Tampereen ammattikorkeakoulu

Keskitetty lokien hallintajärjestelmä Elastic Stackillä

Casper Honkaniemi

OPINNÄYTETYÖ
Marraskuu 2019

Tietojenkäsittely
Tietoverkkopalvelut

TIIVISTELMÄ

Tampereen ammattikorkeakoulu
Tietojenkäsittely
Tietoverkkopalvelut

HONKANIEMI, CASPER:
Keskitetty lokien hallintajärjestelmä Elastic Stackillä

Opinnäytetyö 43 sivua, joista liitteitä 7 sivua
Marraskuu 2019

Opinnäytetyön tavoite oli toteuttaa Elastic Stack -pohjainen lokien tallennus- ja analysointiratkaisu Oy Capnova Ltd:lle. Työn tarkoituksena oli tehostaa lokien prosessointia, madaltaa lokien analysoimisen kynnystä sekä mahdollistaa lokidatan visualisoiminen.

Työn tuloksena kehitettiin suunnitelma, jonka pohjalta lokien tallennus- ja analysointiratkaisu toteutettiin käyttäen Elasticin kehittämää ilmaista Elastic Stack -lokien hallintajärjestelmää. Elastic Stack koostuu neljästä avoimen lähdekoodin sovelluksesta: Elasticsearch, Logstash, Kibana ja Beats sekä Elasticin tarjoamasta X-Pack-lisäominaisuuspaketista. Järjestelmän toteuttamisen lisäksi tuotettiin jatkokehityssuunnitelma, jota käytetään järjestelmän kehittämiseen työn jälkeen.

Työn lopputuloksena Elastic Stackillä toteutetun järjestelmän avulla saadaan kerättyä halutut lokitiedot keskitettyyn tallennuskohteeseen analysoitavaksi. Toteutettua järjestelmää voidaan kuitenkin kehittää useilla tavoilla muun muassa lisäämällä järjestelmään lokien pitkäaikainen säilytyskohde, määrittämällä lokien keräys- ja käsittelykäytännöt sekä määrittämällä lokien automaattisen elinkaaren käytännöt.

Asiasanat: Elastic Stack, loki, keskitetty, tallennus, analysointi

ABSTRACT

Tampereen ammattikorkeakoulu
Tampere University of Applied Sciences
Degree Programme in Business Information Systems
Option of Network Services

HONKANIEMI, CASPER:
Centralized Log Management Platform with Elastic Stack

Bachelor's thesis 43 pages, appendices 7 pages
November 2019

The objective of this thesis was to create a centralized log storage and analysis platform based on Elastic Stack for Oy Capnova Ltd. The purpose of this work was to ease log processing and analysing and to enable log data visualizations.

As the result, a log management platform plan was created and implemented using the free Elastic Stack software by Elastic. Elastic Stack consists of four open source applications, Elasticsearch, Logstash, Kibana and Beats, and additionally X-Pack which adds additional functionality to Elastic Stack. In addition to implementing the platform, a plan to develop the platform further was also created.

As a conclusion, the implemented platform can collect the desired log data to a centralized location for analysis. However, the implemented platform can be further developed in several ways, for example by adding a long-term log storage destination, defining log collection and processing policies, and defining log lifecycle policies.

Key words: Elastic Stack, log, centralized, storing, analysis

SISÄLLYS

LYHENTEET JA TERMIT

API	Application Programming Interface
HTTPS	HyperText Transfer Protocol Secure
Indeksi	Tietokanta johon Elasticsearch tallentaa dokumentit (engl. index)
Instanssi	Yksittäinen käynnissä oleva sovellus
JSON	JavaScript Object Notation.
JVM	Java virtual machine.
Klusteri	Joukko noodeja, jotka toimivat yhtenäisesti.
Noodi	Yksittäinen virtuaalinen tai fyysinen kone, jolla sovellus pyörii (engl. node)
Oop	Ordinary object pointer. JVM:n käyttämä muisti osoitin. Joko pakattu tai ei pakattu.
Pakettivarasto	Varasto, josta voi hakea ohjelmapaketteja. (engl. repository)
Puskuri	Tila, johon tallennetaan tietoa, jota ei vielä olla prosessoitu
RPM	RPM Package Manager.
SIEM	Security information and event management.
Sirpale	Pienempi indeksin osa, josta muodostuu kokonainen indeksi. (engl. shard)
Sisääntulo	Datan vastaanottaja. (engl. input)
SNMP	Simple network management protocol.
SSD	Solid State Drive
SSL	Secure Sockets Layer
Suodatin	Kysely, jolla määritetään mitä dokumentteja käsitellään
Syslog	Protokolla lokien lähettämiseen keskitetylle palvelimelle.
TCP	Transmission Control Protocol
Ulostulo	Kohde, johon käsitelty data lähetetään. (engl. output)
X-Pack	Elastic Stackin lisäominaisuuskirjasto.
YAML	YAML Ain’t Markup Language.

