VILNIAUS GEDIMINO TECHNIKOS UNIVERSITETAS

FUNDAMENTINIŲ MOKSLŲ FAKULTETAS

INFORMACINIŲ SISTEMŲ KATEDRA

ŽURNALINIŲ ĮRAŠŲ SURINKIMAS, AGREGAVIMAS IR ANALIZĖ VIRTUALIOJE INFRASTRUKTŪROJE

Referatas

Darbą atliko: Simonas Riška

Darbą tikrino: prof. dr. [redacted]

VILNIUS 2025

TURINYS

LENTELIŲ SĄRAŠAS

PAVEIKSLŲ SĄRAŠAS

SANTRUMPŲ ŽODYNAS

API (angl. *Application Programming Interface*) – taikomųjų programų programavimo sąsaja.

VM (angl. *Virtual Machine*) – virtuliaoji mašina.

ACL (angl. *Access Control List*) – prieigos kontrolės sąrašas.

IAM (angl. *Identity and Access Management*) – tapatybės ir prieigos valdymas.

SIEM (angl. *Security Information and Event Management*) – saugumo informacijos ir įvykių valdymo sistema.

BDAR – Bendrasis duomenų apsaugos reklamentas.

OS (angl. *Operating System*) – operacinė sistema.

TLS (angl. *Transport Layer Security*) – transportinio lygmens saugumas.

ĮVADAS

Debesų kompiuterijos ir virtualizacijos sprendimai šiuolaikinėje informacinių technologijų eroje keičia tradicinės infrastruktūros sampratą – organizacijos vis dažniau renkasi virtualias mašinas, konteinerizaciją ir debesų paslaugas dėl jų lankstumo, mastelio keitimo galimybių ir efektyvaus resursų panaudojimo, tačiau kartu su šiais privalumais kyla ir nauji iššūkiai, kurie yra susiję su žurnalinių įrašų surinkimu, agregavimu bei analize.

Žurnalai yra vienas pagrindinių šaltinių, kurie leidžia suprasti sistemų veiklą, aptikti saugumo pažeidimus bei atlikti incidentų analizę. Tradiciškai žurnalai yra renkami iš fizinių serverių ir tinklo įrenginių, bet virtualioje infrastruktūroje žurnalinių įrašų sistema tampa gerokai sudėtingesnė – atsiranda dinamiškos virtualios mašinos, trumpalaikiai konteineriai, kelių lygių abstrakcijos (pvz., hipervizoriai, *orchestration* sistemų žurnaliniai įrašai), be to, žurnaliniai įrašai gali būti trumpalaikiai ir priklausomi nuo tiekėjo (pvz., „cloud-native“ žurnaliniai įrašai), kas kelia iššūkių jų išsaugojimui ir analizavimui ilgalaikėje perspektyvoje.

Šio darbo tikslas – išanalizuoti žurnalinių įrašų surinkimo, agregavimo ir analizės procesus virtualioje infrastruktūroje bei palyginti juos su atitinkamais sprendimais fizinėje infrastruktūroje. Darbe bus aptariami aktualūs moksliniai tyrimai, analizuojami realūs pavyzdžiai ir apibendrinamos šiuolaikinės technologijos, leidžiančios užtikrinti saugų – toks teorinis pagrindas leidžia giliau suprasti tiek technologinius, tiek saugumo bei teisės iššūkius tvarkant žurnalinius įrašus šiuolaikinėse virtualizuotose IT aplinkose.

Toliau pateiktose dalyse bus sistemingai nagrinėjami žurnalinių įrašų tipai, jų surinkimo architektūra fizinėje ir virtualioje infrastruktūroje, agregavimo priemonės, analizės metodai, saugumo aspektai, palyginimo lentelė bei pateikiamos apibendrintos išvados.

Darbo objektas – žurnalinių įrašų surinkimo, agregavimo ir analizės ypatumai virtualioje infrastruktūroje.

Darbo tikslas – išanalizuoti žurnalinių įrašų valdymo ypatumus virtualioje infrastruktūroje ir palyginti juos su atitinkamais proesais fizinėje aplinkoje.

Darbo uždaviniai:

1. Apžvelgti žurnalinių įrašų tipus ir jų paskirtį informacinėse sistemose.
2. Išanalizuoti žurnalinių įrašų surinkimo, agregavimo ir analizės procesus fizinėje ir virtualioje infrastruktūroje.
3. Įvardyti pagrindinius iššūkius, saugumo rizikas ir skirtumus virtualioje aplinkoje.
4. Pateikti pagrįstas rekomendacijas dėl žurnalų valdymo virtualioje infrastruktūroje.

