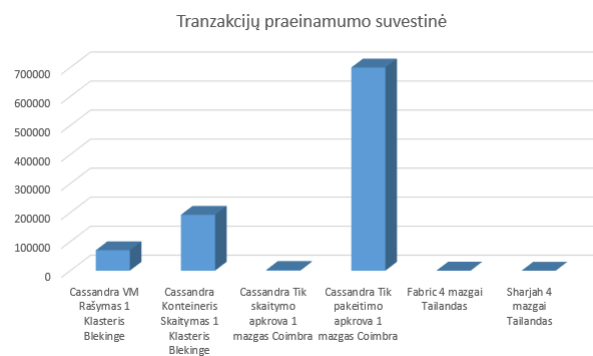
Coimbra politechnikos instituto atliktame tyrime [AB13] buvo lyginamos Cassandra ir MongoDB duomenų bazės. Lyginimas buvo atliktas naudojant tik vieną mazgą todėl buvo lygintas tik įrašų(tranzakcijų) praeinamumas su skirtingomis apkrovomis. Šio darbo tikslui žiūrėsime tik į rezultatus gautus apie Cassandra duomenų bazę. Pastebėta, kad didinant apkrovą tranzakcijų praeinamumas padidėjo. Didžiausias praeinamumas buvo pasiektas kai buvo atlikta tik vieno tipo, o ne mišruotos duomenų operacijos. Aukščiausias duomenų pakeitimo praeinamumas buvo 700000 tranzakcijų per sekundę, o skaitymo praeinamumas buvo 35000 tranzakcijų per sekundę. Mažiausias duomenų praeinamumas buvo pastebėtas su mažesnia apkrova. Žemiausias duomenų pakeitimo praeinamumas buvo 100000 tranzakcijų per sekundę, o žemiausias skaitymo praeinamumas buvo 2325 tranzakcijų per sekundę.

Pagrindinis šių tyrimų palyginimo trūkumas yra tai, kad visi šiame darbe apžvelgti tyrimai buvo atlikti su skirtingo galingumo sistemomis ir skirtingomis duomenų bazių konfigūracijomis. Ieškant literatūros šiam darbui kilo sunkumų surasti tyrimų kuriuose su ta pačia ar panašia kompiuterine įranga ir duomenų bazių konfigūracijomis tiesiogiai būtų lyginamos blokų grandinių duomenų bazės su NoSQL duomenų bazėmis. Sekančiame skirsnyje bus atliekamos Fabric ir Cassandra našumo tyrimas.

1.4. Greičių tyrimų apibendrinimas



9 pav. Transakcijų vėlavimas



10 pav. Transakcijų vėlavimas

Apžvelgtuose darbuose buvo matuojamos Apache Cassandra ir Hyperledger Fabric duomenų bazių praeinamumas ir vėlavimas. Palyginus rezultatus buvo aiškiai matoma, kad Cassandra praeinamumas buvo didesnis ir vėlavimas mažesnis. Net ir palyginus prasčiausią Cassandra praeinamumo rezultatą (2325 transakcijų per sekundę) ir geriausią Fabric rezultatą (300 transakcijų per sekundę) Cassandra duomenų bazės praeinamumas buvo geresnis 775 procentais. Lyginant prasčiausią Cassandra ir geriausią Fabric vėlavimo rezultatą Cassandra vėlavimas buvo 1078 procentų mažesnis negu Fabric.

2. Duomenų bazių palyginimas pagal CAP teoremą

2.1. CAP teorema

Eric A. Brewer 2000 [Bre00] metais pristatė savo teoremą sakančia, kad tinklinis servisas negali užtikrinti trijų savybių vienu metu:

- Neprieštarinumas(Consistency) - kiekviena skaitymo užklausa gauna naujausią informaciją arba klaidos pranešimą.
- Pasiekiamumas(Availability) - kiekviena užklauso gauna atsakymą, be garantijos, kad atsakyme bus naujausias įrašas.
- Skaidinių toleravimas(Partition tolerance) - sistema nesustoja funkcionuoti jeigu būna prarandamas arbitrišką skaičių žinučių.

Šios teoremos įrodymas pateikiamas Seth Gilbert ir Nancy Lynch straipsnyje [GL02] Kuriant paskirstytąsias sistemas, pagal panaudojimo atvejį, svarbu pasirinkti kuriuos du principus tenkins kuriama paskirstytoji sistema. Sekančiame skirsnyje bus apžvelgiama Hyperledger Fabric ir Cassandra atitikimas pagal CAP teoremą.

