Opinnäytetyö (AMK)

Tieto- ja viestintätekniikka

2020

Antti Nousiainen

LOKIHALLINTAJÄRJESTELMÄN KÄYTTÖÖNOTTO

TURKU AMK
TURKU UNIVERSITY OF APPLIED SCIENCES

OPINNÄYTETYÖ (AMK) | TIIVISTELMÄ

TURUN AMMATTIKORKEAKOULU

Tieto- ja viestintätekniikka

2020 | 53 sivua

Antti Nousiainen

LOKIHALLINTAJÄRJESTELMÄN KÄYTTÖÖNOTTO

Opinnäytetyön tavoitteena oli etsiä yrityksen IT-ympäristöön tarvittavat kriteerit täyttävä keskitetty lokihallintajärjestelmä. Järjestelmä mahdollistaisi lokidatan keskitetyn tarkastelun sekä analysoinnin. Kerättyä dataa analysoimalla voitaisiin muodostaa lähes reaaliaikainen tilannekuva kohdeympäristön toiminnasta sekä ratkaista tehokkaasti ongelmia vikatilanteissa.

Työssä käytiin läpi lokihallinnan teoriaa sekä vertailtiin yhteensä kolmea tehtävään sopivaa kaupallista tuotetta niin lisenssikustannusten kuin teknisten ominaisuuksien osalta. Tuotteeksi valikoitui Nagios Enterprisesin Nagios Log Server sen monipuolisten ominaisuuksien ja edullisen lisenssikustannuksen perusteella.

Käyttöönoton tuloksena oli kattavasti kohdeympäristön palvelin- ja tietoverkkoinfrastruktuurin tilaa monitoroiva ja analysoiva järjestelmä joka parantaa kohdeympäristön tietoturvaa sekä mahdollistaa ajankohtaisen kokonaiskuvan muodostamisen.

ASIASANAT:

Lokihallinta, syslog, Nagios, lokipalvelin

BACHELOR´S THESIS | ABSTRACT

TURKU UNIVERSITY OF APPLIED SCIENCES

Information and Communications Technology

2020 | 53 pages

Antti Nousiainen

CENTRALIZED LOG MANAGEMENT IMPLEMENTATION

The purpose of this thesis was to find, implement and manage a centralized log management system based on set criteria. The system should be able to centrally aggregate log data for further analyzation. Logs received by the system should be analyzed to solve existing problems and proactively seek potential problems.

The thesis first reviews several key concepts related into logging and log management. For the practical part of the thesis, in total, three commercial products were evaluated by their technical properties and also by their license pricing. These commercial products were Graylog, Nagios Log Server and Splunk. The Nagios Enterprises' Nagios Log Server was chosen for its diverse set of features and for its affordable licensing cost.

The result of the thesis was a versatile centralized log management system, capable of real-time log aggregation and analyzation. In effect, the system's functionalities enable both greater security for its IT environment and a more coherent state of its infrastructure.

KEYWORDS:

Log management, syslog, Nagios, log server

SISÄLTÖ

KUVIOT

TAULUKOT

KÄYTETYT LYHENTEET

BSD	UNIX-tyyppinen käyttöjärjestelmä (Berkeley Software Distribution)
DNS	Nimipalvelujärjestelmä (Domain Name System)
FQDN	Täydellinen toimialuenimi (Fully Qualified Domain Name)
HTTPD	Web-palveluprosessi (Hypertext Transfer Protocol Daemon)
HTTPS	Tietoliikenneprotokolla (Hypertext Transfer Protocol Suite)
ICMP	Kontrolliprotokolla (Internet Control Message Protocol)
IETF	Internet-standardisointiorganisaatio (Internet Engineering Task Force)
IIS	Palvelinohjelmisto (Internet Information Services)
IPsec	Tietoliikenneprotokolla (IP Security Architecture)
JSON	Tiedostomuoto (JavaScript Object Notation)
LDAP	Verkkoprotokolla (Lightweight Directory Access Protocol)
MAC	Laiteosoite (Media Access Control)
MPLS	Tietoliikenneprotokolla (Multiprotocol Label Switching)
RHEL	Linux-käyttöjärjestelmä (Red Hat Enterprise Linux)
SIEM	Tietoturvahallintajärjestelmä (Security Information and Event Management)
SMB	Tietoliikenneprotokolla (Server Message Block)
SNMP	Tietoliikenneprotokolla (Simple Network Management Protocol)
SQL	Tietokantajärjestelmän kyselykieli (Structured Query Language)
SSH	Hallintaprotokolla (Secure Shell)
SSHD	Hallintaprotokollan ohjelma (Secure Shell Daemon)
TCP	Tietoliikenneprotokolla (Transmission Control Protocol)
TLS	Salausprotokolla (Transport Layer Security)
UDP	Tietoliikenneprotokolla (User Datagram Protocol)
UNIX	Käyttöjärjestelmä (Uniplexed Information and Computing System)

VPN	Virtuaalinen erillisverkko (Virtual Private Network)
XML	Merkintäkieli (Extensible Markup Language)

1 JOHDANTO

Tietotekniikan ja tietoturvan merkitys yrityksissä on korostunut viime vuosien aikana merkittävästi. IT-järjestelmien toimintavarmuus ja saatavuus ulottuu käytännössä kaikkiin yrityksen liiketoimintojen osa-alueisiin. Kasvanut laite- ja palvelumäärä vaatii taustalle monitorointi- ja hallintajärjestelmiä, joilla voidaan mahdollistaa palveluiden jatkuvuus, ja säilyttämällä tietoturva riittävän korkealla tasolla.

Opinnäytetyön tarkoituksena oli ottaa käyttöön keskitetty lokihallintajärjestelmä toimeksiantajayrityksen tietoverkkoympäristössä, jossa tällaista ei aiemmin ollut. Kohdeympäristön tietohallinnon asiantuntijat tarvitsivat järjestelmän, jonka avulla voidaan tarkastella erinäisten verkkolaitteiden ja palvelimien lokitietoja kootusti. Kerättyä lokidataa analysoimalla voitaisiin saada käytännössä reaaliaikaista kokonaiskuvaa verkko- ja palvelininfrastruktuurin tilasta sekä ratkaista piileviä ongelmia proaktiivisesti ennen kuin ne ehtivät kunnolla muodostua.

Ensimmäisenä esimerkkinä lokihallintajärjestelmän tarpeellisuudesta olisi yhtäkkiä ilmennyt laitetason vikatila kriittisessä verkkolaitteessa. Tämäntyyppinen vika voi olla vaikea paikantaa ja selvittää ilman kunnollista monitorointia. Mutta jos laite ilmoittaa lokihallintajärjestelmälle jo etukäteen ilmenevistä virheistä lokidatana, niin asiaan voitaisiin reagoida riittävän nopeasti.

Toisena esimerkkinä olisi sisäverkkoon päässyt haittaohjelma, jota virustorjuntaohjelmat eivät havaitse. Jos haittaohjelma esimerkiksi yrittäisi kirjautua muihin laitteisiin epäonnistuneesti tarpeeksi monta kertaa, niin tästä tulisi lokihallintajärjestelmään hälytys. On hyvin pieni todennäköisyys, että tällainen murtoyritys edes huomattaisiin ilman lokihallintajärjestelmää, sillä sen huomaaminen vaatisi juuri oikean laitteen lokidatan tarkastelua tapahtumahetkellä tai pian sen jälkeen.

Ennen opinnäytetyön tekemistä ei kohdeympäristössä ollut minkäänlaista lokihallintajärjestelmää, eli vikatilanteessa tuli etsiä manuaalisesti relevanttia lokidataa suuresta määrästä laitteita. Tähän ratkaisuna päätettiin ottaa käyttöön kaupallinen tuote, joka mahdollistaa lokidatan keskittämisen ja analysoinnin.

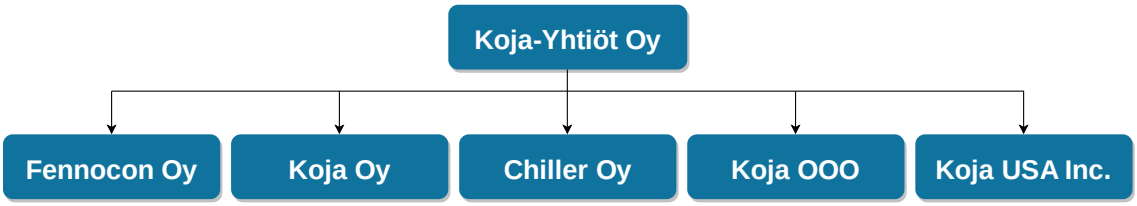
Työn tavoitteena on etsiä ja implementoida kohdeympäristön käyttötarkoitukseen sopiva lokihallintajärjestelmä joka täyttää ominaisuuksiltaan tarvittavat kriteerit. Tärkeimmät työhön liittyvät kysymykset ovat seuraavat.

- Mitä vaatimuksia ohjelmistolla on?
- Mikä kaupallinen tuote on ominaisuuksiltaan sopivin?
- Miten lokihallintajärjestelmän käyttöönottoprosessissa edetään?

Työssä käydään läpi lokihallinnan teoriaa, vertaillaan tehtävään sopivia kaupallisten tuotteiden teknisiä ominaisuuksia, sekä käydään läpi käyttöönottoprojekti vaiheittain.

2 KOHDEYMPÄRISTÖ

Työn toimeksiantajana toimii Tampereella 1930-luvulla perustettu perheyritys Koja-Yhtiöt Oy, joka valmistaa ilmastointijärjestelmiä rakennuksiin ja laivoihin sekä prosessipuhaltimia teollisuuden tarpeisiin. Yrityksellä on toimipisteitä Suomessa, Ruotsissa, Norjassa, Venäjällä ja Yhdysvalloissa. Yrityksen rakenne koostuu emoyhtiönä toimivasta Koja-Yhtiöt Oy -konsernista (kuvio 1) jonka alaisuudessa liiketoiminnot ovat omissa yrityksissään (Koja 2020).



Kuvio 1. Kojan organisaatiorakenne (Koja 2020 pohjalta).

Kojan verkkoinfrastruktuuri koostuu neljästä eri toimipaikoilla sijaitsevista palomuureista ja noin kolmestakymmenestä verkkokytkimestä. Yhteydet päätoimipaikalta (Tampere) muihin toimipisteisiin ja asiakaskohteisiin mahdollistavat MPLS -tekniikka (Multiprotocol Label Switching) ja IPsec-yhteydet (IP Security Architecture). Kesällä 2019 Kojalla on käytössä klusteroitu VMwaren vSphere 6.5 -virtualisointialusta, jossa toimii yhteensä noin 40 kappaletta Windows- ja Linux-palvelimia jotka ovat yrityksen liiketoimintojen osalta kriittisessä roolissa.

3 LOKI

“Loki tarkoittaa aikajärjestyksessä kirjattua tallennetta tapahtumista ja niiden aiheuttajista. Tapahtumat ja muutokset tietojärjestelmissä, sovelluksissa, tietoverkoissa ja tietosisällöissä kirjataan lokiin, eli lokitetaan.” (Näin keräät ja käytät lokitietoja 2019)

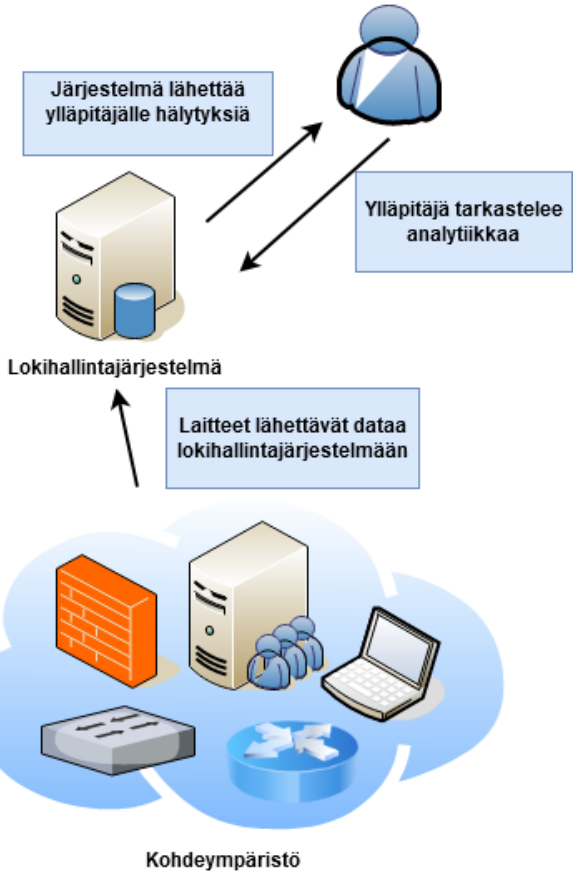
Alun perin lokeilla pystyttiin tutkimaan yksittäisissä laitteissa ilmeneviä ongelmia, mutta vuosien saatossa siitä on tullut keskeinen ympäristöjen teknisen kokonaisturvallisuuden arvioinnin työkalu. Kirjautumisyritykset, tiedostojen avaamiset ja muutokset käyttäjätileissä tallentuvat laitteen lokiohjelmaan, josta ne voidaan lähettää eteenpäin keskistetysti analysoitavaksi (Effective Daily Log Monitoring 2016).

Lokit ovat yksi tärkeimmistä välineistä IT-ympäristön toiminnan seuraamiseen ja vikatilanteissa ongelmien selvittämiseen. Laite tai sovellus tallentaa dataa muista tiedostoista erilliseen sijaintiin ja tyypillisesti lokitiedoston kirjaukset sisältävät seuraavia tietoja (Näin keräät ja käytät lokitietoja 2019).