1. Žurnalinių įrašų paskirtis ir tipai

Žurnalinių įrašų (angl. *log records*) paskirtis informacinių technologijų infrastruktūroje yra itin svarbi, nes vienas iš pagrindinių šaltinių, leidžiančių aptikti anomalijas, atsekti incidentų eigą, užtikrinti atitiktį reikalavimams ir vykdyti retrospektyvų tyrimą tiek fizinėje, tiek virtualioje infrastruktūroje (Cândido et al., 2019). Žurnaliniai įrašai yra būtini ne tik sistemų stebėjimui realiuoju laiku, bet ir retrospektyviam audito duomenų rinkimui – taip užtikrinant sistemos vientisumą ir saugumą.

Pagrindinės žurnalinių įrašų paskirtys:

- Incidentų aptikimas ir analizė – leidžia greitai identifikuoti problemas ir atsekti jų kilmę.
- Atitikties užtikrinimas – daugelis standartų, tokių kaip ISO 27001, BDAR reikalauja saugoti veiklos įrašus.
- Našumo stebėsena – padeda įvertinti paslaugų kokybę ir atpažinti sistemos problemas.
- Automatinis anomalijų aptikimas – pažangūs sprendimai, tokie kaip MoniLog leidžia realiu laiku klasifikuoti neįprastą veiklą (Vervaet, 2021).

Virtualioje infrastruktūroje šios paskirtys tampa dar aktualesnė dėl greit kintančių sistemų, trumpalaikių virtualių mašinų egzistavimo ciklų ir riboto fizinio priėjimo (Preethi D, 2020).

Remiantis sistemine žurnalinių įrašų analizės schema (Cândido et al., 2019) žurnalinius įrašus galima skirstyti pagal jų kilmę:

1 lentelė. Žurnalinių įrašų rūšių klasifikacija pagal infrastruktūros sluoksnį (Cândido et al., 2019; Wang et al., 2025)

Tipas	**Aprašymas**
Sisteminiai žurnaliniai įrašai	Fiksuoja operacinės sistemos veiklą: branduolio klaidas, paleidimo eigą
Programiniai žurnaliniai įrašai	Kuriami programų, pvz., „nginx error.log“, „app.log“
Tinklo žurnaliniai įrašai	Pvz., ugniasienių, maršrutizatorių, IDS/IPS sistemų įrašai
Aplikacijų žurnaliniai įrašai	Vidinės verslo logikos veiklos įrašai: vartotojų prisijungimai, API kvietimai
Auditų žurnaliniai įrašai	Veiksmų sekos įrašai, skirti atitikties įrodymui

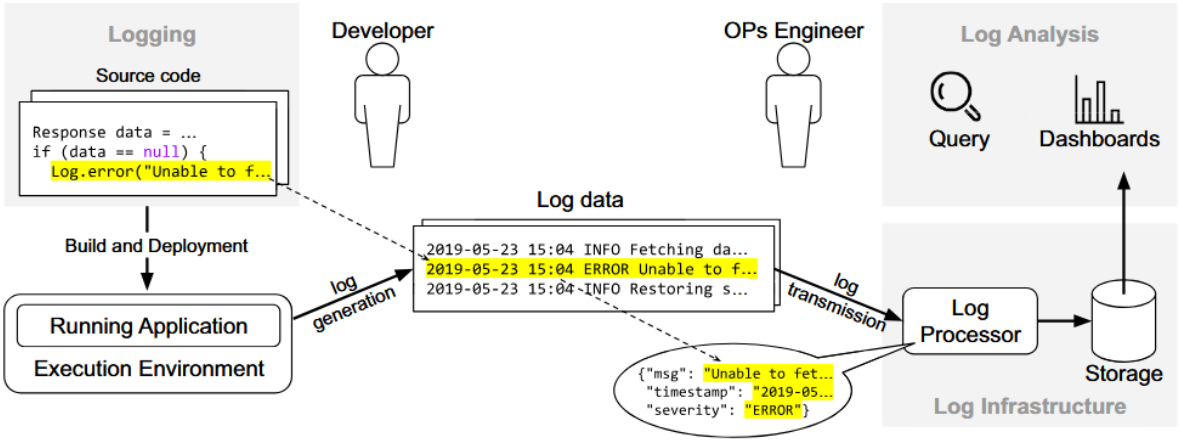
Saugumo žurnalai	Įskaitant autentifikavimo, autorizacijos, įsibrovimo bandymų įrašus
Virtualizacijos žurnalai	Pvz., virtualios mašinos paleidimo, perkėlimo, išjungimo žurnaliniai įrašai (Wang et al., 2025)

Skirtingai nei fizinėje infrastruktūroje, virtualioje aplinkoje žurnaliniai įrašai dažnai yra:

- Trumpalaikiai, nes virtuali mašina gali būti nutraukta ar perkelta.
- Sluoksniuoti, t.y. vienam įvykiui fiksuoti reikia žurnalinių įrašų tiek iš virtualios mašinos, tiek iš hipervizoriaus.
- Priklausomi nuo debesijos tiekėjo, todėl žurnalai kartais yra neprieinami galutiniam naudotojui (Preethi D, 2020).