1 JOHDANTO

Lokit ovat automaattisesti tai manuaalisesti kerättyjä merkintöjä tapahtumista, joita kerätään ennalta määriteltyä tarkoitusta varten ja säilytetään ennalta määrätyn ajan. Kerättyä tietoa hyödynnetään palveluiden ylläpitämisessä ja vikatilanteiden ja tietoturvapoikkeamien selvittämisessä. Lokien käyttöoikeudet tulee olla rajattu siten, että lokeja ei pystytä muuttamaan tai poistamaan. Myös lokien katselu ja lokeja keräävät palvelut tai prosessit tulee suojata, ettei luvattomat henkilöt voi lukea tai kirjoittaa lokeja. Lokien tallennuksen, säilytyksen ja käsittelyn vaatimukset asettavat lait ja asetukset, standardit, sekä organisaatio ja organisaation kumppanit. (Valtionhallinnon tietoturvallisuuden johtoryhmä 2009.)

Opinnäytetyön toimeksiantajana toimii Oy Capnova Ltd (myöhemmin Capnova). Capnova on tamperelainen vuonna 2000 perustettu IT-alan yritys. Capnova myy laajasti erikokoisille yrityksille IT-palveluitaan aina Web-hotelleista, palvelinsali- ja IT-infrastruktuuri ratkaisuihin.

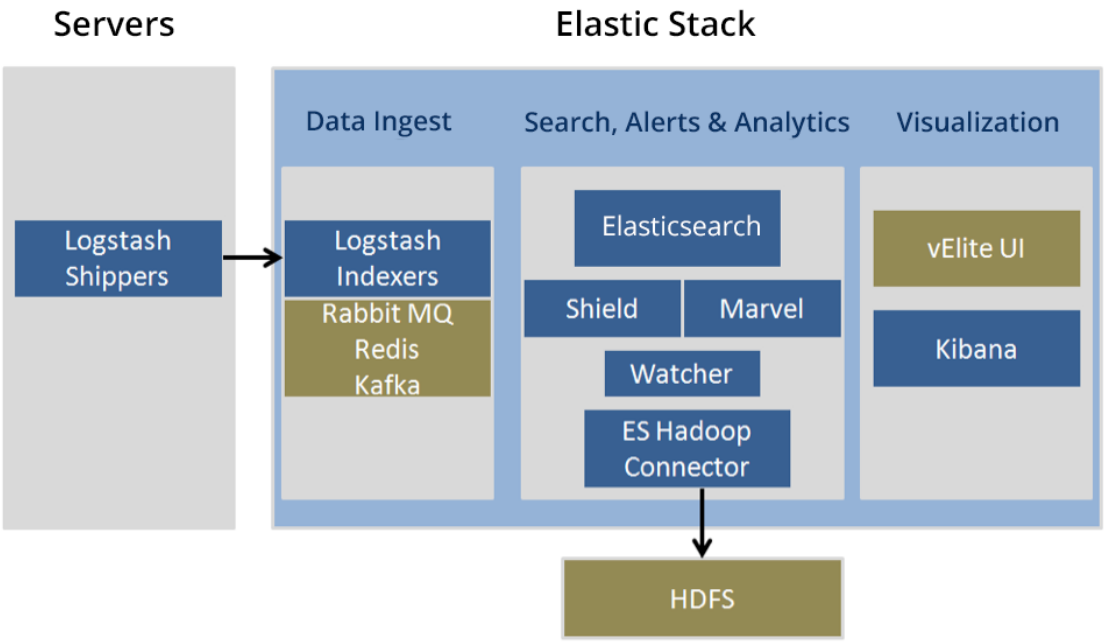
Capnovan nykyinen lokien keräys- ja analysointiympäristö käyttää keskitettyä syslog-palvelinta ja vaatii suuren määrän asiantuntemusta ja aikaa lokien analysoimiseen manuaalisesti komentojonolla. Lisäksi ympäristöllä ei pystytä visualisoimaan lokidataa tai yhdistämään lokeja eri lähteistä.

Opinnäytetyön tavoitteena on luoda Capnovalle keskitetty palvelu lokien tallentamiseen ja analysoimiseen käyttäen Elastic Stack ohjelmistokokonaisuutta. Palvelun on tarkoitus tehostaa lokien prosessointia ja madaltaa lokien analysoinnin kynnystä, sekä visualisoida erillisien palveluiden lokitietoja verkon yleiskuvan saamiseksi. Lisäksi palvelun on tarkoitus toimia pitkäaikaisena lokien tallennus kohteena.

2 ELASTIC STACK

Elastic Stack on Elastic NV:n ylläpitämä avoimen lähdekoodin sovelluskokonaisuus, joka tarjoaa työkalut datan keräämiseen, hakemiseen, analysoimiseen ja visualisoimiseen reaaliajassa. Elastic Stack koostuu Beatseista, Logstashista, Elasticsearchista ja Kibanasta. Näistä Elasticsearch toimii kokonaisuuden keskipisteenä.