1. Aikaleima
2. Tapahtuma ja toimija
3. Käyttöoikeus
4. Tapahtuman lähde
5. Tapahtuman tila

3.1 Lokihallintajärjestelmä

Lokihallintajärjestelmä on ohjelmisto joka mahdollistaa lokidatan keräämisen, varastoinnin ja analysoinnin useista eri lähteistä. Järjestelmän etu on se, että kaikki relevantti data on keskitettynä yhteen paikkaan josta sitä voidaan tarkastella kootusti. (SIEM vs. Log Management).



Kuvio 2. Lokihallintajärjestelmän toimintaperiaate.

Keskittämismetodilla (kuvio 2) pystytään suurissakin ympäristöissä seuraamaan kokonaistilannetta ja arvioimaan teknisen tietoturvallisuuden onnistumista. Verkkoympäristön ylläpitäjä pystyy tarkastelemaan koko ympäristön tilaa yhden järjestelmän kautta.

3.2 SIEM-järjestelmä

SIEM-järjestelmä (Security Information and Event Management) eroaa hiukan lokihallintajärjestelmästä siten, että se sisältää lokihallintajärjestelmien ominaisuuksien lisäksi myös edistyneempiä tietoturvaan liittyvä työkaluja.

- hälytykset
- tapahtumien korrelointi
- näkymät.

SIEM-järjestelmä kykenee havaitsemaan toistuvaa epäilyttävää toimintaa ja tekemään siitä hälytyksen järjestelmän ylläpitäjälle. Esimerkiksi useat epäonnistuneet kirjautumisyritykset, suurien datamäärien siirtäminen tai verkkojen skannaaminen.

Järjestelmä voi käyttää koneoppimista tai erilaisia algoritmeja tunnistaakseen kahden eri tapahtuman liittyvän toisiinsa. Esimerkiksi porttiskannaus ja tämän jälkeen alkaneet epäonnistuneet kirjautumiset voivat olla yhteydessä toisiinsa.

Useat SIEM-tuotteet tarjoavat lisäksi valvomotyyppisiä näkymäsivuja (dashboard) josta voidaan samanaikaisesti tarkastella useita eri tietoja ja siten muodostetaan ympäristön tilannekuva. (SIEM vs. Log Management)

Nykyään lokihallinta- ja SIEM-järjestelmien ero on häilyvä sillä useita SIEM-ominaisuuksilla varustettuja järjestelmiä kutsutaan ja markkinoidaan lokihallintajärjestelminä.

3.3 Lokien hallinta

Lokien hallinnalla tarkoitetaan esimerkiksi lokitietojen keräämistä, jäsentämistä, varastointia, arkistointia ja analysointia. Lokien hallinnassa käytetään tyypillisesti jonkinlaista keskitettyä järjestelmää, jonne kerätty lokidata varastoidaan analysointia varten.

Pelkkä keskitetyn lokihallintajärjestelmän käyttöönottaminen IT-ympäristössä ei välttämättä tuo lisäarvoa sinällään, koska useimmissa tapauksissa analysointi ja raportointi vaativat räätälöintiä kohdeympäristön tarpeiden mukaan. “Toisessa ympäristössä tietty tapahtumasarja voi näyttää hyökkäykseltä, kun taas toisessa se on osa normaalia tapahtumankulkua.” (Näin keräät ja käytät lokitietoja 2019)

Lokien hallinnan eri prosessit voidaan jakaa karkeasti kuuteen eri tyyppiin jotka käymme seuraavaksi läpi.

3.3.1 Lokidatan kerääminen.

Ensimmäinen huomioitava asia lokien keräämisessä on verkon topologia. Lokipalvelimen käyttöönotossa täytyy huomioida monenlaisista järjestelmistä (kuten esimerkiksi palomuurit, palvelimet, verkkokytkimet, tietojärjestelmät, pilvipalvelimet) kerättävä data

ja sen vaatimat verkkoyhteydet. Datan kerääminen ei ole mahdollista jos tietyistä sisäisistä verkoista ei ole pääsyä lokipalvelimen verkkoon, täten eri verkkojen välisiin sääntöihin ja käytäntöihin tulee sallia tiettyjä yhteyksiä.

Toinen huomioitava asia on lokipalvelimen tallennustila, sillä esimerkiksi tietokantapalvelimet voivat parhaimmillaan lähettää kymmeniä lokitapahtumia sekunnissa. Suuren lokimäärän saapuminen ja käsittely vaativat verkolta ja lokipalvelimelta riittävää suorituskykyä (Carstensen 2019).

3.3.2 Lokidatan keskittäminen

Lokidatan keskittämisellä tarkoitetaan kaiken mahdollisen lokidatan keskittämistä yhteen järjestelmään, mutta ei välttämättä pelkästään yhteen laitteeseen sillä useat kaupalliset lokijärjestelmät tarjoavat usean laitteen vikasietoisuutta.

3.3.3 Säilyttäminen ja arkistointi

Yksi merkittävä kysymys lokien hallinnan kannalta on se että kuinka kauan lokeja tulisi säilyttää. Tämä vaihtelee huomattavasti ympäristöstä ja datan tyypistä riippuen, mutta IT-alan yleisten käytäntöjen mukaan dataa olisi hyvä säilyttää kuudesta kuukaudesta pariin vuoteen saakka (Näin keräät ja käytät lokitietoja 2019).

Arkistoinnilla on myös käytännöllinen merkitys; suuren lokidatamäärän säilyttäminen pitkään lokijärjestelmässä voi aiheuttaa laitteen suorituskyvyn kannalta ongelmia. Tämän vuoksi lokidataa olisi järkevää siirtää arkistoitavaksi jollekin toiselle alustalle, kuten esimerkiksi varmuuskopiointijärjestelmään.

3.3.4 Lokikierto

“Lokikierrolla tarkoitetaan lokitiedoston sulkemista ja uuden avaamista, kun lokitiedoston katsotaan olevan täynnä. Lokikierto voidaan ajoittaa tietyn aikataulun (tunneittain, päivittäin, viikoittain jne.) mukaan tai suorittaa, kun lokitiedosto saavuttaa ennalta määritellyn koon” (Lokien säilytys, kerääminen ja suojaaminen 2009).

Lokikierto auttaa myös aiemmin mainituissa suorituskykyongelmissa automaattisesti sulkemalla lokitiedostoja. Tämä vähentää järjestelmän kuormitusta ja siirtää vanhemman datan valmiiksi arkistointia tai poistoa varten.

3.3.5 Lokianalyysi

Lokin analysointi on järjestelmän hyödyllisyyden kannalta yksi merkittävimmistä ominaisuuksista. Pelkkä lokidatan kerääminen ei välttämättä tuo itsessään kovin suurta arvoa, mutta sen analysointi mahdollistaa reaaliaikaisen tilannekuvan muodostamisen ja paremman tietoturvallisuuden. Suurin osa lokihallintajärjestelmistä tarjoaa datan visualisointia, raportointia ja eri tietojen vertailua. Nämä työkalut helpottavat piilevien ongelmien havainnoimista ja toimivat ennaltaehkäisevästi. Myös erilaisilla hälytyksillä voidaan ilmoittaa nopeasti ilmenevistä ja toistuvista tapahtumista, esimerkiksi jos järjestelmään tai käyttäjätiliin yritetään murtautua.

3.4 Lokien tyypit

Keskeisimpinä lokityyppeinä voidaan pitää Linux-järjestelmien sekä verkkolaitteiden käyttämää syslogia, sekä Windows-käyttöjärjestelmän tapahtumalokia. Luvuissa 4 ja 6 käydään läpi syslogin ja Windowsin tapahtumalokin tekniset ominaisuudet ja toimintaperiaatteet.

4 SYSLOG-LOKI

Eric Allman kehitti Syslogin 1980-luvulla työskennellessään Sendmail-ohjelmiston parissa, joka oli yksi ensimmäisistä sähköpostiohjelmista. Tämän jälkeen muutkin UNIX-pohjaiset (Uniplexed Information and Computing System) ohjelmat alkoivat käyttämään syslogia sovellusdatan tallentamiseen (Allman 2014).

Ajan mittaan syslogista alkoi muodostua UNIX-alustoilla (Uniplexed Information and Computing System) yleinen käytäntö lokidatan tallentamiseen ja vuonna 2001 IETF (Internet Engineering Task Force) julkaisi ensimmäisen standardin jossa määriteltiin syslogin tarkat spesifikaatiot, joskaan ne eivät olleet virallisia sääntöjä. Vaikkakin syslog oli kehitetty BSD-alustalle (Berkeley Software Distribution) niin sen hyödyllisyyden takia se käännettiin muillekin käyttöjärjestelmille ja verkkolaitteille (RFC 3164 2001). Kahdeksan vuotta myöhemmin julkaistiin päivitetty standardi RFC 5424 jossa esiteltiin mahdollisuus syslog-viestien salaamiseen käyttäen TLS-protokollaa (Transport Layer Security) (RFC 5424 2009).

Yksi syslogin ongelmista edelleen on se, että vaikkakin standardissa on määritelty syslog-viestin sisältö ja syntaksi, niin kaikki laitevalmistajat eivät tätä noudata ja tämä luo päänvaivaa lokihallintajärjestelmien ylläpitäjille.

4.1 Teknisiä ominaisuuksia

Syslog toimii tyypillisesti portissa 514 käyttäen UDP-protokollaa (User Datagram Protocol) tiedonsiirtoon. UDP-protokolla on luonteeltaan datagram-pohjainen ja perustuu yksittäisten pakettien vaihtoon ilman virheentarkistusta, pakettien järjestyksen tarkastamista tai kaksoiskappaleiden havaitsemista. Myöskään pakettien perillepääsemistä ei varmisteta (RFC 768 1980).

TCP Segment Header Format

Bit #	0 – 7	8 – 15	16 – 23	24 – 31
0	Source Port		Destination Port	
32	Sequence Number			
64	Acknowledgment Number			
96	Data Offset	Res	Flags	Window Size
128	Header and Data Checksum		Urgent Pointer	
160…	Options			

UDP Datagram Header Format

Bit #	0 – 7	8 – 15	16 – 23	24 – 31
0	Source Port		Destination Port	
32	Length		Header and Data Checksum	

Kuvio 3. TCP- ja UDP-protokollien header-kentät (Microchip 2020).

Varmistusmekanismien puuttumisen takia UDP on otsikkotasolla (header) huomattavasti yksinkertaisempi (kuvio 3) kuin TCP-protokolla (Transmission Control Protocol).

4.2 Viestin rakenne

Syslog-paketin koko on minimissään 480 tavua ja maksimina voidaan pitää 2048 tavua, mutta lokin vastaanottaja voi myös vastaanottaa maksimikokoa suuremman paketin. Tällöin vastaanottavan järjestelmän pitäisi lyhentää saapuvaa pakettia tai pudottaa se kokonaan (RFC 5425 2009).

Syslog-paketti koostuu yhteensä viidestä arvosta.

- facility
- severity
- hostname
- timestamp
- message.

4.2.1 Facility-arvo

Facility-arvon avulla pystytään kategorisoimaan saapuvat syslog-viestit niiden lähteen perusteella. Nämä lähteet voivat olla esimerkiksi käyttöjärjestelmä, sovellus tai prosessi. Facility-arvolle voidaan antaa arvoja välillä 0-23 ja viimeiset kahdeksan arvoa ovat varattuja sellaisille prosesseille, jotka eivät sovi ennalta-annetuihin arvoihin. Näitä kahdeksaa arvoa voivat käyttää esimerkiksi erilaiset sovellukset (Deveriya 2005).

4.2.2 Severity-arvo

Lokitietoa lähettävä prosessi, ohjelma tai käyttöjärjestelmä määrittää lokiviestiin severity-arvon. Arvo kuvaa kuinka tärkeä tai kriittinen saapuva viesti on ja näitä kuvataan arvoilla 0-7. (Deveriya 2005)

Taulukko 1 Severity-arvot.

Arvo	**Severity**
0	Emergency: System is unusable
1	Alert: Action must be taken immediately
2	Critical: Critical conditions
3	Error: Error conditions
4	Warning: Warning conditions
5	Notice: Normal but significant condition
6	Information: Information messages
7	Debug: Debug-level messages

Taulukossa 1 on esitettynä eri severity-arvot. Matalin arvo on kaikista kriittisin (Emergency) ja suurin kaikista vähäpätöisin (Debug). Lokiviestejä kerätessä on syytä asettaa severity-arvo riittävän matalaksi, jotta viestejä ei tule liikaa.

4.2.3 Hostname-kenttä

Hostname-kenttä sisältää lokitietoa lähettävän laitteen FQDN-nimen (Fully Qualified Domain Name), staattisen IP-osoitteen, paikallisen isäntänimen tai NILVALUE-arvon. Jälkimmäistä arvoa käytetään jos syslog-applikaatiolla ei ole pääsyä mihinkään muuhun tietoon (RFC 5424 2009).

4.2.4 Timestamp-kenttä

Syslog käyttää ISO 8601 -standardin mukaista päivämäärän ja ajan esittämistapaa, joskin uusimmassa syslogia standardisoivassa julkaisussa (RFC 5424) sille annetaan uusia ehtoja. Kellonaikaa ja aikavyöhykettä päivämäärästä erottavat T- ja Z-merkit tulee kirjoittaa isolla, sekä T-merkki on pakollinen. Lisäksi karkausekunteja ei tule käyttää (RFC 5424 2009).

```
2020-03-14T18:05:20+0200
```

Kuvio 4. Esimerkki aikaleimasta.

Kuviossa 4 esitetään aikaleima ISO 8601 -standardin mukaisesti, jolloin T-merkillä erotetaan päiväys kellonajasta. Aikaleima sisältää lopussa UTC-aikavyöhykkeen erotettuna kellonajasta plusmerkillä.