Tai kelia naujų iššūkių, kurie išsamiau analizuojami tolesniuose skyriuose.

1 pav. Žurnalų duomenų gyvavimo ciklas informacinių sistemų kontekste (Cândido et al., 2019).



2. Žurnalinių įrašų surinkimas: fizinė ir virtuali infrastruktūra

Efektyvus žurnalinių įrašų surinkimas yra būtinas norint užtikrinti sistemų stebėseną, analizę ir incidentų aptikimą tiek tradicinėje fizinėje, tiek virtualioje infrastruktūroje, tačiau šių dviejų aplinkų architektūriniai skirtumai lemia esminius pokyčius duomenų surinkimo būduose ir saugumo užtikrinimo mechanizmuose (Cândido et al., 2019).

Fizinėje infrastruktūroje žurnalai renkami tiesiogiai iš konkrečių serverių, tinklo įrenginių ar programinės įrangos komponentų. Naudojami agentai (pvz., *rsyslog*, *syslog-ng*), kurie perduoda įrašus į centrinę saugyklą, naudodami UDP/TCP protokolus arba SSH (Akbaş, 2024). Kadangi įrenginiai yra pastovūs, identifikuoti šaltinius ir užtikrinti jų patikimumą paprasta – dažniausiai užtenka IP adresų ar MAC identifikatorių.

Be to, tokioje aplinkoje žurnalai yra ilgalaikiai, nes prietaisai veikia nenutrūkstamai, o jų sistemos retai keičiamos ar „išjungiamos“ (Preethi D, 2020) – ši stabili struktūra leidžia naudoti paprastesnius įrankius, nes šaltinių dinamika yra menka.

Virtualioje aplinkoje situacija iš esmės keičiasi – čia žurnalai gali būti generuojami keliuose sluoksniuose vienu metu: virtualios mašinos viduje, hipervizoriuje (pvz., *ESXi*), tinklo virtualizacijos komponentuose (pvz., *vSwitch*), bei debesų paslaugų valdymo lygmenyje (pvz., *AWS CloudTrail*) (Cândido et al., 2019; Wang et al., 2025)– kiekvienas iš šių šaltinių naudoja skirtingą logikos, laiko žymėjimo ir prieigos modelį.

Vienas esminių iššūkių – trumpalaikis egzistavimas – virtualios mašinos gali būti sukurtos, perkeltos arba sunaikintos per sekundes, o jei žurnalai nėra siunčiami į išorinę saugyklą, jie tiesiog išnyksta (Preethi D, 2020), o tai taip pat apsunkina identifikaciją – virtuali mašina gali keisti IP adresus, veikti skirtinguose hostuose ar net regionuose.

Wang et al. (2025) išsamiai aprašo, kaip konteineriuose žurnalų izoliacija yra silpna – vienas konteineris gali netyčia ar tyčia matyti kitų egzempliorių įrašus, nes bendrinama ta pati saugykla, todėl jie siūlo „POGs“ sistemą, kuri leidžia kiekvienam konteineriui turėti atskirą žurnalų konfigūraciją, saugyklą ir prieigos teises.

2 lentelė. Žurnalinių įrašų valdymo aspektų skirtumai tarp fizinės ir virtualios infrastruktūros

Savybė	**Fizinė infrastruktūra**	**Virtuali infrastruktūra**	**Šaltinis**
Įrenginių pastovumas	Pastovūs serveriai	Dinamiškos VM ir konteineriai	(Preethi D, 2020)
Žurnalų ilgaamžiškumas	Stabilus, kaupiami vietoje ar SIEM sistemoje	Trumpalaikiai, jei neišsaugomi centralizuotai	(Wang et al., 2025)
Prieigos kontrolė	Fizinė, OS pagrindu	Reikalauja daugiasluoksnės izoliacijos	(Wang et al., 2025)
Surinkimo technologijos	Rsyslog, syslog-ng, SNMP	Fluentd, Filebeat, cloud API	(Cândido et al., 2019)
Šaltinių įvairovė	Ribota (serveriai, įranga)	Daugialypė: VM, konteineriai, debesų paslaugos	(Cândido et al., 2019)

3. Žurnalinių įrašų agregavimas ir orkestravimas

Surinkti žurnalinius įrašus iš įvairių sistemų nepakanka – būtina juos centralizuotai kaupti, struktūrizuoti ir valdyti, o tam pasitelkiami agregavimo ir orkestravimo sprendimai – tokie sprendimai bei procesai leidžia analitikams, sistemų administratoriams ar saugumo specialistams greitai pasiekti, filtruoti ir analizuoti įrašus, gaunamus iš daugelio skirtingų šaltinių (Cândido et al., 2019).