Elastic Stack -kokonaisuuteen integroitujen Beats-tiedonkerääjien lisäksi voidaan Elastic Stackiin liittää monia muita komponentteja käyttötarpeiden mukaan, kuten esimerkiksi Verizon on yhdistänyt useita eri sovelluksia omaan ratkaisuunsa (kuvio 1). Lisäksi Elastic Stackin avoimen lähdekoodin ansiosta voidaan luoda täysin uusia kustomoituja ratkaisuja kuten Burkhart ja Warneck (2018) esittämä (Blizzardin luoma) BEAM-monitorointiratkaisu.



KUVIO 1. Verizonin Elastic Stack pohjainen lokien analysointiratkaisu (The Elastic Stack-powered evolution of Verizon Wireless… n.d.)

Elastic Stack on avointa lähdekoodia ja käyttää Apache 2.0 lisenssiä, joten kuka tahansa voi käyttää sitä ilmaiseksi. Tämän lisäksi aikaisemmin suljettua lähdekoodia ollut X-Pack-lisäominaisuuskirjasto on siirretty julkiseksi lähdekoodiksi,

mutta X-Pack säilytti aikaisemman kaupallisen lisenssin. Tämä tarkoittaa, että kuka tahansa voi kehittää X-Pack-lisäominaisuuskirjastoa, mutta jotkut X-Pack-kirjaston ominaisuuksista ovat maksullisia. X-Packin koodin julkiseksi siirtämisen yhteydessä osa aikaisemmin maksullisia X-Pack kirjaston ominaisuuksia on siirretty ilmaiseen Elastic Stack Basic -lisenssiin, mutta kehittyneemmät ominaisuudet vaativat edelleen Elastic Stack Gold- tai Platinum-lisenssin. (We Opened X-Pack n.d.)

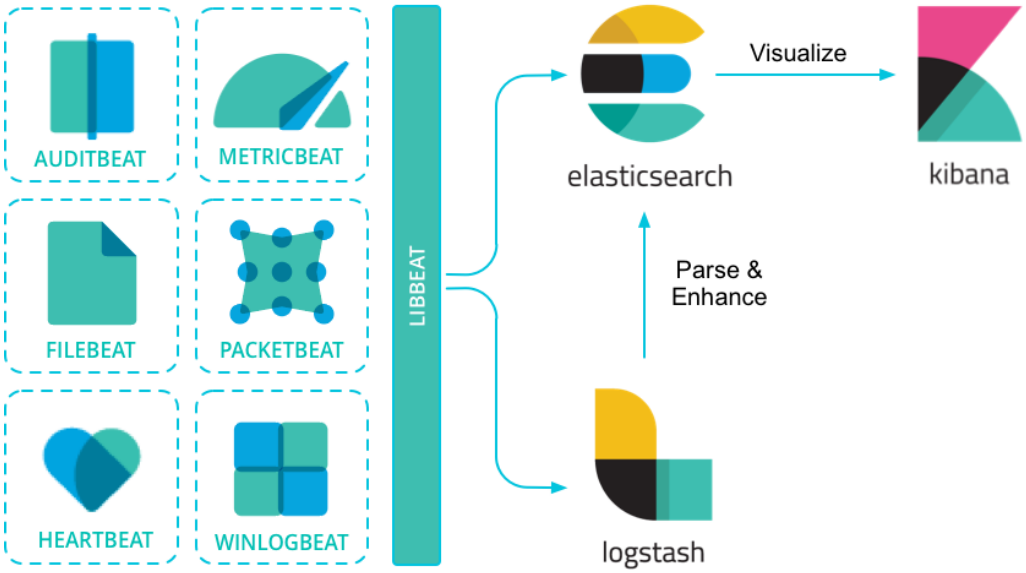
2.1 Beats

Beatsit ovat Elastic Stackiin sisältyviä kevyitä datasensoreita. Jokaisella Beatsilla on oma tehtävänsä ja ne keräävät ja lähettävät vain tietynlaista dataa. Beatsit asennetaan agenteiksi palvelimille, joilta halutaan lähettää dataa Elastic Stack -klusteriin joko suoraan Elasticsearchille tai Logstashin kautta prosessoidusti. Elasticin ylläpitämiä virallisia Beatseja on kahdeksan (taulukko 1), joista lokien keräämiseen tärkein on Filebeat, joka kerää ja lähettää lokitiedostoja Elastic Stack -klusteriin. Virallisien Beatsien lisäksi on 92 yhteisön ylläpitämää Beatsia. Kaikki Beatsit pohjautuvat libbeat-kirjastoon. Libbeat-kirjasto yksinkertaistaa yhteisön ylläpitämien Beatsien luontia ja tarjoaa API:n, joilla Beatsit voivat mm. lähettää dataa Elastic Stackiin ja konfiguroida syötteen asetukset (kuvio 2). (Beats overview n.d.)