2.2. Hyperledger Fabric pagal CAP teoremą

Romos ir Southampton universitetų mokslininkų darbe [ASA⁺18] buvo tiriama kaip skirtingi susitarimo algoritmai įtakoja Ethereum blokų grandinę pagal CAP teoremą. Šiame darbe CAP buvo pateiktas kitos, labiau blokų grandinės atitinkantis CAP teoremos apibrėžimas:

- Neprieštarinumas(Consistency) - blokų grandinė yra neprieštaringa, jeigu yra išvengta išsišakojimų. Neprieštarinumas yra pasiekiamas susitarimo algoritmų. Jeigu neprieštarinumas nebūna pasiekiamas turime nurodyti, ar jis bus pasiektas vėliau(galiausiai neprieštaringas) ar nebus pasiektas niekada(prieštaringas)
- Pasiekiamumas(Availability) - blokų grandinė yra pasiekama, jeigu klientų pateiktos transakcijos yra apdorojamos ir galiausiai patvirtinamos ir visam laikui pridedamos prie grandinės
- Skaidinių toleravimas(Partition tolerance) - kai įvyksta tinklo skaidymasis, valdžia yra paskirstoma į atskiras grupes taip, kad skirtingos grupės negali komunikuoti tarpusavyje. Blokų grandinė turi toleruoti 1. Periodus kai tinklas veikia asinchroniškai 2. Atitinkama skaičių Bizantinių valdžių siekiančių sutrikdyti pasiekiamumą ir neprieštarinumą.

Naudojantis šiais CAP apibrėžimais sekančiuose skyriuose bus įvertinta Hyperledger fabric blokų grandinių duomenų bazė.

2.3. Hyperledger Fabric neprieštaringumas

Hyperledger Fabric yra vadinamas „galiausiai neprieštaringas“, nes norint įrašyti transakciją į blokų grandinę užtrunka laiko ją patvirtinti. Klientas sudaro savo transakciją ir išsiunčia ją patvirtinimo perą, patvirtinimo lygiarangis simuliuoja transakcijos pridėjimą į grandinę siekdamas užtikrinti pridėjimo teisingumą. Jeigu simuliacija įvyko sėkmingai tuomet transakciją gauna patvirtinimo parašą ir keliauja atgal pas klientą, kuris išsiunčia jau patvirtiną transakciją į rikiavimo servisą kuris išsiunčia naują grandinės būseną visiems tinklo perams. Per laiko tarpą, kuris praeina nuo kliento transakcijos išsiuntimo patvirtinimo perams iki laiko kol rikiavimo servisas visiems tinklo perams išsiunčia naują grandinės būseną kiti tinklo klientai gali atlikti savo transakcijas nežinodami, kad kiti klientai taip pat daro transakciją. Duomenų neprieštaringumą užtikrina rikiavimo servisas naudodamas susitarimo algoritmą todėl galiausiai visi tinklo perai gaus atnaujint informaciją, tačiau nebus užtikrinta, kad grandinė kurią gavo lygiarangis yra pačios naujausios versijos. [IBM18]

2.4. Hyperledger Fabric pasiekiamumas

Pasiekiamumą Fabric užtikrina pasiekiamumas laikydama grandinės kopijas visuose peruose, todėl klientas visados galės atlikti transakcijas į blokų grandinę. [IBM18]

2.5. Hyperledger Fabric skaidinių toleravimas

Skaidinių toleravimas yra užtikrintas visą tinklą paskirsčius ant daugelio mazgų. Tinklas išlieka veikiantis, net ir tam tikras mazgų skaičius dingtų, vartotojai vis tiek galėtų siųsti transakcijas į kitus mazgus[IBM18].