4.2.5 Message-kenttä

Message-osio sijaitsee syslog-viestin lopussa ja sisältää laitteen facilityn kertoman viestin. Viestillä ei ole mitään tarkkaa formaattia ja sen olisi hyvä tarjota jotain lisäinformaatiota tapahtumasta.

```
Mar 14 18:23:35 CORE systemd[22646]: Startup finished in 365ms.
```

Kuvio 5. Esimerkki syslog-viestin message-kentästä.

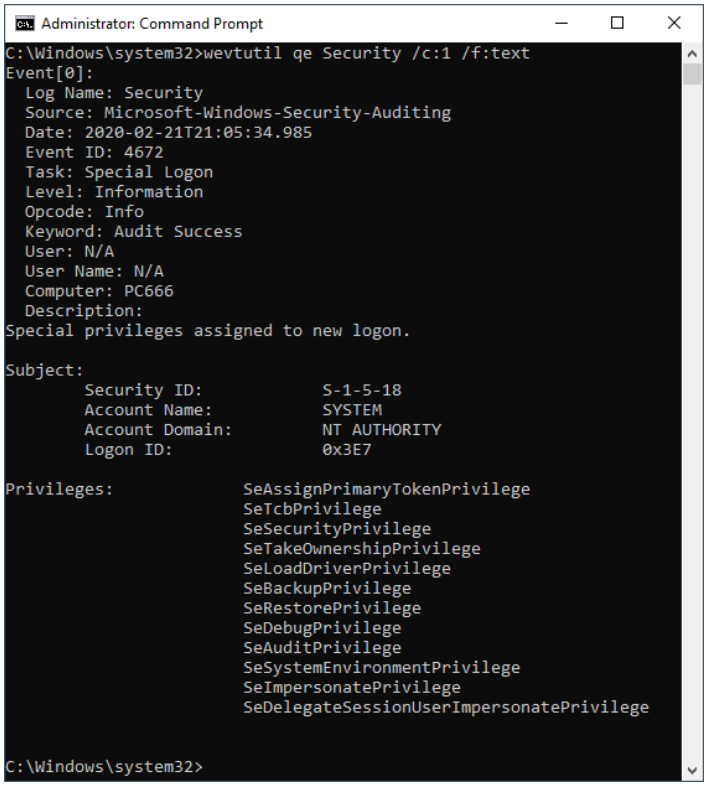
Kuviossa 5 esitetään Linux-palvelimen systemd-prosessin ilmoittama käyttöjärjestelmän käynnistymisen kesto millisekunteina. Viesti antaa tapahtumasta lisätietoa lokin tarkastelijalle.

5 WINDOWS EVENT LOG -LOKI

Vuonna 2008 tapahtuneesta Windows Vistan ja Windows Server 2008 julkaisusta lähtien Microsoft on käyttänyt Windows-järjestelmissään lokien tallentamiseen XML-muodossa (Extensible Markup Language) olevaa EVTX-tiedostopäätettä. Aiemmissa Windows-järjestelmissä oli käytössä EVT-tiedostopääte josta puuttui esimerkiksi kehittyneempiä tapahtumien (event) ominaisuuksia, XML-formaatti, lokien jäsentely kahteen kategoriaan ja se sisälsi vanhan Event Viewer-ohjelman (Charter, B. 2008).

5.1 Yleistä

Loki jaetaan kahteen kategoriaan: Windows-loki sekä sovellukset ja palvelut -loki. Koska lokidata on XML-muodossa, sen lukemiseen tulee käyttää Windows-käyttöjärjestelmästä löytyvää graafista Event Viewer -ohjelmaa tai vaihtoehtoisesti komentolinjapohjaista wevtutil-ohjelmaa. Käytössä on muitakin eri ohjelmia, mutta tyypillisesti Windows-ympäristössä käytetään Event Viewer -ohjelmaa (Hicks, J. 2012).



Kuvio 6. SYSTEM-tunnuksen kirjautumisen lokitieto wevtutil-ohjelmassa esitettynä tekstiformaatissa.

5.1.1 Windows-loki

Windows-loki-kategoriasta löytyy yhteensä viisi eri lokityyppiä. Application-loki sisältää laitteella toimivien ohjelmistojen lokitietoja. Security-lokissa näkyy käyttäjätilien onnistuneet ja epäonnistuneet kirjautumiset sekä korotetut käyttäjäoikeudet. Setup-loki on Windows-käyttöjärjestelmän asennuksen loki. System on itse Windows-käyttöjärjestelmän loki. Forwarded Events -loki sisältää muista laitteista lähetetyt lokit (Ultimate Guide to Logging 2020). Kuviossa 6 nähdään esimerkki Security-lokiin kirjatusta kirjautumistapahtumasta.

5.1.2 Sovellukset ja palvelut -loki

Applikaatiot ja palvelut -loki koostuu admin- ja operational-lokeista. Admin-loki eroaa muista siten että se tarjoaa yksityiskohtaisempaa tietoa tapahtumista kuin muut lokit ja se ehdottaa korjaavia toimenpiteitä. Esimerkiksi jos Active Directory -toimialueen palvelimien välinen replikointi ei toimi, niin loki saattaa ehdottaa palomuurisääntöjen tarkastelua. Operational-loki seuraa tiettyjen tapahtumien ilmenemisiä kuten esimerkiksi uusien levyjen havaitsemista (Palmer, M. 2008).

5.2 Jaottelu

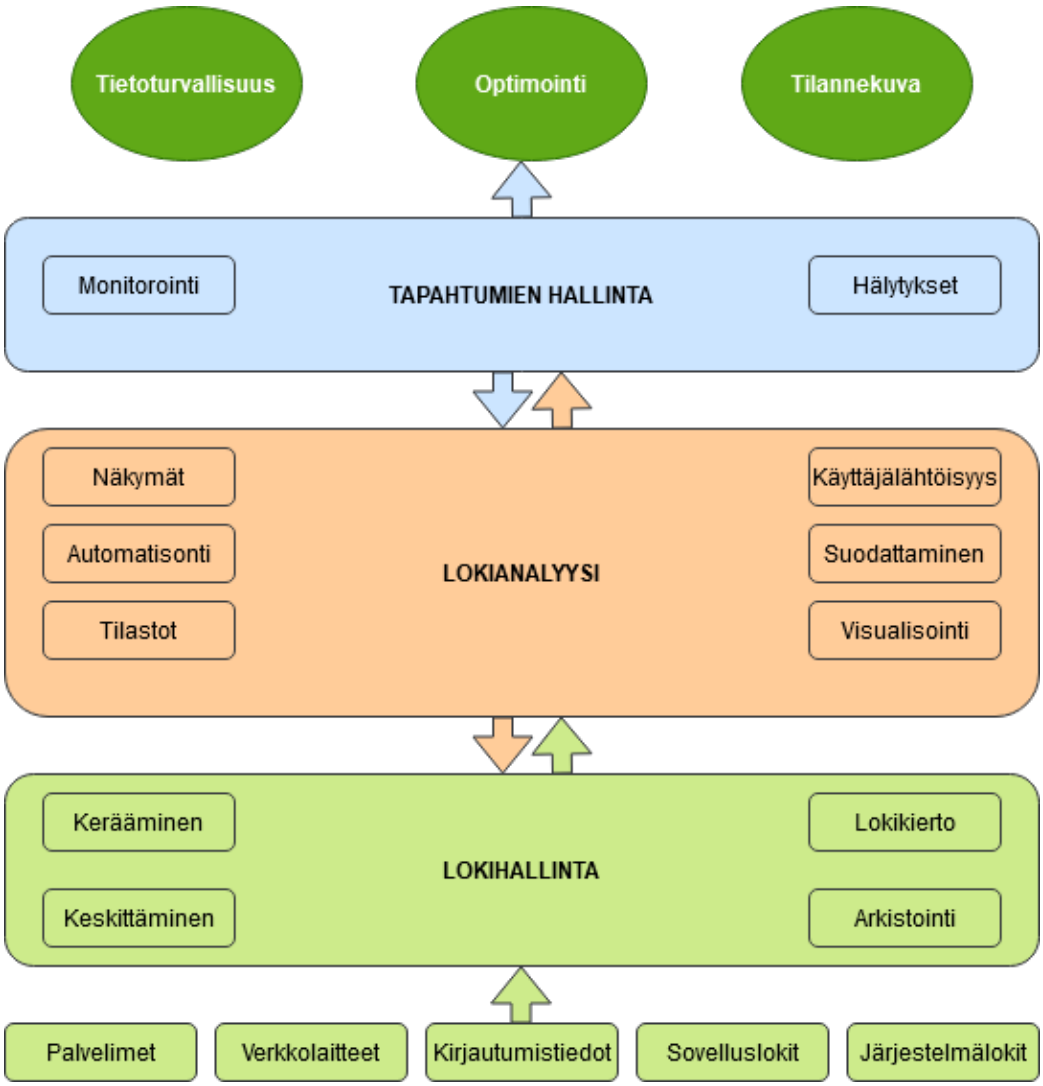
Jokainen Windowsin tapahtumanlokin tapahtuma (event) jaotellaan sen kriittisyyden perusteella yhteensä viiteen eri tyyppiin.

Taulukko 2. Event login eri tyypit (Event Types 2018).

Arvo	**Kuvaus**
1	Error: loss of data or loss of functionality
2	Warning: not necessarily significant, but may indicate a possible future problem
3	Information: describes the successful operation of an application, driver, or service
4	Success Audit: an audited security access attempt that is successful
5	Failure Audit: an audited security access attempt that fails

6 VAATIMUSMÄÄRITTELY

Vaatimusmäärittelyn tarkoituksena on kuvata soveltuvan järjestelmän tavoitteita ja teknisten ominaisuuksien vaatimuksia. Tarkoituksena on löytää mahdollisimman hyvin eri kriteereihin sopiva kaupallinen tuote. Kuviossa 7 on esitetty keskeisimpiä lokihallintajärjestelmän toiminnallisuuksia ja ominaisuuksia.



Kuvio 7. Lokihallintajärjestelmän toiminnot (Agrawal ym. 2015).

6.1 Lokidatan kerääminen

Järjestelmän tulisi tukea aiemmin läpikäytyjä (kuvio 7) lokihallinnan keskeisimpiä toimintatapoja: keräämistä, keskittämistä, arkistointia, lokien kiertoa sekä niiden analysoimista. On myös ehdottoman tärkeää että järjestelmä tukisi kaikkia kohdeympäristön laitteita, joista lokidataa halutaan kerätä.

Kohdeympäristön keskeisimpiä laitteita ovat esimerkiksi Windows -ja Linux-palvelimet, verkkokytkimet, palomuurit ja reitittävät kytkimet. Kaikki näistä laitteista eivät välttämättä tue tai käytä syslog-protokollaa, joten tulisi olla myös mahdollisuus käyttää vaihtoehtoista lokien kuljetusmetodia, esimerkiksi SNMP:tä (Simple Network Management Protocol).

Järjestelmän tulisi tukea hälytyksiä, eli jos havaitaan esimerkiksi poikkeuksellisen suuri määrä epäonnistuneita kirjautumisyrityksiä, niin tästä lähtisi ilmoitus kohdeympäristön ylläpidolle.

6.2 Analysointi

Yksi tärkeimmistä ominaisuuksista olisi kerätyn lokidatan analysointi ja visualisointi. Järjestelmän täytyy tukea lokidatan suodatusta erilaisin hakumetodein, sekä tuottaa helppolukuisia kaavioita ja diagrammeja joista dataa voidaan helpommin tulkita. Lokihallintajärjestelmän keskeisimmät hyödyt olisivat tietoturvallisuuden parantuminen ja ajankohtaisen ympäristön tilannekuvan luominen.

Tiettyjen lokiviestien etsimisen tulisi olla helppoa ja olisi hyvä että aiemmin luotuja hakuja ja näkymiä voisi tallentaa myöhempää käyttöä varten. Useimmissa lokihallintajärjestelmissä on käytössä dashboard-tyyppisiä näkymäratkaisuja.

6.3 Toiminnan jatkuvuus

Järjestelmän toiminnan jatkuvuus ja elinkaari ovat vaatimusmäärittelyssä huomioon otettavia asioita. On tärkeää että järjestelmä olisi suhteellisen vakaa, saisi säännöllisesti päivityksiä eikä kuormittaisi kohdeympäristön IT-osastoa jatkuvilla ylläpitotöillä. Lisäksi järjestelmän käyttöliittymän tulisi olla järkevästi suunniteltu ja sen käyttö intuitiivista, jotta peruskäyttö ei vaatisi kovin pitkää perehdytystä.

6.4 Teknisiä ominaisuuksia

Tehtävään soveltuvan järjestelmän tulisi tukea virtualisointia, sillä se tultaisiin ottamaan käyttöön VMware vSphere -virtualisointialustalla. Pohjana olevan käyttöjärjestelmän tulisi olla joko Windows Server 2016 tai Linux-puolen CentOS- tai Ubuntu-palvelinkäyttöjärjestelmä.

Järjestelmän tulisi olla suorituskyvyltään skaalautuva, jotta sille voitaisiin tarpeen mukaan antaa lisää resursseja, jotta lokien prosessoinnin osalta ei pullonkaulaa muodostuisi.

7 TUOTEKATSAUS

Lokihallintajärjestelmäksi arvioitiin kolmea eri kaupallista tuotetta. Nämä ovat Graylog, Nagios Log Server ja Splunk. Jokainen järjestelmistä sisältää haluttuja ominaisuuksia ja seuraavaksi käymme ne läpi.

7.1 Graylog

Graylog on vuonna 2009 perustettu avoimen lähdekoodin projekti josta on sittemmin tullut kaupallinen tuote. Se on yritystason ratkaisu joka mahdollistaa lokien keskitetyn keräämisen, varastoinnin ja analysoinnin reaaliajassa. Tuotteen kilpailukyky perustuu skaalautuvuuteen, käyttäjälähtöisyyteen sekä suureen määrään edistyneitä ominaisuuksia (About Graylog 2020).