TAULUKKO 1. Viralliset Beatsit

Beat	**Lähetettävä data**
Auditbeat	Linux-järjestelmien auditointidata
Filebeat	Lokitiedostot
Functionbeat	Pilvidata
Heartbeat	Saatavuusdata
Journalbeat	Systemd journalit (kokeellinen)

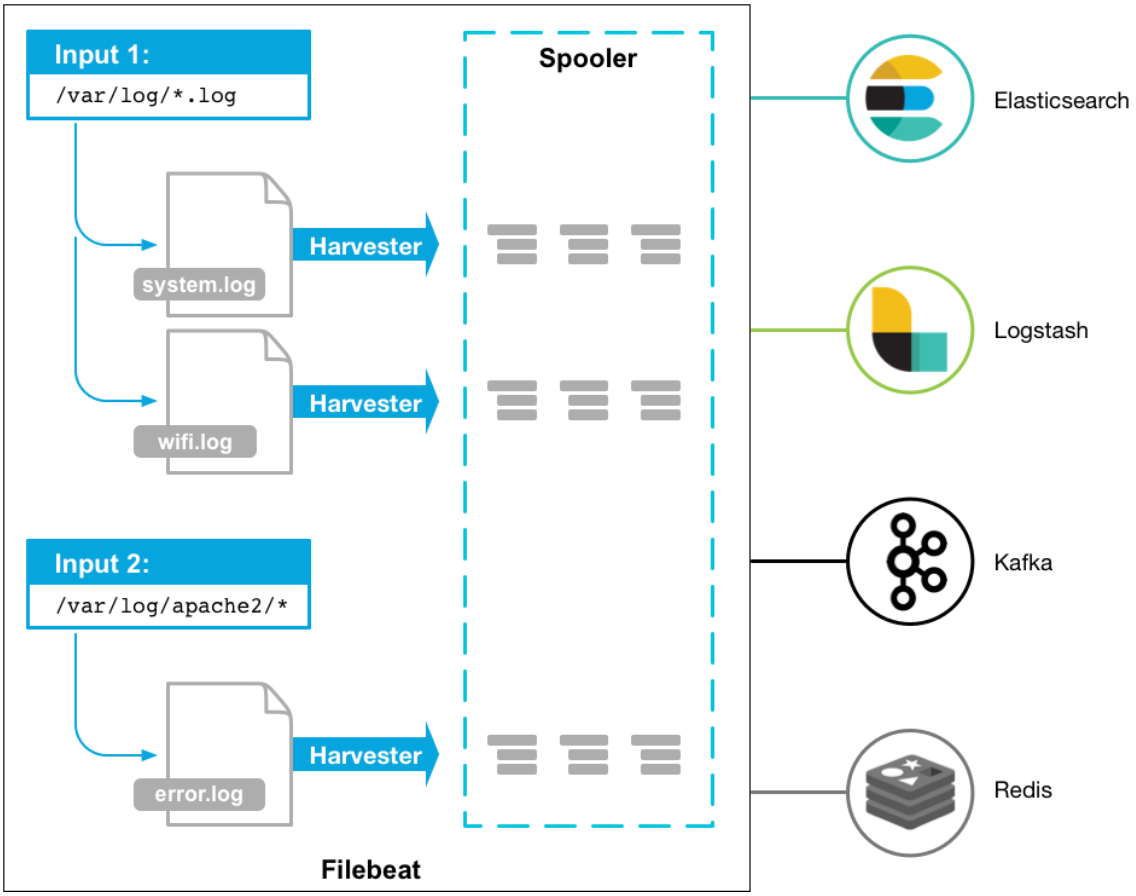
Metricbeat	Metrics
Packetbeat	Verkkoliikenne
Winlogbeat	Windows-tapahtumalokit



KUVIO 2. Beatsien toiminta (Beats overview n.d.)

2.1.1 Filebeat

Yksi beatseista on nimeltään Filebeat. Filebeat on kevyt lokien lähettäjä osana Elastic Stackiä. Filebeat seuraa määritettyjä lokitiedostoja ja sijainteja, joista se kerää ja lähettää lokit joko Elasticsearchille tai Logstashille. Filebeat koostuu kahdesta pääkomponentista, joista toinen on sisääntulo ja toinen kerääjä. Sisääntulot määrittelevät minkä tyyppisiä lokeja ja mistä kyseisiä lokeja Filebeat hakee. Kun sisääntulo löytää tiedoston, joka vastaa siihen määriteltyjä sääntöjä, se käynnistää tiedostolle kerääjän, jonka tehtävänä on lukea kyseisen tiedoston sisältö rivi riviltä. (kuvio 3). (Filebeat overview n.d.)