2.6. Cassandra neprieštaringumas

Kai duomenys yra rašomi į duomenų bazę užtrunka laiko duomenys pateks į visus tinklo mazgus. Kai kurie mazgai gali būtų nepasiekiami. Cassandra yra „galiausiai neprieštaringa“ duomenų baze, nėra užtikrinta, kad duomenys kurie yra skaitomi yra naujausios versijos visame tinkle. Cassandra duomenų bazėje yra įgyvendintas pasirenkamasis pasiekiamumas kai klientas pats gali nurodyti kokios lygio neprieštaringumo jis nori skaitant duomenis(mazgų skaičius kurie turi identiškus duomenis) ir rašant duomenis(į kiek mazgų bus įrašyti duomenys). Cassandra duomenų bazėje nėra užtikrinamas neprieštaringumas, nes jeigu kažkuri mazgas yra nepasiekiamas vykdant skaitymą arba rašymą ir jį pasiekti būtina siekiant užtikrinti neprieštaringumą operacija neįvyks. [CasDesk]

2.7. Cassandra pasiekiamumas

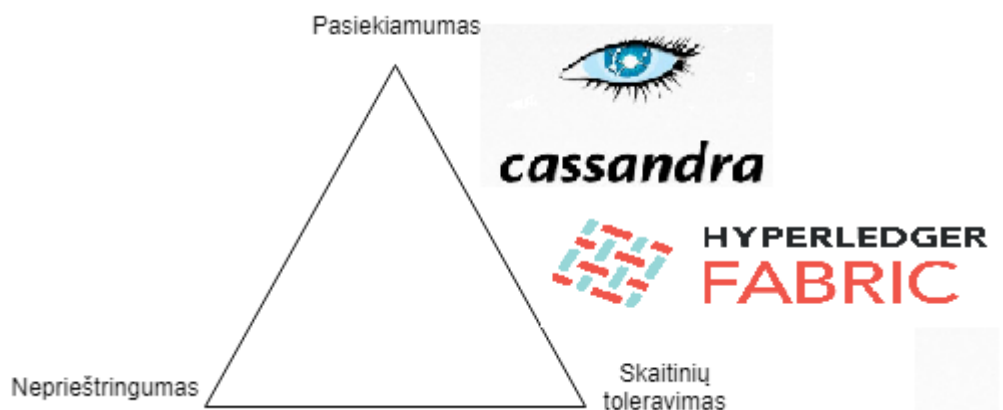
Rašant duomenis į Cassandra duomenų bazę yra padaromos kelios kopijos ir išsiunčiamos skirtingiems klasterio mazgams. Tai yra užtikrinamas, kad jeigu mazgas tampa nebepasiekiamu

duomenys nebūna prarandami.[CasDesk] Duomenų replikavimo lygį galima nurodyti kuriant duomenų bazę.

2.8. Cassandra skaidinių toleravimas

Cassandra duomenų bazė tenkina skaidinių toleravimo principą, nustojus veikti tam tikram skaičiui mazgų tinkle sistemos veikimas nenutrūksta. [CasDesk]

2.9. Apibendrinimas



11 pav. Transakcijų vėlavimas

Hyperledger Fabric ir Apache Cassandra duomenų bazės įgyvendina pasiekiamumo ir patirčių toleravimo kriterijus CAP teoremoje. Kuriant Cassandra duomenų bazę yra galimybė pakeisti konfigūracija taip, kad rašymo operacija būtų laiko sėkminga tik tuomet, kai visiems klasterio mazgams yra padaroma replika, panašiai ir skaitant galima skaityti informacija iš visų tinklo mazgų ir pasirinkti tą informaciją, kuri yra didžiojoje daugumoje mazgų. Pasirinkus tokią konfigūraciją Cassandra labiau pradeda atitikti neprieštaravimo ir patirčių toleravimo savybes tačiau tokia konfigūracija reikalauja paaukoti transakcijų greitį. Hyperledger Fabric nesuteikia tokios konfigūravimo galimybės.

3. Rezultatai ir išvados

3.1. Rezultatai

1. Šiame darbe buvo palygintas Cassandra ir Hyperledger Fabric transakcijų praeinamumas ir vėlavimas.
2. Aprašytas transakcijų praeinamumo ir vėlavimo priklausomybė nuo mazgų skaičiaus tinkle.
3. Išskirti Hyperledger Fabric parametrai kuriuos keičiant keičiasi sistemos veikimo greitis.
4. Įvertinta Cassandra ir Hyperledger architektūros atitikimas CAP teoremos principus.

3.2. Išvados

1. Yra trūkumas tyrimų lyginančių blokų grandinių duomenų bazių greičius su paskirstytoms duomenų bazės.
2. Numatytoji Cassandra ir Hyperledger Fabric duomenų bazių realizaciją, CAP teoremoje, atitinka pasiekiamumo ir skaidymo toleravimo principus.
3. Cassandra konfigūracija yra lankstesnė ir leidžia kuriant duomenų bazę padidinti duomenų neprieštarinumą sumažinus transakcijų įvykdymo greitį.
4. Cassandra duomebų bazė yra ženkliai greitesnė(iš apžvelgtų straipsnių bent 7 kartus) negu Hyperledger Fabric.