Graylog on pääasiallisesti lokipalvelin mutta se sisältää myös SIEM-tuotteille tyypillisiä ominaisuuksia kuten Correlation Engine -ratkaisun joka havaitsee toistuvia samankaltaisia tapahtumia, kuten esimerkiksi tietoverkon skannausyrityksiä. Analytiikan osalta Graylog tarjoaa aikataulutettua raportointia, sekä mahdollisuuden integroida Graylogin data kolmannen osapuolen järjestelmiin käyttämällä ohjelmointirajapintaa (Features 2020).

Muita ominaisuuksia ovat myös seuraavat.

- tuki usealle lokidatatyypille
- nopeat ja tehokkaat hakumetodit
- edistyneet dashboard-ominaisuudet
- järjestelmän käyttäjien toiminnan auditointi
- roolipohjainen pääsynhallinta
- vikasietoisuus. (Features 2020)

Graylog käyttää vuosittaista lisenssimaksua joka perustuu päivittäin kerättävien lokitietojen määrään. Jos lokidataa kerätään vuorokaudessa alle viisi gigatavua, niin lisenssi on ilmainen (Graylog 2020). Graylogin lisenssihinnasto on esitettynä taulukossa 3.

Taulukko 3. Graylogin hinnasto.

Päivittäinen data-määrä	**Hinta vuodessa (USD)**
5 GB	$5400
10 GB	$7500
20 GB	$12000
50 GB	$20000

7.2 Nagios Log Server

Nagios Log Server on vuonna 2014 Nagios Enterprisesin toimesta julkaistu lokihallinta-järjestelmä. Järjestelmä sisältää runsaasti lokipalvelimelle tyypillisiä ominaisuuksia ja on käyttöliittymältään helppokäyttöinen (Nagios Log Server 2020).

Nagios sisältää lisenssikustannuksiinsa (taulukko 4) suhteutettuna useita edistyneitä ominaisuuksia joita ei tyypillisesti tämän hintatason tuotteesta löytyisi.

- usean instanssin vikasietoisuus
- konfiguroitavat hälytykset
- yksinkertaistettu uusien lokilähteiden lisääminen
- integrointimahdollisuus kolmannen osapuolen järjestelmiin
- reaaliaikainen lokidatan monitorointi. (Nagios Log Server 2020)

Nagios Log Serverin hinnoittelu on instanssipohjainen, eli mitä useampi lokipalvelin klusterissa on, niin sen enemmän se kustantaa. Lisenssimaksu on myös kertaluontoinen eikä sitä tarvitse uusia. Nagios Log Server tarjoaa myös ilmaisversiota, mutta siinä lokidatan määrä on rajoitettu puoleen gigatavuun päivässä joka ei välttämättä yritysympäristöihin riitä (Nagios Log Server 2020).

Taulukko 4. Nagios Log Serverin hinnasto.

Instanssien lukumäärä	**Hinta (USD)**
1	$3995
2	$4995
3	$5995
4	$6995
10	$14995

7.3 Splunk Enterprise

Splunk Enterprise on järjestelmä joka mahdollistaa lokidatan keräämisen, etsimisen, analysoinnin ja visualisoinnin. Splunk pystyy keräämään dataa verkkosivuilta, sovelluksista sensoreista ja erilaisista laitteista. Muihin lokihallintajärjestelmiin verrattuna Splunk Enterprise sisältää monia edistyneempiä ominaisuuksia esimerkiksi raportoinnin ja analytiikan osalta (Splunk Enterprise 2020).

Edistyneiden korrelaatio-ominaisuuksiensa takia Splunkia voidaan pitää enemmänkin SIEM-tuotteena kuin lokihallintajärjestelmänä. Splunk mahdollistaa yhteyksien löytämisen eri tapahtumien välillä jotka voivat viitata esimerkiksi tietomurtoon. Tyypillisesti lokihallintajärjestelmistä ei tällaisia ominaisuuksia löydy.

Splunk käyttää kerätyn lokidatan analysoimisessa koneoppimista samankaltaisuuksien ja korrelaatioiden löytämiseen. Splunkin tekoäly luo myös statistiikoista ennusteita perustuen ennalta opittuun tietoon (Splunk Enterprise 2020).

Splunk tukee aiemmin mainittujen lokihallintajärjestelmien tapaan seuraavia ominaisuuksia.

- indeksointi
- haku, metriikat ja trendien ennustaminen
- hälytykset
- konfiguroitavat näkymät
- raportit ja pivot-taulukot
- tiedon mallintaminen. (Splunk Enterprise 2020)

Splunk Enterprisen lisensointi perustuu järjestelmään saapuvan lokidatan määriin. Lisenssiä on saatavilla vuosittaisena tai kertaluontoisena lisenssinä. Kolmesta vertailussa olleesta järjestelmästä Splunk on huomattavasti kallein. Jos Splunkin kustannuksia verrataan esimerkiksi Graylogin vastaavaan vuosittaiseen lisenssimaksuun (taulukko 4) niin ero on merkittävä.

Taulukko 5. Spunk Enterprisen hinnasto.

Päivittäinen data-määrä	**Vuoden lisenssi**	**Jatkuva lisenssi**
1 GB	$1800	$5400
2 GB	$3000	$9000
5 GB	$6000	$18000
10 GB	$10000	$30000
20 GB	$15200	$54000
50 GB	$38000	$114000
100 GB	$60000	$180000

7.4 Tuotevalinta

Tuotevalintaa tehdessä vertailtiin kolmen eri järjestelmän ominaisuuksia suhteessa vaatimusmäärittelyyn (luku 5) ja työn tavoitteisiin.

Käyttöön otettavaksi järjestelmäksi valittiin Nagios Log Server sen helppokäyttöisyyden, monipuolisten ominaisuuksien ja edullisen lisenssikustannuksen vuoksi. Valintaa tehdessä koettiin SIEM-tuotteille tyypilliset edistyneet tekoälyominaisuudet tarpeeseen nähden liian monimutkaisiksi. Keskeinen tekijä eri tuotteiden ilmaisversioita kokeillessa oli uusien lokilähteiden lisäämisen yksinkertaisuus sekä web-hallintapaneelin helppokäyttöisyys.

8 NAGIOS LOG SERVER

Nagios Log Serverin toimintaperiaate perustuu ympäristössä luodun lokidatan keskittämiseen, indeksointiin, suodattamiseen, hakemiseen ja analysointiin. Järjestelmän sisällä toimii kolme avoimen lähdekoodin ohjelmistoa jotka vastaavat eri toiminnoista: Logstash, Elasticsearch ja Kibana (Full Architecture Overview 2019).

8.1 Logstash

Logstash on Nagios Log Serverin kriittisin ohjelmistokomponentti, sillä se vastaa saapuvan raa'an lokidatan vastaanottamistesta, jäsentelemisestä ja siirrosta Elasticsearchille indeksointia varten. Logstashin toiminta perustuu kolmeen vaiheeseen: syötteet (inputs), suodattimet (filters) ja tulosteet (outputs) (How Logstash Works 2020).

8.1.1 Syötteet

Logstashin syötevaihe kuuntelee ennaltamääriteltyjä TCP- ja UDP-portteja saapuvan lokidatan varalta. Kaikkiin määriteltyihin verkkoportteihin saapuva lokidata otetaan vastaan ja sille voidaan asettaa tyyppiarvo jotta suodattaminen olisi jatkossa helpompaa (Full Architecture Overview 2019).

Nagios Log Server sisältää oletuksena esimerkiksi syslogille ja Windowsin tapahtumalokille omat valmiit syötteensä. Käyttäjä voi itse luoda uusia syötteitä järjestelmän webhallinnassa. Tyypillisesti eri laitetyypit asetetaan lähettämään dataa eri verkkoportteihin.

Esimerkiksi UDP-porttiin 2099 saapuvalle datalle voidaan asettaa tyypiksi "firewall", jotta arvon perusteella olisi jatkossa helpompi luoda suodattimia ja hakuja:

```
udp {
    type => 'firewall'
    tags => 'firewall'
    port => 2099
}
```

Kuvio 8. Esimerkki Logstashin syötteestä.

8.1.2 Suodattimet

Suodatusvaihe on Logstashin monimutkaisin prosessi ja se vastaa datan jäsentelystä omiin kenttiinsä. Järjestelmässä on sisäänrakennettuna useita suodattimia kuten esimerkiksi *grok*, jonka avulla voidaan raa'asta datasta poimia elementtejä omiin kenttiinsä. Tämä mahdollistaa jäsentelemättömän datan järjestämisen omiin kenttiinsä joiden perusteella voidaan myöhemmin suorittaa hakuja ja indeksointeja (Full Architecture Overview 2019).

```
if [type] == 'firewall' {

    grok {
        match => [ 'src', '%{IP:srcip}:%{DATA:srcinfo}' ]

        match => [ 'dst', '%{IP:dstip}:%{DATA:dstinfo}' ]
    }
}
```

Kuvio 9. Esimerkki Logstashin suodattimesta.

Kuviossa 9 suodatetaan kuviossa 8 aiemmin luotua firewall-tyyppiä käyttäen *grok*-suodatinta ja if-lauseketta. Suodatin osaa poimia jäsentelemättömästä viestistä IP- ja MAC-osoitteet ja antaa niille omat kenttänsä.

Eräs toinen hyödyllinen suodatin on *geoip* joka mahdollistaa poimitun IP-osoitteen syöttämisen tietokantaan, joka kertoo IP-osoitteen alkuperän kaupungin tarkkuudella. Kaikki haun palauttamat tiedot jaetaan omiin kenttiinsä joiden avulla voidaan analysoida ympäristöön saapuvaa liikennettä esimerkiksi maatasolla. Kuviossa 10 on esitettynä ulkoverkosta yrityksen palomuuriin saapunut yhdistämisyrityksen sijaintidataosio JSON-formaatissa.

```
"geoip": {
      "city_name": "Falls Church",
      "continent_code": "NA",
      "country_code2": "US",
      "country_code3": "US",
      "country_name": "United States",
      "dma_code": 511,
      "ip": "[redacted]",
      "latitude": 38.864,
      "longitude": -77.1922,
      "postal_code": "22042",
      "region_name": "Virginia",
      "region_code": "VA",
      "timezone": "America/New_York",
      "location": [
        -77.1922,
        38.864
      ]
}
```


Kuvio 10. Esimerkki *geoip*-hausta JSON-muodossa.

8.1.3 Tulosteet

Logstashin tulosteet mahdollistavat jäsennellyn datan viemistä kolmannen osapuolen sovelluksiin, kuten esimerkiksi Graphite-monitorointityökaluun. Nagios Log Serverissä data viedään automaattisesti Elasticsearchille indeksoitavaksi (How Logstash Works 2020).

Koska Nagios Log Server koostuu vain yhdestä palvelimesta, niin tulosteiden konfiguroimiselle ei välttämättä ole tarvetta. Muiden tuotteiden toteutuksissa saattaa olla esimerkiksi yksi palvelin lokien keräämistä varten ja toinen sen analysoimista varten, missä tapauksessa tulisi konfiguroida tuloste lähettämään dataa eteenpäin.

8.2 Elasticsearch

Elasticsearch on tietokantaratkaisu joka vastaa lokidatan indeksoimisesta ja replikoinnista usean instanssin välillä. Elasticsearchissa indeksi on joukko lokidataa joka on esimerkiksi kerätty tietyn ajanjakson aikana.

Elasticsearch säilyttää indeksejä JSON-formaatissa, jossa data on jaoteltu erilaisiin kenttiin. Kentät voivat sisältää monenlaisia arvoja, esimerkiksi merkkijonoja, lukuja, totuusarvoja, päivämääriä ja koordinaatteja (What is Elasticsearch? 2020). Esimerkki JSON-formaatista nähtiin kuviossa 10.

8.3 Kibana

Kibana toimii Nagios Log Serverin web-komponenttina ja se vastaa Elasticsearchin indeksoiman datan visualisoinnista. Kibanan avulla voidaan poimia kerätyistä lokiviesteistä kenttiä tai arvoja joiden perusteella voidaan luoda esimerkiksi pylväsdiagrammeja, ympyrädiagrammeja, taulukoita ja karttoja. Näitä yhdistelemällä voidaan luoda erilaisia näkymiä (dashboard) joista voidaan käytännössä reaaliaikaisesti seurata ympäristön eri tapahtumia (What is Kibana 2020). Myös kaikki ylläpitoon liittyvät toimet tapahtuvat Kibanan web-hallintapaneelin kautta.

9 KÄYTTÖÖNOTTO

Nagios Log Serverin käyttöjärjestelmävaatimuksina ovat RHEL/CentOS, Debian tai Ubuntu (Administrator Guide 2020). Alustaksi valittiin CentOS sen vakauden takia ja koska kohdeympäristössä on pyritty vakioimaan palvelinkäyttöjärjestelmiä yhtenäiseksi.

Lokihallintajärjestelmä asennettiin virtualisoituna CentOS 7.6 -käyttöjärjestelmälle VMwaren vSphere 6.5 -virtualisointialustalle. Kohdeympäristön virtualisointiklusterissa oli käytössä useampi VMwaren ESXi -palvelin joten virtuaalikone saatiin ympäristön vikasietoisuuden piiriin heti alusta alkaen. Järjestelmän dokumentaation mukaan seuraavat laitteistovaatimukset olisivat suositeltuja.

- neljä prosessoriydintä
- vähintään 8 Gt käyttömuistia
- vähintään 1 Tt tallennustilaa. (Administrator Guide 2020)

Asennusvaiheessa käytössä olivat suositellut laitteistovaatimukset (poislukien tallennustila), joskin myöhemmin huomattiin että pitkältä aikaväliltä lokeja tarkastellessa järjestelmän prosessori- ja muistikapasiteetti eivät täysin riittäneet. Tästä syystä VMwaren hallinnasta virtuaalikoneelle lisättiin enemmän resursseja käytettäväksi. Lopulta virtuaalipalvelimella oli käytössään kuusi prosessoriydintä sekä 10 Gt käyttömuistia.