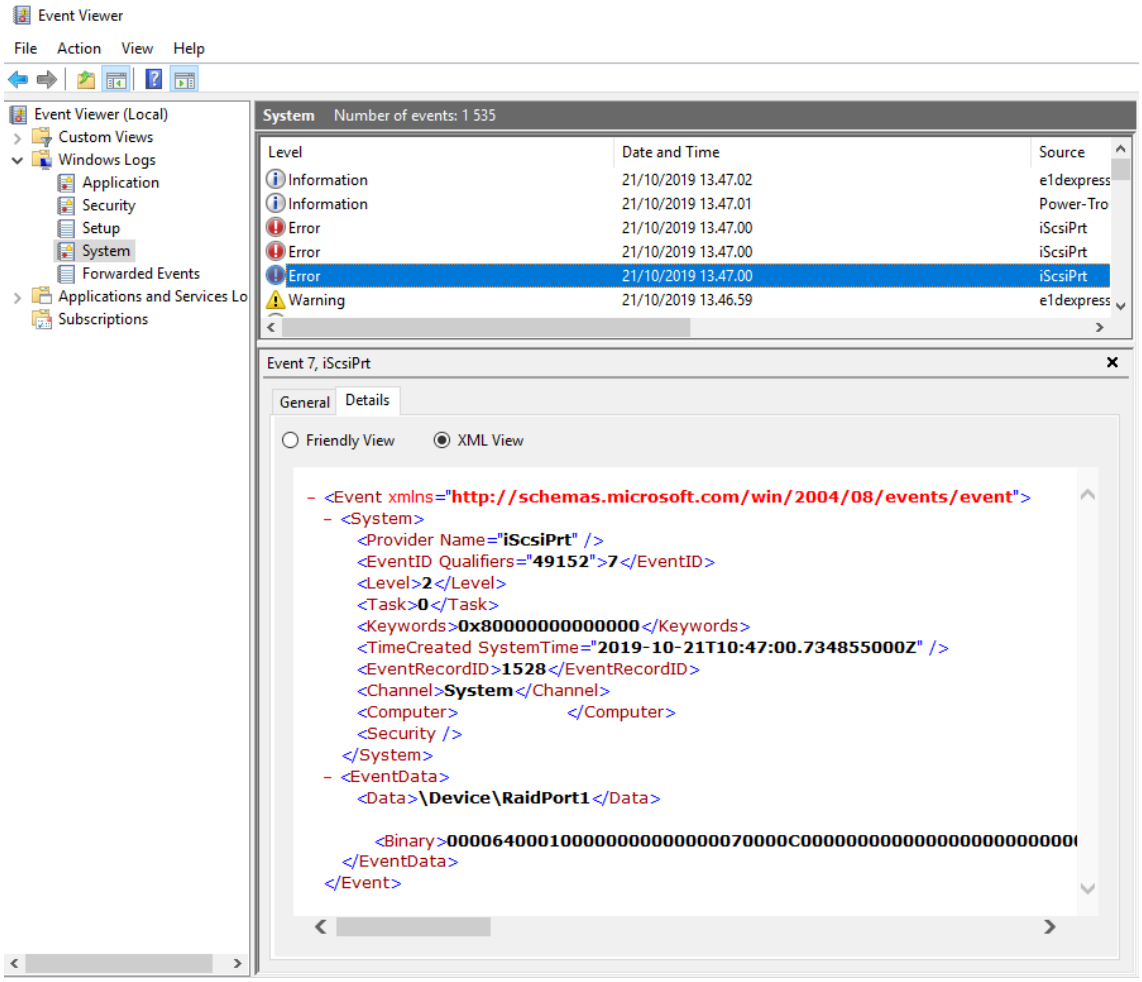
KUVIO 3. Filebeatin toiminta (Filebeat overview n.d.)

Filebeat pitää rekisterissä ylhäällä, kuinka pitkälle kerääjä on lukenut tiedostoa ja mitkä rivit on lähetetty Elastic Stackiin, joten Filebeat voi jatkaa lokien lähettämistä siitä, mihin se on jäänyt. Jos kohde, johon lokit lähetetään ei ole saatavilla, jatkaa Filebeat lokien lähettämistä uudelleen kohteen ollessa jälleen saatavilla. Näin Filebeat pystyy varmistamaan, että jokainen lokin rivi toimitetaan ainakin kerran Elastic Stackiin ilman datan menettämistä. (How Filebeat works n.d.)

2.1.2 Winlogbeat

Toinen beatseista, jota tässä työssä käsitellään, on Winlogbeat. Winlogbeat lähettää Windows tapahtuma lokeja Elastic Stackiin, joko Elasticsearchille tai Logstashille. Winlogbeat toimii Windowsissa palveluna taustalla. Winlogbeatille voi määrittää mitä tapahtuma kanavia se seuraa Windowsissa (kuva 1). Samoin Filebeatin tavalla Winlogbeat pystyy jatkamaan tapahtumien lähettämistä siitä

mihin jäätiin yhteyskatkoksien tai muiden katkoksien ajalta. (Winlogbeat Overview n.d.)