Virtualioje infrastruktūroje žurnalai dažnai saugomi virtualios mašinos viduje, tačiau tokie įrašai yra laikini – jei virtuali mašina išjungiama ar pašalinama, visi lokaliai saugoti žurnalai išnyksta, todėl būtina juos perduoti į išorinę saugyklą. Cândido et al. (2019) pažymi, kad turėti tik įrašus nepakanka – reikia turėti visus reikiamus komponentus, kad būtų galima tiksliai sekti įvykio šaltinį ir laiką.

Wang et al. (2025) išskiria, kad bendros žurnalų saugyklos tarp virtualios mašinos ar konteinerių gali sukelti saugumo riziką, pavyzdžiui, vienas konteineris gali matyti kito veiklos įrašus, jeigu nėra įgyvendinta tinkama izoliacija, todėl saugus agregavimas turi užtikrinti ne tik duomenų perdavimą, bet ir privatumo išlaikymą.

Tam naudojami šie pagrindiniai komponentai:

- Surinkimo agentai, tokie kaip *Filebeat*, *Fluentd* ar *Logstash*, kurie renka įrašus iš VM, konteinerių ar kitų įrenginių ir siunčia juos į centrinę sistemą (Cândido et al., 2019).
- Saugojimo ir indeksavimo sistemos, kaip *Elasticsearch*, kurios leidžia greitai ieškoti įrašų pagal datą, vartotoją ar kitus kriterijus (Cândido et al., 2019).
- Vizualizacijos sprendimai, kaip *Kibana*, leidžiantys kurti valdymo skydus (dashboard), filtruoti duomenis ir sekti įvykius realiuoju laiku (Vervaet, 2021).
- Saugumo analizės sistemos, kaip *Wazuh* ar *Splunk*, kurios leidžia taikyti koreliacijos taisykles, aptikti atakas ar pažeidimus (Akbaş, 2024)

Šie įrankiai dažnai veikia kartu – pavyzdžiui, „ELK Stack“ (*Elasticsearch, Logstash, Kibana*) yra vienas populiariausių atvirojo kodo sprendimų žurnalų surinkimui ir peržiūrai (Cândido et al., 2019).

Agregavimas surenka įrašus, o orkestravimas juos struktūrizuoja, identifikuoja, priskiria šaltiniui ir paruošia analizei, o tai ypač svarbu, kai žurnalai gaunami iš skirtingų sistemų, kurių įrašai naudoja skirtingus formatus (Preethi D, 2020).

Karanjai et al. (2024) savo darbe apie sistemą „LogBabylon“ pabrėžia, kad įvairūs įrašų šaltiniai generuoja skirtingus, dažnai neapibrėžtus tekstinius įrašus, kuriuos sunku apjungti – jie siūlo naudoti didžiuosius kalbos modelius kartu su papildoma informacijos paieška (angl .Retrieval-Augmented Generation), kad žurnalai būtų susieti ir suprasti automatiškai.

Šiuolaikinėse sistemose, kaip rašo Vervaet (2021), įrašų kiekis gali siekti milijonus įrašų per minutę, todėl be automatizuotos analizės ir gerai organizuotos infrastruktūros, žmogaus darbas tampa praktiškai neįmanomas.

Debesų paslaugų tiekėjai, tokie kaip *Amazon Web Services*, *Microsoft Azure* ar *Google Cloud*, siūlo savo žurnalų agregavimo ir analizės sistemas, pvz.:

- *AWS CloudWatch Logs*
- *Azure Monitor*
- *GCP Cloud Logging*

Tačiau Preethi D (2020) pastebi, kad šios paslaugos dažnai yra uždaros, priklausomos nuo tiekėjo ir ne visada leidžia pasiekti žurnalų įrašus tokia forma, kokios reikia teisinei ar forensinei analizei.

4. Žurnalinių įrašų analizė ir anomalių įvykių aptikimas

Surinkti ir suagreguoti žurnalų įrašai įgauna prasmę tik tada, kai jie yra tinkamai analizuojami – analizė leidžia ne tik nustatyti sistemos būseną, bet ir aptikti neįprastą veiklą, kuri gali rodyti saugumo pažeidimus, sistemos klaidas ar kitus incidentus (Cândido et al., 2019).

Tradiciškai žurnalų įrašų analizė buvo atliekama rankiniu būdu. Sistemų administratoriai naudodavo komandų eilutės įrankius, tokius kaip *grep*, *awk* ar *less*, kad galėtų ieškoti tam tikrų raktinių žodžių ar klaidų žinučių (Akbaş, 2024), tačiau tai tinka tik mažo masto sistemoms, kuriose įrašų kiekis ribotas.

Didelėse infrastruktūrose, ypač virtualioje aplinkoje, žurnalų kiekiai auga eksponentiškai. Vervaet (2021) pastebi, kad debesijos platformose per vieną minutę gali būti sugeneruojama šimtai tūkstančių ar net milijonai įrašų. Rankinis metodas tokiu atveju tampa nepraktiškas, todėl būtina naudoti automatizuotus sprendimus.