9.1 Asennusvaihe

Ennen asennusprosessin aloittamista varmistettiin CentOS-käyttöjärjestelmän pakettien olevan ajan tasalla ajamalla seuraavat komennot pääkäyttäjänä:

```
yum clean all
yum -y upgrade
```

Nagios Log Server -ohjelmiston asennusprosessi on hyvin yksinkertainen. Aluksi palvelimelle muodostetaan SSH-etähallintayhteys (Secure Shell), jonka jälkeen suoritetaan komento joka hakee sekä ajaa asennusskriptin Nagios Enterprisesin sivuilta käyttäen curl-ohjelmaa:

```
curl https://assets.nagios.com/downloads/nagios-log-server/install.sh | sh
```

Ohjelma voidaan myös asentaa käsin lataamalla asennuspaketti, purkamalla se väliaikaiskansioon ja suorittamalla fullinstall-skripti.

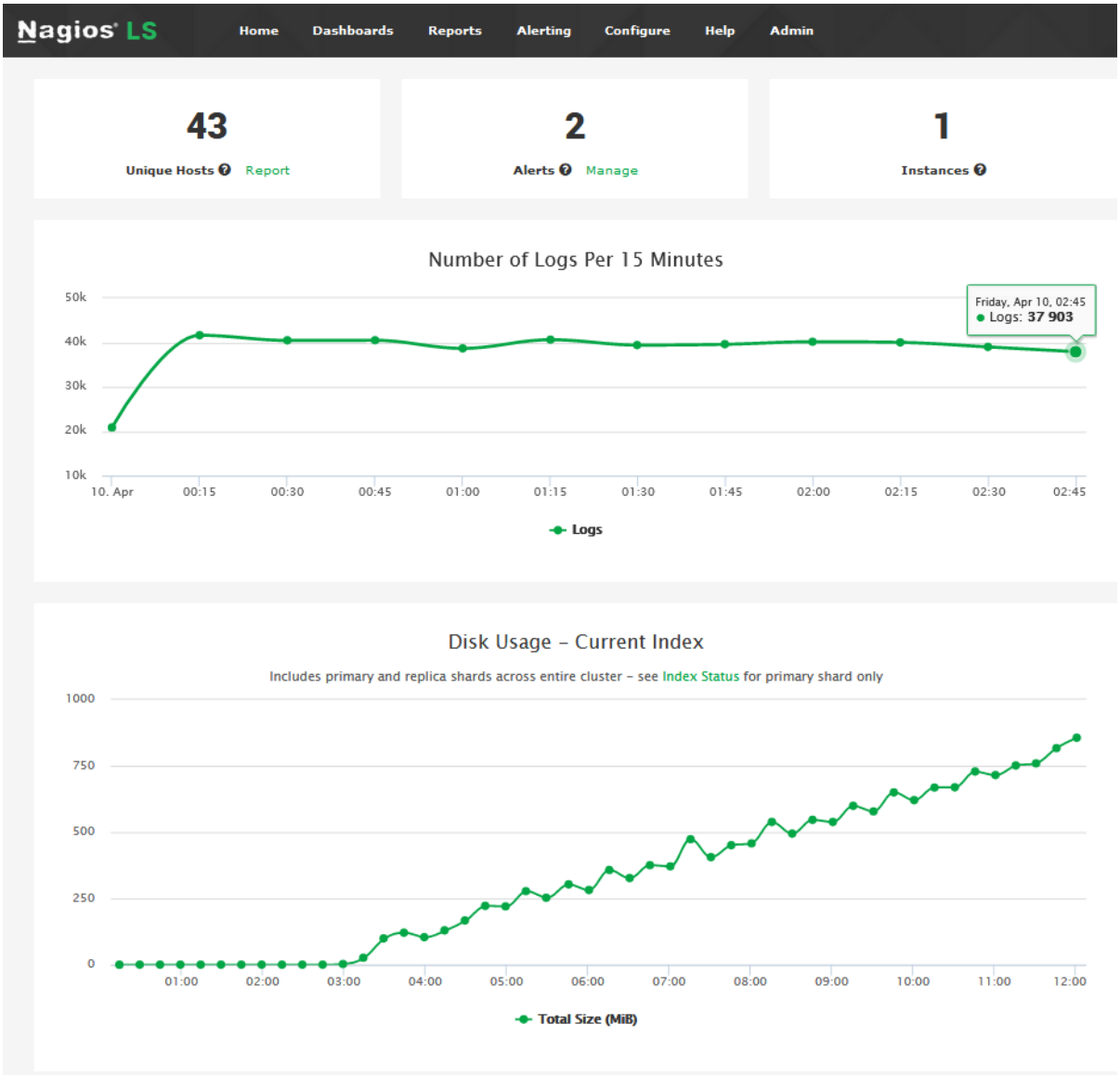
```
cd /tmp
wget  https://assets.nagios.com/downloads/nagios-log-server/nagioslogserver-latest.tar.gz
tar xzf nagioslogserver-latest.tar.gz
cd nagioslogserver
./fullinstall
```

Asennuksen jälkeen yhdistettiin järjestelmän web-hallintaan osoitteessa `http://nagios-log/nagioslogserver/`, jossa viimeisteltiin asennusprosessi syöttämällä lisenssikoodi ja määrittelemällä pääkäyttäjätunnus. Asennusta viimeistellessä oli mahdollisuus valita uusi asennus, tai lisätä palvelin jo olemassaolevaan klusteriin. Tämän jälkeen järjestelmä oli valmis käytettäväksi.

Onnistuneen asennuksen jälkeen palvelimella oli kolme uutta palvelua (service): Logstash, Elasticsearch sekä HTTPD-palveluprosessi (Hypertext Transfer Protocol Daemon) joka vastaa web-palvelimen toiminnasta.

9.2 Verkkohallinta

Järjestelmään kirjauduttua aukeaa eteen oletusnäkymä jossa näkyy tietoja järjestelmän toiminnasta, sekä pääsy erilaisille hallinnollisille välilehdille (kuvio 11).

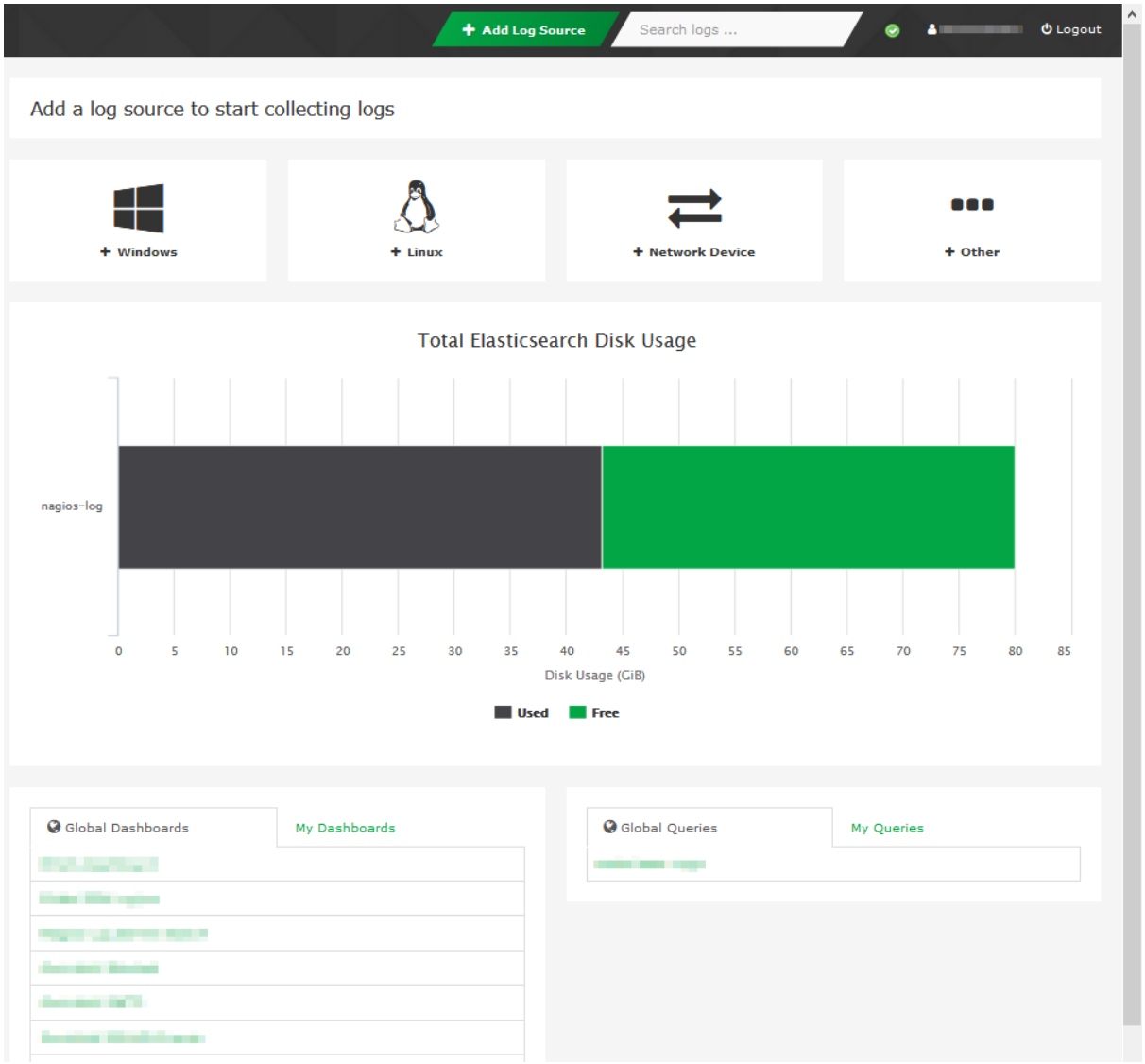


Kuvio 11. Nagios Log Serverin oletusnäkymä.

Kuviossa 11 näkyy kuvan yläosassa eri välilehdet joilta järjestelmän tilaa voidaan tarkastella ja konfiguroida. Tämän alla näkyy tietoja järjestelmään yhteydessä olevien laitteiden määrästä, aktiivisena olevista hälytyksistä sekä klusterissa olevien instanssien määrästä.

Kuvan keskiosassa on kaksi käyrää jossa ylemmässä näkyy viimeisen vuorokauden aikana vastaanotetun lokidatan määrä ja alemmassa tällä hetkellä auki olevien indeksien koko levyllä.

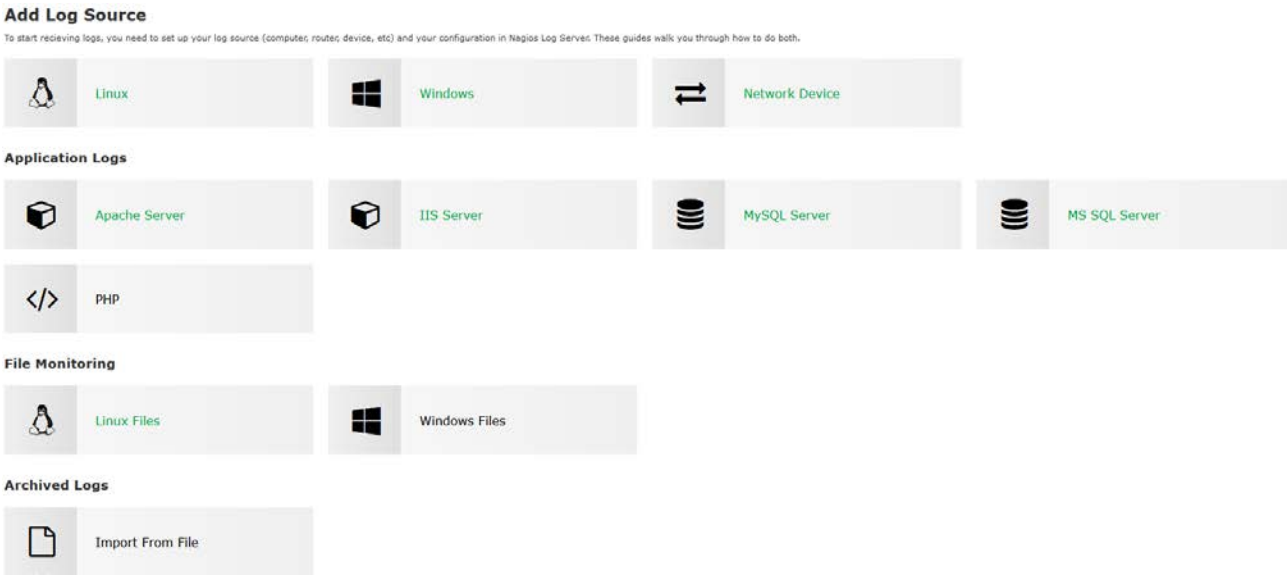
Sivun toisella puolella oleva puoli (kuvio 12) sisältää osion josta voidaan aloittaa uusien lokilähteiden lisääminen, diagrammin joka näyttää järjestelmän käytössä olevaa levytilaa, sekä listan jossa on käyttäjien luomat näkymät (dashboardit).



Kuvio 12. Nagios Log Serverin oletusnäkymän toinen puoli.

9.3 Laitteiden lisääminen ja analysointi

Lokidatan lähettäminen järjestelmään alkaa Nagios Log Serverin web-hallinnan etusivulta (kuvio 11), josta siirrytään Configure-välilehdelle. Sivulla on valittavissa useita lokilähteitä, esimerkiksi Windows- ja Linux-laitteita, verkkolaitteita, Apache- ja IIS-verkkopalvelimia ja kahdentyyppisiä SQL-tietokantapalvelimia (Structured Query Language).



Kuvio 13. Lokilähteiden lisääminen web-hallinnan kautta.