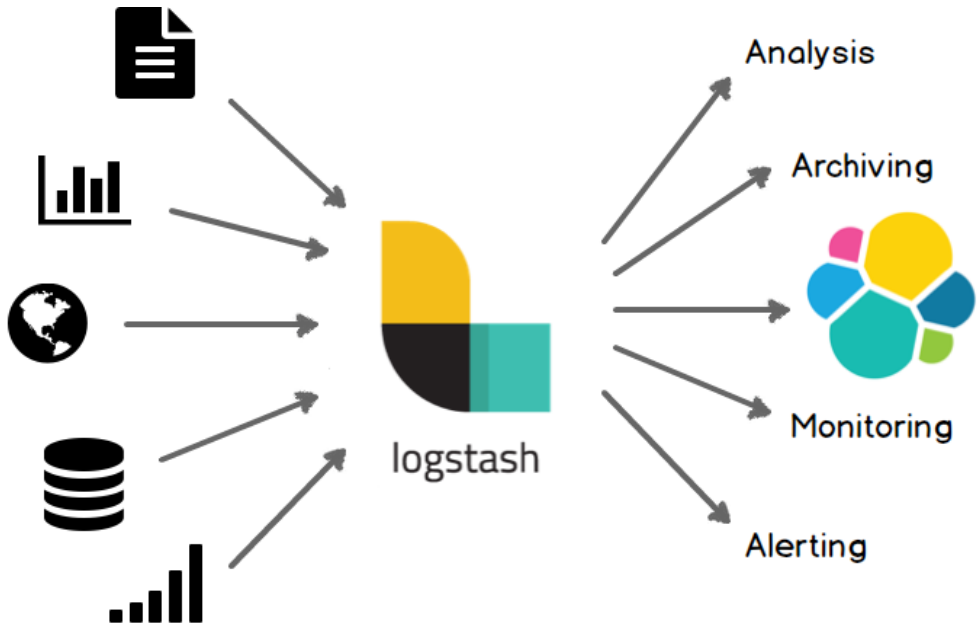


KUVA 1. Windows tapahtumienvalvonta-sovelluksessa järjestelmä kanavan virhe XML-muodossa

2.2 Logstash

Logstash on datankeräysmoottori, joka pystyy dynaamisesti yhdistämään dataa ja normalisoimaan datan eri kohteille. Logstash kerää, prosessoi ja lähettää datan eteenpäin ennalta määritetyllä tavalla (kuvio 4). Logstash voi esimerkiksi kerätä dataa Beatseilta, syslogilta tai SNMP:ltä. Logstash käyttää kanavia (engl. pipeline), tunnistamaan yhden palvelimen käyttämät lähteet toisistaan. Logstashin suurin hyöty Elastic Stackissä on, että sillä voi prosessoida dataa

käyttämällä monia erilaisia suodatukseen tarkoitettuja lisäosia ennen datan indeksointia. Esimerkiksi voidaan pudottaa virheenjäljityslokit pois, lisätä GeoIP-tiedot, muokata päivämäärät käyttämään samaa formaattia ja anonymisoida tietoa. Prosessoinnin jälkeen Logstash lähettää datan määritetyille kohteille, esim. Elasticsearchiin tai muille kohteille tallennettavaksi tai analysoitavaksi. (Logstash Introduction n.d.)