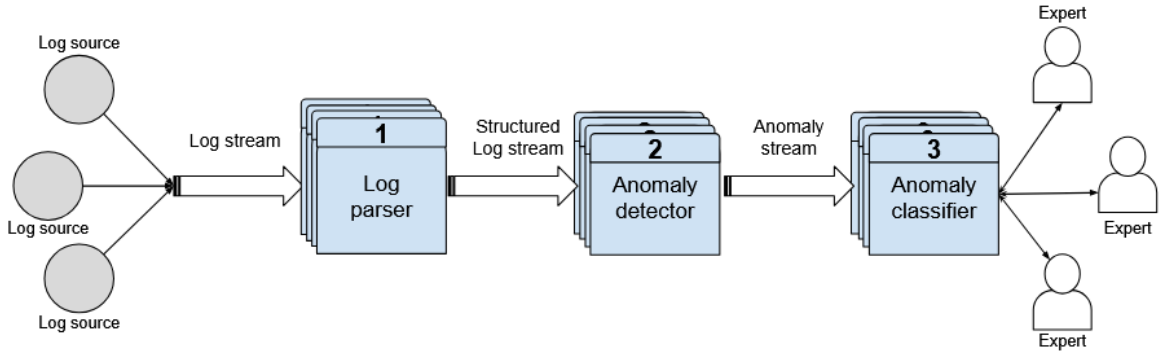
Pirmas automatizavimo žingsnis yra naudoti filtravimą pagal raktinius žodžius arba reguliarias išraiškas (*regex*) – tai leidžia greitai surasti tam tikras klaidas ar įvykius, toks metodas yra paprastas, tačiau jis negali aptikti sudėtingesnių sekų, kurios iš pirmo žvilgsnio atrodo normalios, bet jų kombinacija rodo grėsmę (Cândido et al., 2019).

Akbaş (2024) pažymi, kad tikėtini pažeidimai gali būti paskleisti tarp kelių žurnalų šaltinių – pvz., autentifikavimo bandymai viename serveryje ir neleistini API kvietimai kitame – todėl būtinas įrašų koreliavimas.

Dirbtinis intelektas tampa vis svarbesnis analizuojant žurnalų įrašus. Jis leidžia aptikti anomalias sekas, kurių negalima numatyti iš anksto – Vervaet (2021) sukūrė sistemą „MoniLog“, kuri remiasi trijų etapų metodu:

1. Žurnalų struktūrizavimas ir loginių laukų atpažinimas.
2. Anomalijų sekų identifikavimas.
3. Aptiktų įvykių klasifikavimas pagal svarbą.

2 pav. „MoniLog“ žurnalų analizės sistemos struktūra (Vervaet, 2021)



Ši sistema išmoksta iš administratoriaus veiksmų ir palaipsniui gerina klasifikavimo tikslumą – tai leidžia sumažinti klaidingų pavojaus signalų kiekį ir susikoncentruoti į realias grėsmes.

Karanjai et al. (2024) taip pat siūlo pažangų sprendimą „LogBabylon“, kuris analizuoja įrašus naudodamas didžiuosius kalbos modelius kartu su paieškos sistemomis (angl. *Retrieval-Augmented Generation*) – tokia sistema leidžia sugretinti tarpusavyje nesusijusius įrašus ir pateikti žmogui suprantamą įžvalgą apie incidentą.

Daug organizacijų naudoja SIEM (saugumo informacijos ir įvykių valdymo) sistemas, tokias kaip *Wazuh*, *Splunk* ar *IBM QRadar* – šios sistemos ne tik kaupia žurnalus, bet ir taiko taisykles bei koreliacijos algoritmus, kad būtų galima greitai identifikuoti pažeidimus (Akbaş, 2024), tačiau, kaip pažymi Akbaş (2024), net ir modernios SIEM sistemos patiria sunkumų, kai duomenų apimtis labai didelė arba kai reikia jungti duomenis iš skirtingų infrastruktūros lygių (pvz., virtualios mašinos, konteineriai, debesų API).

Norint atlikti patikimą analizę po incidento, būtina užtikrinti, kad žurnalų įrašai būtų neištrinti, nepakoreguoti ir laiku sužymėti – Preethi D (2020) siūlo naudoti šifruotus įrašus bei vadinamąją „praeities įrašo įrodymo“ (angl. *Proof of Past Log*) schemą, kuri leidžia įrodyti, kad įrašas egzistavo tam tikru momentu. Jie taip pat naudoja „Bloom“ filtrus, kad greitai būtų galima patikrinti, ar konkretus įrašas priklausė konkrečiam įvykiui – toks modelis ypač svarbus debesų kompiuterijoje, kur įrašai dažnai yra prieinami tik per tiekėjo valdomas sistemas.