9.3.1 Windows

Windows-järjestelmien lokien lähettämiseen käytetään Nagioksen suosittelemaa NXLog-ohjelmaa. NXLog (Community Edition) on avoimen lähdekoodin ohjelmisto joka tukee lokidatan useita formaatteja kuten esimerkiksi syslog, XML ja JSON. (NXLog Community Edition 2020).

Ohjelmisto on ladattavissa suoraan Nagios Log Serverin web-hallinnasta. Ohjelman asennuksen jälkeen avataan kohdejärjestelmässä NXLogin konfiguraatiotiedosto, jonne liitetään Nagios Log Serverin valmiiksi antama konfiguraatio (kuvio 14).

```
<Input eventlog>
    Module im_msvistalog
</Input>

<Output out>
    Module om_tcp
    Host nagios-log
    Port 3515

    Exec  $tmpmessage = $Message; delete($Message); rename_field("tmpmessage","message");
    Exec  $raw_event = to_json();
</Output>
```

Kuvio 14. Ote NXLogin konfiguraatiosta.

Kuviossa 14 näkyvä konfiguraatio sisältää asetukset joilla Windows-järjestelmän tapahtumalokin data lähetetään lokipalvelimelle (nagios-log) TCP-portissa 3515. Lähetettävä data muunnetaan raa'asta XML-muodossa olevasta event log -datasta JSON-formaattiin. Konfiguraation lisäämisen jälkeen tulee käynnistää NXLog-ohjelman prosessi uudelleen, jotta asetukset tulevat voimaan.

Nagios Log Server -palvelimella käytetään jo valmiiksi konfiguroitua syötettä jossa määritellään Windows-laitteiden tapahtumalokin vastaanottaminen oikeassa portissa, sekä JSON-muodossa. Syötevaiheessa annetaan datan tyypiksi *eventlog* jotta myöhemmin voidaan luoda sen perusteella suodattimia (kuvio 15).

```
tcp {
    type => 'eventlog'
    port => 3515
    codec => json
}
```

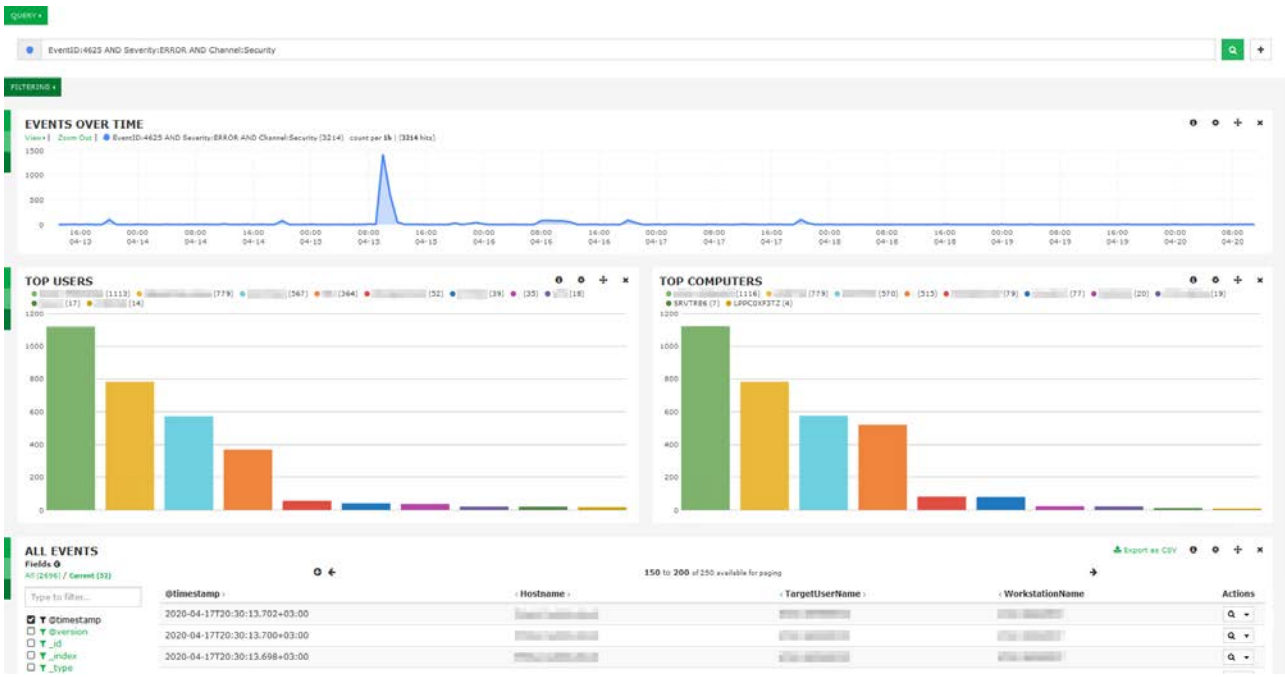
Kuvio 15. Windows-tapahtumalokin syöte.

Windowsin tapahtumaloki käyttää eri tapahtumatyyppeihin liitettäviä ID-lukuja, Severity-arvoja sekä Windows-lokin alta löytyviä kategorioita, esimerkiksi Security (katso luku 5.1.1). Näitä arvoja voidaan käyttää Nagiosin web-hallinnassa kyselyinä (query), jonka pohjalta lähdetään rakentamaan näkymää. Voidaan esimerkiksi luoda kysely joka näyttää Active Directory -toimialueen palvelimien lokiin jääneet epäonnistuneet kirjautumisyritykset seuraavaa tapahtumatyyppiä tarkastellessa:

```
EventID:4625 AND Severity:ERROR AND Channel:Security
```

Poimimalla lokiviestin eri kentistä dataa, voidaan alkaa muodostamaan analytiikkaa seuraamalla esimerkiksi laitteita joihin epäonnistuneet kirjautumisyritykset kohdistuvat, sekä

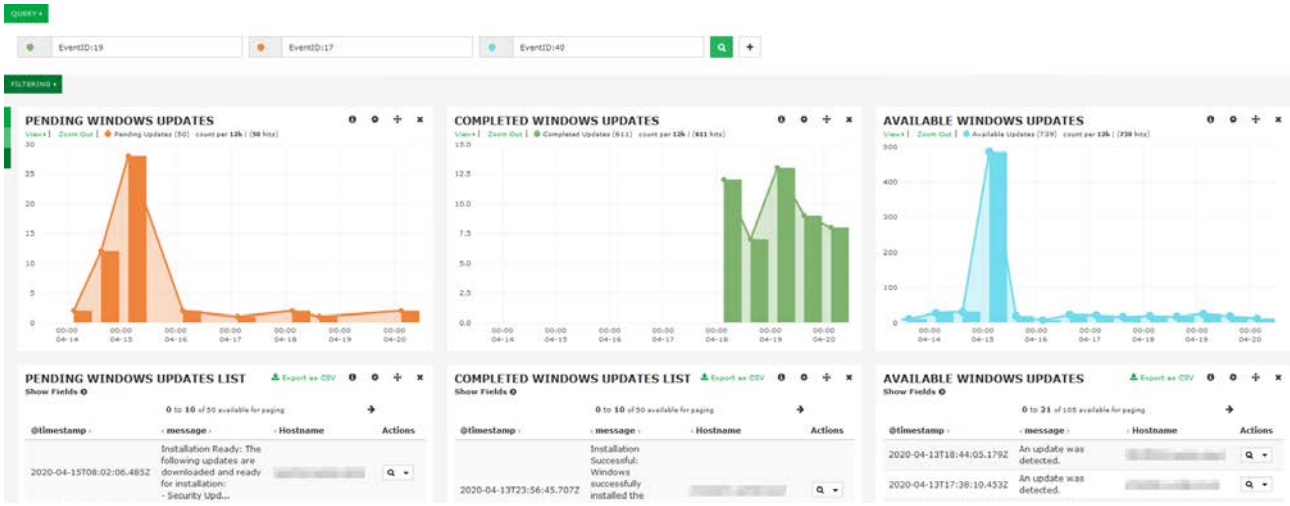
käyttäjätilejä joilla niitä tapahtuu. Nagioksen web-hallinnasta voidaan poimia näitä yksittäisiä elementtejä joiden pohjalta voidaan rakentaa esimerkiksi pylväsdiagrammeja ja aikajanan muodossa olevan käyrän.



Kuvio 16. Epäonnistuneet Windows-kirjautumiset.

Kuviossa 16 näkyy ylimpänä aikajanalle asetettu käyrä joka kuvaa tapahtumien lukumäärää ajan funktiona. Käyrän alla on molemmin puolin pylväsdiagrammit havainnoimassa käytettyjä käyttäjätunnuksia sekä laitteita joilta epäonnistuneita kirjautumisia on tullut. Näiden alla näkyy viimeisimpänä ilmenneet tapahtumat.

Kirjautumisyrityksien seurannan lisäksi luotiin näkymät joissa seurattiin Windows-palvelimien applikaatiolokin virheilmoituksia, System-lokin prosessivirheitä, sekä näkymä jossa seurattiin Windows-päivitysten tilaa palvelimien osalta (kuvio 17).



Kuvio 17. Windows-päivitystilanne.

Kuviossa 17 kuvataan toimialueen Windows-palvelimien päivitysten tilaa saatavilla olevien, asennusta odottavien ja asennettujen päivitysten osalta.

9.3.2 Linux

Jos Linux-palvelin käyttää käyttöjärjestelmänään CentOS:ää, Fedoraa, Red hatia, Ubuntua tai Debiania, niin se voidaan lisätä Nagios Log Serverin valvonnan piiriin käyttämällä automatisoitua skriptiä. Skripti noudetaan ja ajetaan Nagios Log Server -palvelimelta käyttäen komentokehoitetta. Skriptin ajovaiheessa annetaan Nagios-palvelimen nimi ja käytettävä portti:

```
curl -sS -O http://nagios-log/nagioslogserver/scripts/setup-linux.sh
sudo bash setup-linux.sh -s nagios-log -p 5544
```

Skripti havaitsee käyttääkö järjestelmä syslog-ohjelmanaan rsyslog:ia vai syslog-ng:tä ja luo tarvittavat konfiguraatiot sekä uudelleenkäynnistää tarvittavat prosessit.

Kuten Windows-laitteissa, Nagios-palvelimelta löytyy valmis syöte Linux-palvelimien syslog-dataa varten.

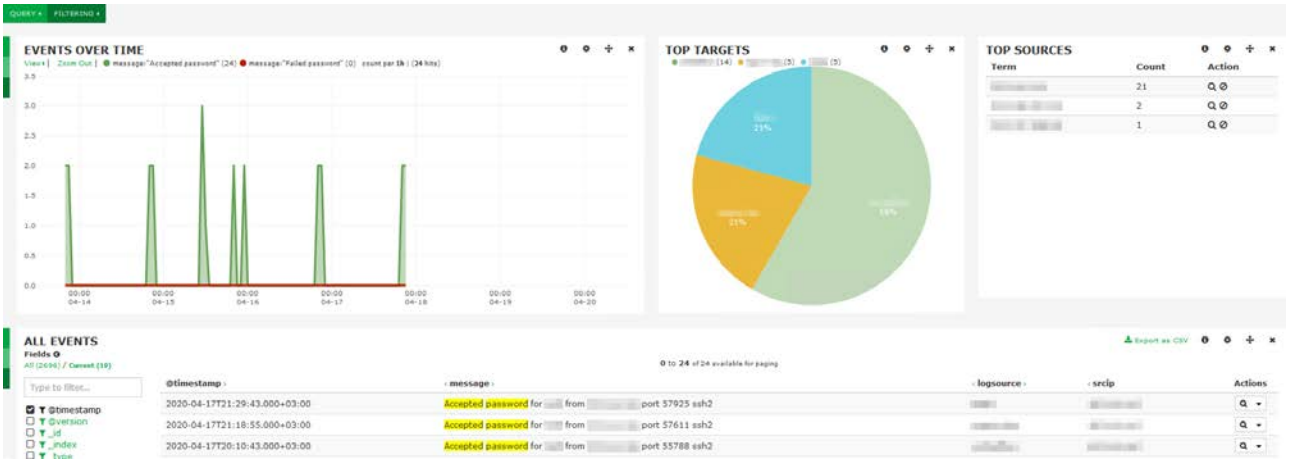
Linux-palvelimien osalta haluttiin kerätä tietoa onnistuneista ja epäonnistuneista sisäänkirjautumisista SSH-yhteyden kautta. Nagios-palvelimelle saapuvaa syslog-dataa ei oltu

jäsennelty Linux-palvelimien osalta kovin paljoa, joten viestit koostuivat vain aikaleimasta, facility-arvosta, lähdelaitteesta ja pitkästä viestikentästä. Tästä syystä lähdettiin koostamaan suodatinta joka osaisi poimia yhdistäjän IP-osoitteen.

```
if [program] == "sshd" {
    grok {
        match => [ "message", "%{IP:srcip}" ]
    }
}
```

Kuvio 18. Linux-palvelimien suodatin.

Kuviossa 18 luodaan if-lauseke, jota suoritetaan, jos lokiviestin facility-arvona on hallintaohjelma SSHD (Secure Shell Daemon). Lausekkeessa käytetään *grok*-parseria poimimaan viestikentästä IP-osoite ja sille luodaan oma kenttänsä (*srcip*). Tämän uuden kentän avulla voidaan Windows-palvelimien tavoin alkaa luomaan analytiikkaa Nagioksen web-hallinnan kautta.



Kuvio 19. Linux-palvelimien SSH-kirjautumiset.

Kuviossa 19 SSH-kirjautumisia seurattiin *srcip*-kentän avulla ja sen perusteella pystyttiin seuraamaan IP-osoitteita joista tulee eniten epäonnistuneita kirjautumisia. Käyttäen Nagioksen web-hallinnan query-työkalua luotiin näkymään myös mahdollisuus seurata onnistuneita kirjautumisia etsimällä syslog-paketin viestikentästä merkkijonoa “Accepted password”. Onnistuneiden ja epäonnistuneiden kirjautumisien käyrä näkyy kuviossa 19 ylävasemmalla.