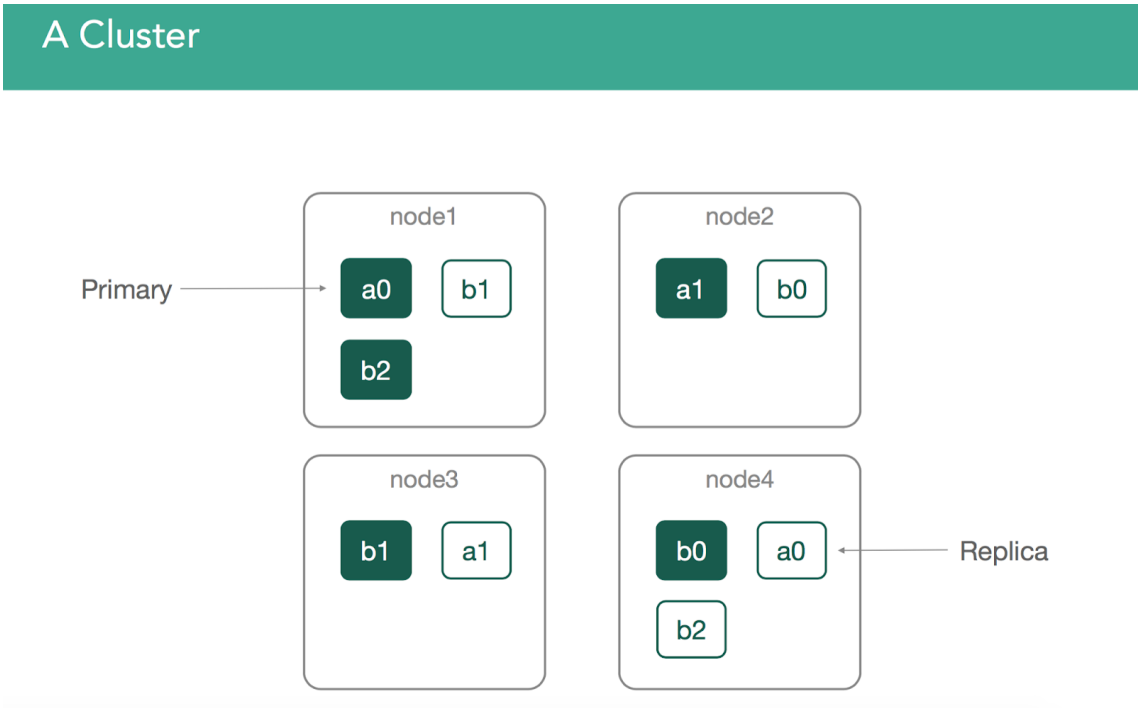


KUVIO 4. Logstash prosessi (Logstash Introduction n.d.)

2.3 Elasticsearch

Elasticsearch on hajautettu haku- ja analyysimoottori, joka toimii Elastic Stackin keskuksena. Elasticsearch tallentaa datan dokumentteina, jotka ovat indeksoitavia datayksikköjä. Dokumentit ovat JSON muodossa. Elasticsearch organisoi dokumentit indeksitietokantoihin. Indeksit voidaan luoda esimerkiksi päivän mukaan, jolloin yksi indeksi sisältää kaikki päivän dokumentit. Indeksit voivat olla valtavia, jolloin niiden prosessoiminen yhdellä noodilla kestää pitkään. Elasticsearch voi jakaa indeksit pienempiin osiin, joita kutsutaan sirpaleiksi. Indeksiä luodessa voi määrittää kuinka monesta sirpaleesta indeksi koostuu. Jokainen

sirpale on erillinen pieni indeksinsä ja mahdollistaa täten Elasticsearchin skaalaamisen useille noodeille. Sirpaleista voidaan luoda useita kopioita, replikoita, joka mahdollistaa Elasticsearchin vikasietoisuuden. Tästä myös tulee se hyöty, että noodi, jossa replika sijaitsee voi myös hakea sirpaleesta tietoa nopeuttaen hakuja (kuvio 5). (Basic Concepts n.d.)



KUVIO 5. Elasticsearch klusterin sirpaleiden sijoitus (Joshua Backing 2016)

2.4 Kibana

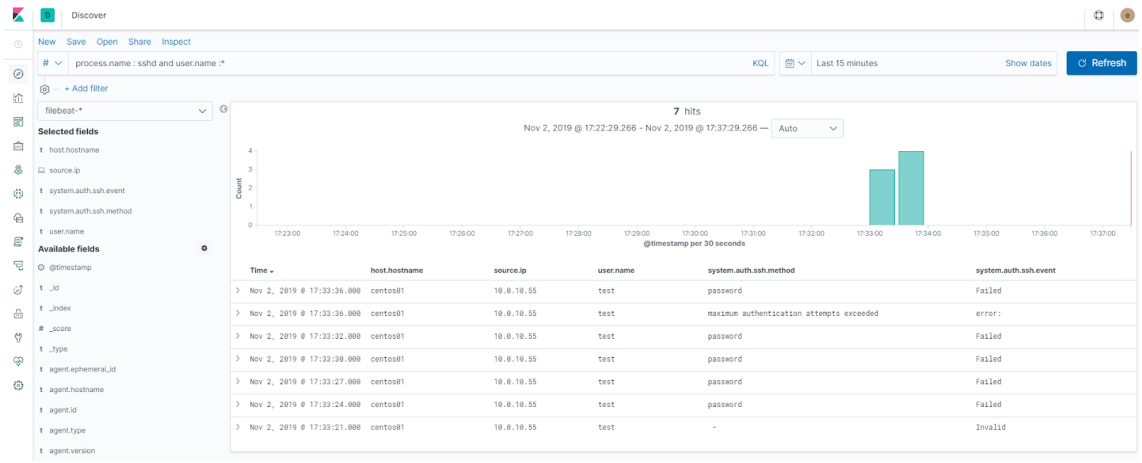
Kibana on Elasticsearchille suunniteltu analysointi ja visualisointi alusta, jota käytetään hakemaan, katsomaan ja tutkimaan dataa, jota on tallennettu Elasticsearch indekseihin. Kibanalla pystytään helposti analysoimaan ja visualisoimaan dataa eri muodoissa ja eri lähteistä. (Introduction n.d.) Kibana toimii Elastic Stack -klusterin graafisena käyttöliittymänä.