5. Žurnalinių įrašų saugumas ir teisiniai aspektai

Virtualioje infrastruktūroje žurnalinių įrašų saugumas tampa ypač svarbus, nes šie duomenys dažnai naudojami kaip pagrindiniai įrodymai saugumo incidentų tyrimuose, audituose ar net teisiniuose ginčuose (Preethi D, 2020), Jei įrašai yra pažeidžiami, ištrinti, netikslūs arba prieinami neteisėtiems asmenims, organizacija gali ne tik prarasti svarbią informaciją, bet ir pažeisti teisės aktus.

Viena didžiausių grėsmių virtualioje aplinkoje yra žurnalų įrašų praradimas, kai virtuali mašina ar konteineris yra sustabdomas ar ištrinamas (Wang et al., 2025). Jei įrašai nėra iš karto perduodami į saugią, išorinę saugyklą, jie tampa neprieinami net tyrimams. Cândido et al. (2019) pabrėžia, kad tai ypač svarbu, kai žurnalai saugomi laikinuose failuose ar RAM diskuose.

Kita svarbi problema – įrašų vientisumo stoka. Kai įrašus galima modifikuoti be aiškios prieigos kontrolės ar be šifravimo, atsiranda rizika, kad užpuolikas galės paslėpti savo veiksmus. Preethi D (2020) siūlo naudoti viešojo rakto kriptografiją: kiekvieno naudotojo veiklos įrašai šifruojami jo viešuoju raktu, o analizę gali atlikti tik turintysis privatų raktą – tai padidina konfidencialumą ir neleidžia nei kitiems naudotojams, nei debesų tiekėjui be leidimo matyti duomenų.

Norint, kad žurnalų įrašai būtų pripažįstami kaip patikimi įrodymai, būtina užtikrinti jų autentiškumą. Vienas iš sprendimų – naudoti „praeities įrašo įrodymo“ (angl. *Proof of Past Log*) mechanizmą (Preethi D, 2020) – jis naudoja kriptografinius parašus arba vienkrypčius maišos algoritmus (pvz., HMAC), kad kiekvienas įrašas būtų neatskiriamai susietas su prieš tai buvusiu – tokie mechanizmai ne tik leidžia atsekti, ar duomenys buvo pakeisti, bet ir garantuoja įrašų nuoseklumą. Karanjai et al. (2024) pabrėžia, kad be tokios technologijos dirbtinis intelektas, naudojamas analizei, gali remtis klaidingais duomenimis, o tai iškraipo rezultatus.

Dėl jautraus žurnalų pobūdžio būtina apriboti, kas gali matyti ar keisti įrašus. Akbaş (2024) nurodo, kad viena dažniausių pažeidimų priežasčių yra netinkamas teisių valdymas, kai net įprasti naudotojai gali skaityti ar ištrinti įrašus, todėl rekomenduojama naudoti:

- specialias prieigos kontrolės politikas (ACL),
- identiteto patvirtinimo mechanizmus (pvz., *Kerberos*, IAM),
- šifruotą perdavimą (pvz., TLS) ir saugyklas (pvz., S3 su šifravimu).

Be to, kai kurie debesų paslaugų tiekėjai riboja žurnalų laikymo trukmę, todėl, kaip pažymi Preethi D (2020), organizacijos turi kopijuoti žurnalus į savo nepriklausomas saugyklas, kad būtų užtikrintas ilgalaikis jų prieinamumas, ypač kai reikia laikytis teisinių reikalavimų, pvz., BDAR ar ISO 27001.

Įvairūs standartai ir teisės aktai reikalauja, kad būtų kaupiami žurnalų įrašai tam tikrą laiką ir kad būtų užtikrintas jų vientisumas. Pavyzdžiui:

- BDAR (Bendrasis duomenų apsaugos reglamentas) reikalauja, kad bet koks duomenų tvarkymo veiksmas būtų atsekamas (Akbaş, 2024),
- PCI-DSS – reikalauja saugoti žurnalus bent 1 metus ir analizuoti juos kasdien (Akbaş, 2024),
- ISO 27001 – rekomenduoja žurnalų analizę kaip vieną iš saugumo valdymo procesų (Cândido et al., 2019).

Nesugebėjimas užtikrinti šių reikalavimų gali lemti teisines sankcijas, reputacijos praradimą arba incidentų nenustatymą laiku.

6. Virtualios ir fizinės infrastruktūros palyginimas

Žurnalinių įrašų rinkimo ir analizės tikslai tiek fizinėje, tiek virtualioje infrastruktūroje yra panašūs, tačiau jų įgyvendinimas skiriasi iš esmės – šie skirtumai kyla dėl architektūros, resursų valdymo ir saugumo modelių, taigi, norint tinkamai taikyti analizės bei saugojimo priemones, būtina suprasti šių dviejų aplinkų ypatumus (Akbaş, 2024; Cândido et al., 2019).