9.3.3 Palomuurit

Palomuurien osalta Nagios Log Server ei tarjoa valmista ohjetta laitteen lisäämisestä, koska useimmissa laitteissa on aivan omat komentorivisyntaksinsa ja verkkohallintansa. Periaate eri laitteiden välillä on kuitenkin sama: määritellään kohteeksi Nagios Log Server, määritellään verkkoportti sekä haluttu Severity-arvo (luku 4.2.2).

Nagios Log Serverin web-hallinnassa konfiguroitiin uusi syöte palomuureja varten. Vaikka palomuurit käyttävät syslog-protokollaa, niin niiden luomat viestit eivät täysin noudata syslogin viestiformaattia. Tästä syystä laiteryhmälle valittiin omaksi portikseen 2099.

```
udp {
    type => 'firewall'
    tags => 'firewall'
    port => 2099
}
```

Kuvio 20. Palomuurien syöte.

Kuviossa 20 liikenteelle annettiin oma luokituksensa: firewall, jonka perusteella dataa voidaan myöhemmin suodattaa.

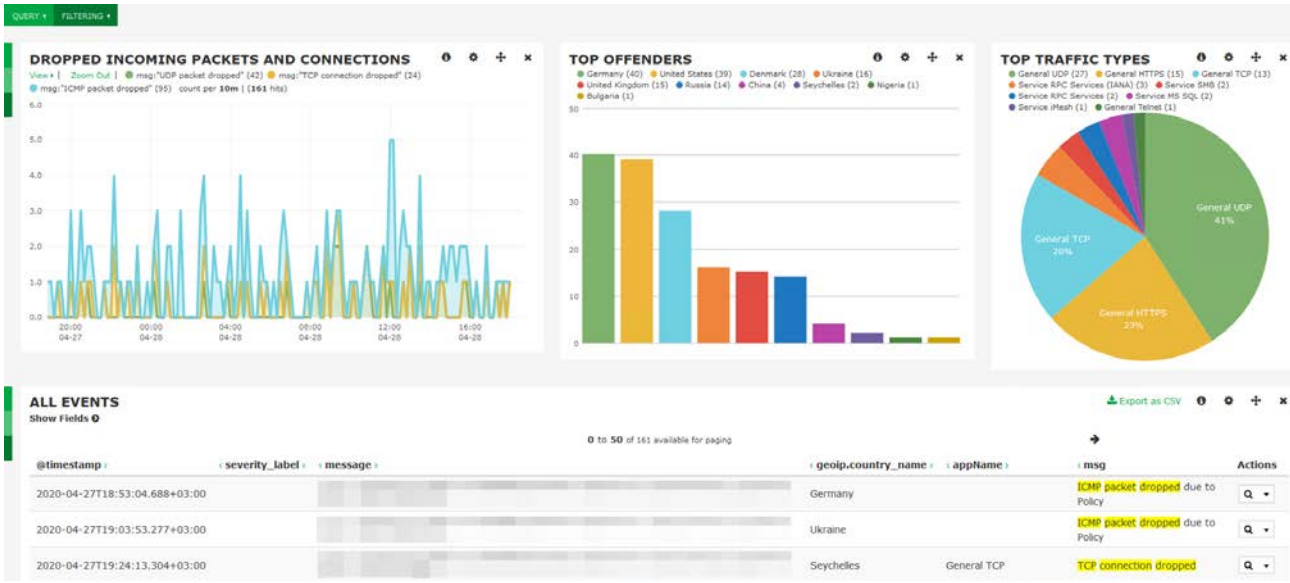
Lokipalvelimeen saapuvat viestipaketit sisältävät message-kentässä jonkin verran turhaa tietoa, kuten esimerkiksi lähde- ja kohdelaitteiden MAC-osoitteet. Tästä syystä ne päätettiin suodattaa pois. Tapahtumien lähde- ja kohde-IP-osoitteet haluttiin poimia omiksi kentikseen ja selvittää selvittää etenkin lähdeosoitteiden alkuperämaat.

```
1  if [type] == 'firewall' {
2
3      grok {
4          match => [ 'src', '%{IP:srcip}:%{DATA:srcinfo}' ]
5          match => [ 'dst', '%{IP:dstip}:%{DATA:dstinfo}' ]
6          remove_field => [ 'srcinfo', 'dstinfo' ]
7      }
8
9      geoip { source => 'srcip' }
10 }
```

Kuvio 21. Palomuurien suodatin.

Kuviossa 21 on esitettynä palomuureja varten konfiguroitu Logstashin parseri. Rivillä 3 käytetään *grok*-suodatinta poimimaan viestipaketista arvoja omiksi kentikseen. Samassa lausekkeessa myös poistetaan rivillä 6 MAC-osoitteet IP-osoitteiden perästä. Rivillä 9 lähde-IP-osoitteet syötetään *geoip*-liitännäiseen joka selvittää alkuperämaan ja parhaimmillaan jopa kaupungin. Liitännäinen käyttää avointa Maxmind GeoLite2 -tietokantaa IP-osoitteiden sijaintitiedoista (Logstash Reference 2020).

Luodut kentät ja erityisesti sijaintidata mahdollistivat esimerkiksi ympäristöön saapuvan estetyn liikenteen seurannan. Nagios Log Serveriin luodussa näkymässä seurattiin esimerkiksi saapuvia protokollia, lähtömaita ja palveluita, kuten esimerkiksi HTTPS (Hypertext Transfer Protocol Daemon)- ja SMB-liikenne (Server Message Block).



Kuvio 22. Palomuurien estetyt verkkoyhteydet

Kuviossa 22 on esitettynä Nagios Log Serveriin konfiguroitu näkymä, josta voidaan tarkastella kohdeympäristöön saapuvaa liikennettä joka on estetty. Yläosassa vasemmalla näkyy kolme päällekkäistä käyrää, jotka kuvaavat eri protokollien (TCP, UDP ja ICMP) osuutta kokonaisliikennemäärästä. Keskimmäisenä on pylväsdiagrammi, jossa on viimeisen viikon ajalta listattu eniten liikennettä tuottavat lähdemaat ja oikealla on eriteltynä eri palveluiden prosenttiosuus liikenteestä piirakkadiagrammin muodossa. Kuvan alaosassa näkyy viimeisimpänä saapuneet lokiviestit palomuureilta.

Kuvion 22 näkymän lisäksi konfiguroitiin kaksi muuta näkymää. Ensimmäisessä näkymässä seurattiin ulkoverkosta tulleita VPN-kirjautumisia (Virtual Private Network) sekä palomuurien hallintaosoitteisiin kohdentuneita kirjautumisyrityksiä.

9.3.4 Verkkokytkimet

Hewlett-Packardin verkkokytkinten osalta laitteisiin muodostettiin SSH-hallintayhteys ja ajettiin seuraavat komennot

```
logging 10.0.0.185 udp 5555
logging facility syslog
logging severity warning
```

Ensimmäisellä komennolla annettiin Nagios Log Serverin IP-osoite, määriteltiin käytettävä protokolla (UDP) ja annettiin verkkoportiksi 5555. Severity-arvoksi määriteltiin *warning.*

Nagios Log Serverin puolella konfiguroitiin uusi syöte käyttäen aiemmin määriteltyjä arvoja. Samalla annettiin verkkokytkinten syslog-viestipaketeille tyypiksi “switch-log”.

```
udp {
    type => 'switch-log'
    tags => 'switch-log'
    port => 5555
}
```

Kuvio 23. Verkkokytkimien syslog-syöte.

Verkkokytkimien osalta mitään merkittävää analytiikkaa ei kerätty. Näkymään luotiin kuvaaja esittämään saapuvien viestien lukumäärää suhteessa aikaan. Jo pelkästään viestimääriä tarkastelemalla voidaan pystyä havaitsemaan vikatilanne, sillä tyypillisesti verkkokytkimet luovat paljon viestejä jos esimerkiksi jokin komponentti on pettämässä. Kytkimissä syslog-viestien warning-taso lähettää lokidataa palvelimelle vain jos jokin vika havaitaan. Tästä syystä kytkimien osalta ei kovin paljoa viestejä tule.

9.3.5 Virtualisointipalvelimet

Yksi kohdeympäristön keskeisimpiä laitekokonaisuuksia ovat klusterissa toimivat VMwaren virtualisointipalvelimet. Klusteroitu palvelinkokonaisuus vastaa ympäristön virtualisoitujen palvelimien vikasietoisesta toiminnasta.

Aivan ensimmäiseksi Nagios Log Serveriin tuli avata palomuurista uusi verkkoportti, sillä VMwaren-palvelimet käyttävät syslog-datan lähettämiseen oletuksena UDP-protokollan porttia 514, joka ei oletusarvoisesti ole Nagios-palvelimella auki. Nagios Log Serverille yhdistettiin SSH-hallintayhteydellä ja ajettiin seuraava komentopari joka lisää uuden verkkoportin sekä uudelleenkäynnistää palomuurin prosessin:

```
firewall-cmd --zone=public --add-port=514/udp –permanent
firewall-cmd --reload
```

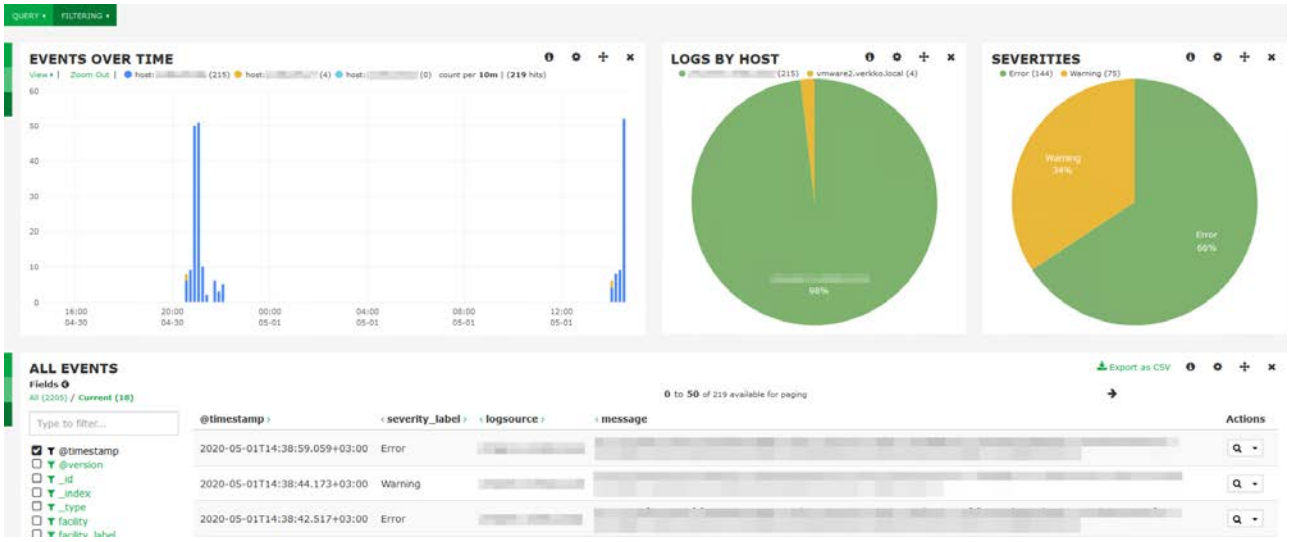
Tämän jälkeen voitiin luoda uusi syöte Nagiosin web-hallinnasta. Koska VMwaren ESXi-palvelimet lähettävät syslog-viestejä sen oikeassa formaatissa, niin UDP-lausekkeen sijaan luotiin syslog-lauseke:

```
syslog {
    type => 'syslog-esxi'
    port => 514
}
```

Kuvio 24. VMware-palvelimien syöte.

Syötteen luomisen jälkeen kirjauduttiin pääkäyttäjätunnuksella VMwaren vCenter-palvelimelle ja konfiguroitiin kaikki ESXi-palvelimet lähettämään lokidataa Nagios-palvelimelle.

Kuviossa 24 määritelty syslog-formaatin syöte mahdollistaa sen, että Nagios-palvelin osaa automaattisesti poimia saapuvista VMware-palvelimien lokiviesteistä severity-arvon. Tätä hyödyntämällä voidaan Nagioksen web-hallinnassa alkaa rakentamaan analytiikkanäkymää. Dashboard-sivulla luotiin uusi näkymä kuvaamaan VMwaren ESXi -palvelimien lokissa havaittuja virheilmoituksia.



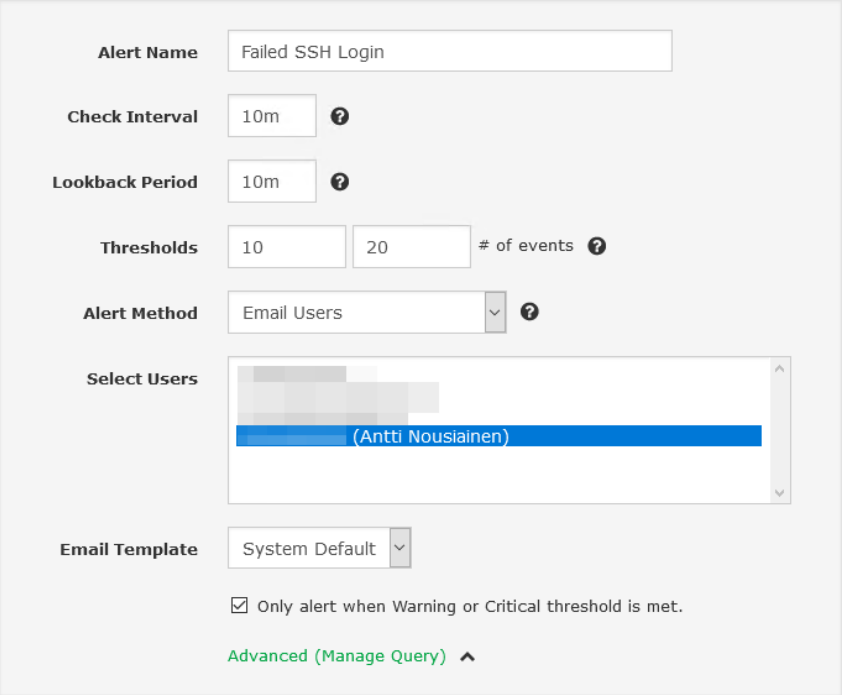
Kuvio 25. ESXi-palvelimien dashboard.