3.3. Rekomendacijos ateities darbams

1. Atlikti Cassandra ir Hyperledger Fabric praeinamumo ir vėlavimo tyrimą atlikti naudojant tą pačią kompiuterinę įrangą.
2. Atlikti Cassandra praeinamumus ir vėlavimo tyrimą keičiant konfigūraciją link transakcijų neprieštarinumo pusės ir lyginti lėtėjantį greitį su Hyperledger Fabric greičiu.
3. Atlinkti greičio lyginimo tyrimus su skirtingomis Fabric ir Cassandra versijomis.

4. Priedai

4.1. Žodynas

- Didžioji knyga(angl. ledger) - transakcijų kaupimo vieta.
- Grandinės kodas(angl. chaincode) - programa parašyta Go, Java arba NodeJS kalbomis kuri vykdo verslo logiką sutartą tarp tinklo narių.
- Išmanusis kontraktas(angl. smart contract) - kompiuteriu protokolas kurio paskirtis yra patvirtinti susitarimą.
- Lygiarangis(angl. peer) - tinklo dalyvis gaunantis ir siunčiantis informaciją.
- Mazgas(angl. node) - galutinis taškas arba įrenginys kompiuterių tinkle.
- Paskalos protokolas(angl. gossip protocol) - procedūra kai vieną rangių tinkle informaciją vienas tinklo mazgas perduoda savo kaimynams, o kaimynai savo kaimynams iki tol kol visas tinklas gauna informaciją.
- Skaidymas(angl. partition) - loginės duomenų bazės struktūros padalijimas.
- Transakcija(angl. transaction) - loginis vienetas kuriame yra vienas arba daugiau duomenų perdavimo sakinių.

Literatūra

- [AB13] Veronika Abramova ir Jorge Bernardino. Nosql databases: mongodb vs cassandra, 2013.
- [AF16] ANZ ir Wells Fargo. Distributed ledger technology and opportunities in correspondent banking, 2016.
- [ASA⁺18] De Angelis, Stefano, Aniello, Leonardo ir k.t. Pbft vs proof-of-authority: applying the cap theorem to permissioned blockchain, 2018.
- [Bre00] Eric A. Brewer. Towards robust distributed systems. (invited talk) principles of distributed computing, portland, oregon, 2000.
- [Dat17] Datastax. Modernizing banks with datastax enterprise, 2017.
- [FSW81] E.B. Fernandez, R.C. Summers ir C. Wood. Database security and integrity, 1981.
- [GL02] Seth Gilbert ir Nancy Lynch. Brewer's conjecture and the feasibility of consistent, available, partition-tolerant web services, 2002.
- [Goo19] Google. <https://trends.google.com/trends/explore?date=all&geo=us&q=bitcoin,ethereum,liteco> 2019.

- [Hyp16] Hyperledger. Linux foundation's hyperledger project announces 30 founding members and code proposals to advance blockchain technology, 2016.
- [Hoe08] Jaap-Henk Hoepman. Distributed double spending prevention, 2008.
- [IBM18] IBM. Hyperledger fabric: a distributed operating system for permissioned blockchains, 2018.
- [Nak08] S. Nakamoto. A peer-to-peer electronic cash system, 2008.
- [Nak09] S. Nakamoto. Bitcoin: a peer-to-peer electronic cash system, 2009.
- [NQT⁺18] Qassim Nasir, Ilham A. Qasse, Manar Abu Talib ir Ali Bou Nassif. Performance analysis of hyperledger fabric platforms, 2018.
- [PST17] Suporn Pongnumkul, Chaiyaphum Siripanpornchana ir Suttipong Thajchayapong. Performance analysis of private blockchainplatforms in varying workloads, 2017.
- [SLC19] Sogand Shirinbab, Lars Lundberg ir Emiliano Casalicchio. Performance comparision between scaling of virtual machines and containers using cassandra nosql database, 2019.
- [TNV18] Parth Thakkar, Senthil Nathan N ir Balaji Viswanathan. Performance benchmarking and optimizing hyperledger fabric blockchain platform, 2018.