Palyginimo aspektai yra šie:

1. Įrašų šaltinis – fizinėje aplinkoje šaltiniai dažniausiai yra konkretūs, nekintantys serveriai ar tinklo įrenginiai, o virtualioje – įrašai gali būti sugeneruoti virtualios mašinos viduje, hipervizoriuje, tinklo virtualizacijos sluoksnyje ar debesų paslaugų kontrolės lygyje (Preethi D, 2020; Wang et al., 2025).
2. Įrenginių stabilumas – fiziniai serveriai dažniausiai veikia ilgą laiką be perkėlimo, o virtualios mašinos gali būti sustabdytos, perkeltos į kitą hipervizorių ar ištrintos per kelias sekundes – dėl to žurnalų duomenų tęstinumas yra sudėtingesnis (Cândido et al., 2019).
3. Prieiga prie įrašų – fizinėje infrastruktūroje administratorius dažnai turi tiesioginę prieigą prie įrenginio disko ir gali atkurti ar peržiūrėti įrašus, virtualioje aplinkoje įrašai gali būti saugomi tik debesies paslaugų tiekėjo valdomose sistemose, kur prieiga ribojama ar apmokestinama (Akbaş, 2024; Preethi D, 2020)
4. Duomenų kiekis – virtualiose aplinkose įrašų apimtis dažniausiai didesnė, nes automatiniai procesai generuoja daugiau informacijos: orkestratoriai, konteinerių platformos, vidiniai API kvietimai ir kt. Vervaet (2021) pažymi, kad debesijos sistemose vienos paslaugos žurnalų kiekis gali viršyti milijoną įrašų per valandą.

5. Saugumo kontrolė – fizinėje aplinkoje saugumo kontrolė grindžiama operacinių sistemų teisių valdymu, virtualioje – be to dar reikalinga izoliacija tarp konteinerių ar virtualios mašinos, naudojant papildomus įrankius ir konfigūracijas, kaip siūlo Wang et al. (2025) su POGs sistema.
6. Duomenų praradimo rizika – virtualioje aplinkoje žurnalų duomenys gali būti prarasti, jei jie neperduodami realiu laiku į išorinę sistemą, fizinėje aplinkoje jie saugomi diske ilgiau, kol įrenginys veikia (Preethi D, 2020).

3 lentelė. Virtualios ir fizinės IT infrastruktūros skirtumai žurnalinių įrašų kontekste

Kriterijus	**Fizinė infrastruktūra**	**Virtuali infrastruktūra**	**Šaltinis**
Šaltinių stabilumas	Pastovūs serveriai	Dinamiškos VM ir konteineriai	(Cândido et al., 2019)
Prieigos lygis	Pilna prieiga prie disko	Priklausoma nuo debesijos tiekėjo	(Akbaş, 2024)
Saugumo kontrolė	OS lygio	Reikalinga papildoma izoliacija	(Wang et al., 2025)
Duomenų kiekis	Vidutinis	Labai didelis (ypač debesijoje)	(Vervaet, 2021)
Praradimo rizika	Maža, kol veikia serveris	Didelė be centralizuoto perdavimo	(Preethi D, 2020)
Įrašų sklaida per sistemas	Lokali, vienos sistemos lygmeniu	Plati: VM, hipervizorius, valdymo plokštės	(Cândido et al., 2019)

Virtuali infrastruktūra siūlo didesnį lankstumą, tačiau tuo pačiu reikalauja sudėtingesnės žurnalinių įrašų tvarkymo infrastruktūros – reikia ne tik daugiau įrankių, bet ir griežtesnių taisyklių – tiek techninių, tiek teisinių. Kaip pastebi Preethi D (2020) – be tinkamos apsaugos organizacija rizikuoja ne tik netekti duomenų, bet ir prarasti galimybę vykdyti veiklos auditą ar ginti savo poziciją teisminiuose procesuose, tuo tarpu fizinė infrastruktūra yra labiau nuspėjama ir stabili, tačiau ji riboja lankstumą ir gebėjimą greitai reaguoti į apkrovų pokyčius ar diegimo poreikius (Cândido et al., 2019).

Išvados

Šiame darbe buvo išnagrinėtas žurnalinių įrašų surinkimo, agregavimo ir analizės procesas virtualioje infrastruktūroje, lyginant jį su fizine infrastruktūra. Remiantis moksliniais tyrimais, galima daryti keletą esminių išvadų, kurios padeda suprasti šios srities iššūkius ir galimybes.