Kuviossa 25 näkyy VMwaren ESXi -palvelimien näkymäsivu. Näkymä rakennettiin syslogin severity-arvoa hyödyntämällä. Järjestelmään saapuvasta suuresta datamäärästä on esitettynä vain warning-tasosta ylöspäin olevan lokiviestit. Ylävasemmalla näkyy saapuneiden warning- ja error-viestien määrä viimeisen vuorokauden ajalta. Oikealla näkyvät kaksi piirakkadiagrammia, vasemmanpuoleisessa eri ESXi-palvelimien osuus datamäärästä, sekä oikealla warning- ja error-arvojen prosenttiosuus kokonaisviestimäärästä.

9.4 Hälytykset

Nagios Log Serverin web-hallinnan Alerting-välilehdellä on mahdollista konfiguroida erilaisia hälytyksiä. Hälytys voi laueta jos esimerkiksi riittävän monta tietyntyyppistä tapahtumaa havaitaan tietyllä aikavälillä. Tämän lisäksi on myös valittavissa reaaliaikainen hälytysmahdollisuus. Hälytyksen ilmoitustapa voi olla esimerkiksi sähköposti, mutta on myös mahdollista viedä ilmoitus kolmannen osapuolen järjestelmiin.

Nagiokseen konfiguroitiin epäonnistuneille Windows-kirjautumisille sekä SSH-yhteyksille omat intervallihälytyksensä, joista jälkimmäinen on esitettynä kuviossa 27.



Kuvio 26. Hälytysten konfigurointi.

Kuviossa 26 esitetään SSH-yhdistämisiä varten konfiguroitu hälytys. Tarkastusintervalliksi (check interval) asetettiin 10 minuuttia ja hälytyksen laukaiseviksi arvoiksi (thresholds) 10 ja 20 tapahtumaa. Ensimmäisen tapahtuma-arvon saavutettuaan järjestelmä lähettää warning-ilmoituksen sähköpostilla ylläpitäjille ja jälkimmäisen arvon kohdalla critical-ilmoituksen.

9.5 Käyttäjähallinta

Nagios Log Serverin web-hallinnan Admin-välilehti mahdollistaa uusien ylläpitäjä- ja operaattorikäyttäjien luomisen sekä hallinnan. Oletuksena myös jokaikinen kirjautuminen tallentuu vain ylläpitäjille näkyvään auditointilokiin joka on erotettuina muista lokitiedoista.

Järjestelmästä löytyy Active Directory -toimialueen tunnistautumista käyttävä liitännäinen joka päätettiin ottaa käyttöön käyttäjähallinnan helpottamiseksi.

Liitännäisen käyttöönottaminen tapahtuu Nagios Log Serverin Admin-välilehdeltä. Aluksi tulee määritellä toimialueen nimijärjestelmä LDAP-muodossa (esimerkiksi DC=contoso, DC=com), käyttäjätunnuksien jälkiliite (esimerkiksi @contoso.com) ja listata toimialueen Domain Controller -palvelimien osoitteet tai DNS-nimet (Domain Name System). Tämän

jälkeen voidaan tuoda toimialueelta jo olemassaolevia käyttäjätilejä Nagios Log Serverin järjestelmään.

Active Directory -liitännäinen tukee myös toimialueen sisäiseen käyttöön luotuja TLS-sertifikaatteja joita ei puuttumisen takia konfiguroitu.

9.6 Järjestelmäpäivitykset

Nagios Log Server saa järjestelmäpäivityksiä useasti vuodessa, mikä on tärkää järjestelmän toiminnan jatkuvuuden kannalta. Jos uusi järjestelmäpäivitys on saatavilla, niin Nagios ilmoittaa siitä näkyvästi web-hallinnan etusivulla.

Itse päivitysprosessi on hyvin yksinkertainen ja pysyy samana järjestelmäversiosta riippumatta. Päivittäminen aloitetaan muodostamalla SSH- tai terminaaliyhteys Nagios-palvelimelle. Jos käytössä on usean Nagios-instanssin klusteri, niin tulee disabloida lokidatan replikoinnin pirstaloituminen (shard allocation) ajamalla seuraava komento missä tahansa klusterin instanssissa:

```
curl -XPUT localhost:9200/_cluster/settings -d '{"transient":{"cluster.routing.allocation.enable":"none"}}'
```

Kyseinen komento muuttaa instanssin datan pirstalointiasetusta ja uusi asetus replikoituu myös muihinkin klusterin instansseihin.

Tämän jälkeen ajetaan komento joka hakee sekä suorittaa paikalliselta Nagios-palvelimelta päivitysskriptin:

```
wget -O upgrade.sh https://assets.nagios.com/downloads/nagios-log-server/upgrade.shsh ./upgrade.sh
```

Skripti suorittaa päivittämisen automaattisesti eikä vaadi käyttäjältä mitään toimintoja. Onnistuneen päivityksen jälkeen palautetaan shard allocation -asetus takaisin ennalleen seuraavalla komennolla:

```
curl -XPUT localhost:9200/_cluster/settings -d '{"transient":{"cluster.routing.allocation.enable":"all"}}'
```

10 YHTEENVETO

Työn idea syntyi kohdeympäristön tarpeesta saada keskitetty lokihallintajärjestelmä jollaista ympäristössä ei aikaisemmin ollut. Työn tarkoituksena oli etsiä, implementoida ja konfiguroida yrityksen kohdeympäristöön vaatimusmäärittelyssä listatut ominaisuudet täyttävä kaupallinen tuote.

Työssä vertailtiin kolmea soveltuvaa järjestelmää niin teknisten ominaisuuksien kuin lisenssikustannusten osalta. Järjestelmäksi valikoitui Nagios Enterprisesin vuonna 2014 lanseeraama Nagios Log Server. Lokihallintajärjestelmän käyttöönottoprosessi tapahtui suunnitelmien mukaan ja jo hyvin vähäisen konfiguroinnin jälkeen saatiin eri lokilähteistä kattavaa tilannekuvaa.

Tuloksena oli IT-ympäristön toimintavarmuutta ja tietoturvaa lisäävä keskitetty lokihallintajärjestelmä jolla voidaan lähes reaaliaikaisesti tarkastella usealta kannalta ympäristön kokonaistilannetta. Kerättyä lokidataa analysoimalla voidaan ratkaista tieto- ja palveluinfrastruktuurissa olevia ongelmia ja havaita tiettyjä elementtejä ennen kuin ne ehtivät muodostuvat ongelmiksi. Tästä esimerkkinä on tuotannossa sijaitseva verkkokytkin, joka alkoi lähettää jatkuvasti lokihallintajärjestelmälle ilmoituksia viallisista datapaketeista. Saaduista tiedoista pääteltiin vian olevan eräässä verkkokaapelissa, jonka vaihtamisen jälkeen ongelma poistui.

Työssä aikaavievin osuus oli oikeanlaisten suodattimien konfiguroiminen, jotta saapuvaa datamassa saataisiin jäsenneltyä järkevämpään muotoon. Myös eri näkymien konfigurointi vei melko paljon aikaa. Järjestelmän eri ominaisuuksiin oli riittävästi aikaa perehtyä, mutta esimerkiksi hälytyksiä olisi voitu määritellä enemmän. Myös lokien arkistointiin ja rotaatioon olisi voitu perehtyä enemmän.

Työn aihepiiri oli mielenkiintoinen ja työtä oli mielekästä tehdä. Aikaisempi kokemus monitorointijärjestelmistä sekä Linux-alustoista osoittautuivat hyödyllisiksi.

Pelkkä lokihallintajärjestelmän olemassaolo ei itsessään välttämättä tuo merkittävää lisäarvoa, vaan järjestelmää tulee jatkuvasti ylläpitää ja kehittää eteenpäin. Katsoisin että esimerkiksi ympäristön palomuureista saatava liikennemetriikka sekä tiedostopalvelimen auditointilokin tarkastelu voisi tuoda tuotteelle vielä enemmän lisäarvoa.

Oikeaoppisen lokikierron varmistamiseksi voitaisiin hyödyntää Nagios Log Serverin varastointiominaisuutta, jossa esimerkiksi yli kuukauden vanhaa lokidataa siirrettäisiin automaattisesti ulkopuoliseen järjestelmään, kuten verkkolevylle. Tämä mahdollistaisi järjestelmän suorituskyvyn paranemisen sekä vanhojen lokien tarkastelua pidemmältä aikaväliltä.

LÄHTEET

Agrawal K. ym. 2015. Data Analysis and Reporting using Different Log Management Tools. International Journal of Computer Science and Mobile Computing, vol.4:7. S. 224−229.

Carstensen, N. 2019. What is log management? A complete logging guide. Viitattu 18.3.2020 https://www.graylog.org/post/what-is-log-management-a-complete-logging-guide/.

Charter, B. 2008. EVTX and Windows Event Logging. Viitattu 20.3.2020 https://www.sans.org/reading-room/whitepapers/logging/paper/32949/.

Deveriya, A. 2005. Network Administrators Survival Guide. Indianapolis, Indiana, USA: Cisco Press. 552 s.

Elastic 2020. How Logstash Works. Viitattu 28.4.2020 https://www.elastic.co/guide/en/logstash/current/pipeline.html/.

Elastic 2020. Logstash Reference. Viitattu 28.4.2020 https://www.elastic.co/guide/en/logstash/current/plugins-filters-geoip.html/.

Elastic 2020. What is Elasticsearch? Viitattu 30.3.2020 https://www.elastic.co/what-is/elasticsearch/.

Elastic 2020. What is Kibana? Viitattu 28.4.2020 https://www.elastic.co/what-is/kibana/.

Graylog 2020. About Graylog: Helping Organizations At a Fraction of the Cost. Viitattu 26.3.2020 https://www.graylog.org/about/.

Graylog 2020. Graylog enterprise features. Viitattu 26.3.2020 https://www.graylog.org/features/.

Graylog 2020. Open Source vs. Enterprise. Viitattu 26.3.2020 https://www.graylog.org/products/open-source-vs-enterprise/.

Hicks, J. 2012. Command Line Event Logs. Viitattu 18.3.2020 https://www.petri.com/command-line-event-log/.

Internet Hall of Fame 2014. Eric Allman. Viitattu 4.3.2020 https://www.internethalloffame.org/inductees/eric-allman/.

Koja 2020. Kojan tarina – pienestä metallipajasta kansainväliseksi toimijaksi. Viitattu 4.3.2020 https://www.koja.fi/koja/kojan-tarina/.

Koja 2020. Kotimainen Koja-Yhtiöt. Viitattu 4.3.2020 https://www.koja.fi/koja/koja-yhtiot/.

Microchip 2020. TCP vs. UDP. Viitattu: 14.3.2020 https://microchipdeveloper.com/tcpip:tcp-vs-udp/.

Microsoft 2018. Event Types. Viitattu 20.3.2020 https://docs.microsoft.com/en-us/windows/win32/eventlog/event-types/.

Nagios Enterprises 2019. Full Architecture Overview. Viitattu 28.4.2020 https://support.nagios.com/kb/article.php?id=98/.

Nagios Enterprises 2020. Administrator Guide. Viitattu 4.4.2020 https://assets.nagios.com/downloads/nagios-log-server/guides/administrator/.

Nagios Enterprises 2020. Nagios Log Server. Viitattu 26.3.2020 https://www.nagios.com/products/nagios-log-server/.

NXLog 2020. The modern open source log collector. Viitattu 20.4.2020 https://nxlog.co/products/nxlog-community-edition/.

Palmer, M. 2008. Hands-On Microsoft Windows Server 2008 Administration. Boston, Massachusetts, USA: Cengage Learning. 524 s.

PCI Security Standards Council 2016. Effective Daily Log Monitoring. Viitattu 24.3.2020 https://www.pcisecuritystandards.org/documents/Effective-Daily-Log-Monitoring-Guidance.pdf/.

RFC 3164. 2001. The BSD syslog Protocol. Internet Engineering Task Force. 29 s.

RFC 5424. 2009. The Syslog Protocol. Internet Engineering Task Force. 38 s.

RFC 768. 1980. User Datagram Protocol. Internet Engineering Task Force. 3 s.

Solarwinds 2020. Ultimate Guide to Logging. Viitattu 20.3.2020 https://www.loggly.com/ultimate-guide/windows-logging-basics/.

Splunk 2020. Splunk Enterprise. Viitattu 26.3.2020 https://www.splunk.com/en_us/software/splunk-enterprise.html/.

Sumologic 2020. SIEM vs. Log Management. Viitattu 28.3.2020 https://www.sumologic.com/glossary/siem-log/.

Traficom 2019. Näin keräät ja käytät lokitietoja. Viitattu 4.3.2020 https://www.kyberturvallisuuskeskus.fi/fi/ajankohtaista/ohjeet-ja-oppaat/nain-keraat-ja-kaytat-lokitietoja/.

Valtiovarainministeriö 2009. Lokien säilytys, kerääminen ja suojaaminen. Viitattu 18.3.2020 https://www.vahtiohje.fi/web/guest/lokien-sailytys-keraaminen-ja-suojaaminen/.