1. Virtualioje aplinkoje žurnalinių įrašų architektūra yra žymiai sudėtingesnė. Įrašai gaunami iš kelių sluoksnių – VM, hipervizorių, tinklų ir debesijos paslaugų. Dėl to sunku juos sinchronizuoti ir analizuoti, ypač kai trūksta standartizuotų formatų (Cândido et al., 2019; Preethi D, 2020).
2. Žurnalai virtualioje infrastruktūroje yra trumpalaikiai ir gali būti prarasti. Jeigu duomenys nėra perduodami į išorinę sistemą realiuoju laiku, jų gali nebelikti net prieš prasidedant analizei (Wang et al., 2025). Tai reiškia, kad būtina taikyti centralizuotus ir saugius perdavimo bei saugojimo sprendimus.
3. Žurnalinių įrašų analizė fizinėje infrastruktūroje yra paprastesnė, bet mažiau lanksti. Fizinės sistemos generuoja mažesnius duomenų kiekius, o jų analizė dažnai vykdoma vietoje. Tačiau tai riboja analizės greitį ir mastelio keitimą, palyginti su debesų sprendimais (Akbaş, 2024).
4. Automatizuota analizė ir dirbtinis intelektas tampa būtini. Rankinė analizė tampa neefektyvi esant dideliam įrašų kiekiui. Todėl tokios sistemos kaip „MoniLog“ ar „LogBabylon“ leidžia taikyti pažangias klasifikavimo priemones bei aptikti anomalias veiklas realiu laiku (Karanjai et al., 2024; Vervaet, 2021).
5. Saugumas ir teisinis atitikimas yra kritiniai veiksniai. Be įrašų vientisumo užtikrinimo ir prieigos kontrolės organizacija rizikuoja ne tik patirti duomenų nutekėjimą, bet ir būti nepasiruošusi audito ar teismo procesui (Akbaş, 2024; Preethi D, 2020).

Remiantis išvadomis, galima pateikti šias rekomendacijas:

1. Naudoti centralizuotas žurnalų kaupimo ir analizės sistemas. Rekomenduojama diegti sprendimus, paremtus ELK, Wazuh ar panašiais įrankiais, kurie leidžia apjungti įrašus iš kelių šaltinių ir juos analizuoti vieningoje aplinkoje (Cândido et al., 2019).
2. Užtikrinti duomenų vientisumą ir šifravimą. Reikėtų naudoti mechanizmus, tokius kaip „Proof-of-Past-Log“, viešojo rakto infrastruktūrą (PKI) ir šifravimą tiek perdavimo, tiek saugojimo metu (Preethi D, 2020).

3. Automizuoti analizę naudojant mašininį mokymąsi. Tokios priemonės leidžia sumažinti klaidingų pavojaus signalų kiekį ir greičiau reaguoti į realius incidentus (Karanjai et al., 2024; Vervaet, 2021).
4. Apgalvoti infrastruktūros saugojimo strategiją. Virtualioje aplinkoje būtina turėti papildomus saugyklos sprendimus, kurie nepriklauso nuo debesijos tiekėjo ir leidžia išlaikyti ilgalaikį duomenų prieinamumą (Akbaş, 2024).
5. Periodiškai tikrinti teisinį atitikimą. Būtina peržiūrėti, ar žurnalų kaupimas ir analizė atitinka BDAR, ISO/IEC 27001 ar PCI-DSS standartus, kad būtų išvengta rizikos organizacijos veiklai (Akbaş, 2024).

LITERATŪROS SĄRAŠAS

Akbaş, E. (2024). Evaluating SIEM RADAR: A New Metric for Enhancing Regulatory and Compliance Efficiency. *European Conference on Cyber Warfare and Security*, *23*(1), 18–26. https://doi.org/10.34190/ECCWS.23.1.2346

Cândido, J., Aniche, M., & Deursen, A. Van. (2019). Log-based software monitoring: a systematic mapping study. *PeerJ Computer Science*, *7*, 1–38. https://doi.org/10.7717/PEERJ-CS.489

Karanjai, R., Lu, Y., Alsagheer, D., Kasichainula, K., Xu, L., Shi, W., & Huang, S.-H. S. (2024). LogBabylon: A Unified Framework for Cross-Log File Integration and Analysis. *Proceedings of the 40th ACM/SIGAPP Symposium on Applied Computing*, 1953–1960. https://doi.org/10.1145/3672608.3707883

Preethi D. (2020). *CLOUD LOG REASSURING SOUNDNESS AND SECRECY THEME FOR CLOUD FORENSICS*. *7*, 77. www.jetir.org

Vervaet, A. (2021). MoniLog: An automated log-based anomaly detection system for cloud computing infrastructures. *Proceedings - International Conference on Data Engineering*, *2021-April*, 2739–2743. https://doi.org/10.1109/ICDE51399.2021.00317

Wang, K., Wu, S., Cui, Y., Huang, Z., Fan, H., & Jin, H. (2025). System log isolation for containers. *Frontiers of Computer Science*, *19*(5). https://doi.org/10.1007/S11704-024-2568-8