2.4.1 Kibanan työkalut

Kibanassa on useita työkaluja, joista lokien analysointiin tärkeimmät ovat Discover, Visualize, Dashboard ja SIEM. Lisäksi Kibana sisältää työkaluja klusterin valvontaan ja ylläpitoon. Kaikki Kibanan sisältävät työkalut ovat seuraavat: Discover, Visualize, Dashboard, Canvas, Maps, Machine Learning, Infrastructure, Logs, APM, Uptime, SIEM, Dev Tools, Stack Monitoring ja Management. (Kibana Guide n.d.)

Discover

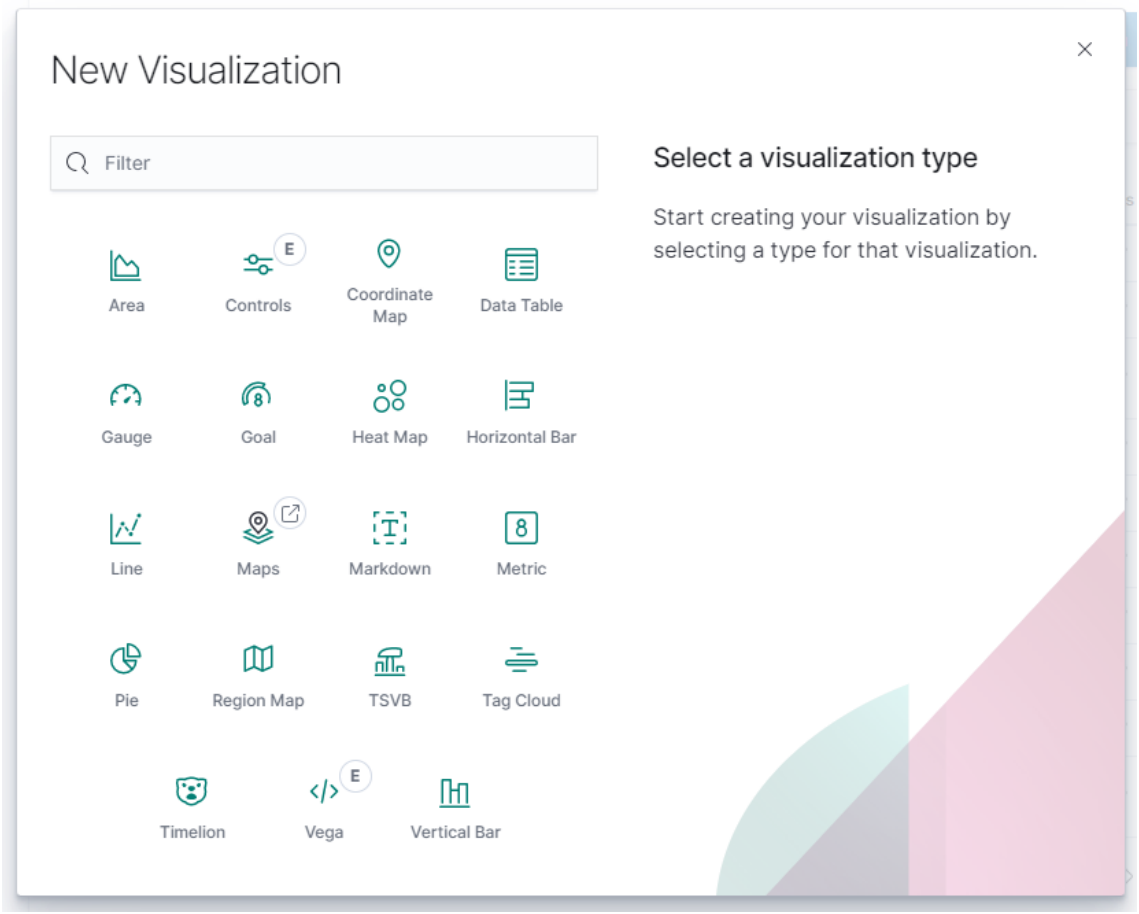
Discover välilehteä käytetään datan hakemiseen ja suodattamiseen. Välilehti sisältää aikasuodattimen, indeksimallin valinnan, hakukentän, filtterikentän, kenttävalinnan, histogrammin ja lokit (Discover n.d.). Esimerkiksi haulla “event.dataset:system.auth AND system.auth.ssh.event:*” voidaan hakea palvelimille tehtyjä SSH kirjautumisyrityksiä. Lisäksi valitsemalla kenttä valinnasta: “host.hostname”, “source.ip”, “system.auth.ssh.event”, “system.auth.ssh.method” ja “user.name” voidaan rajata kentät, jotka Kibana näyttää lokinäkymässä lokin analysoinnin helpottamiseksi (kuva 2). Lisäksi on mahdollista tallentaa haku myöhemmin käytettäväksi esimerkiksi visualisoinnin tekoon (Saving searches n.d.).



KUVA 2. Discover välilehti

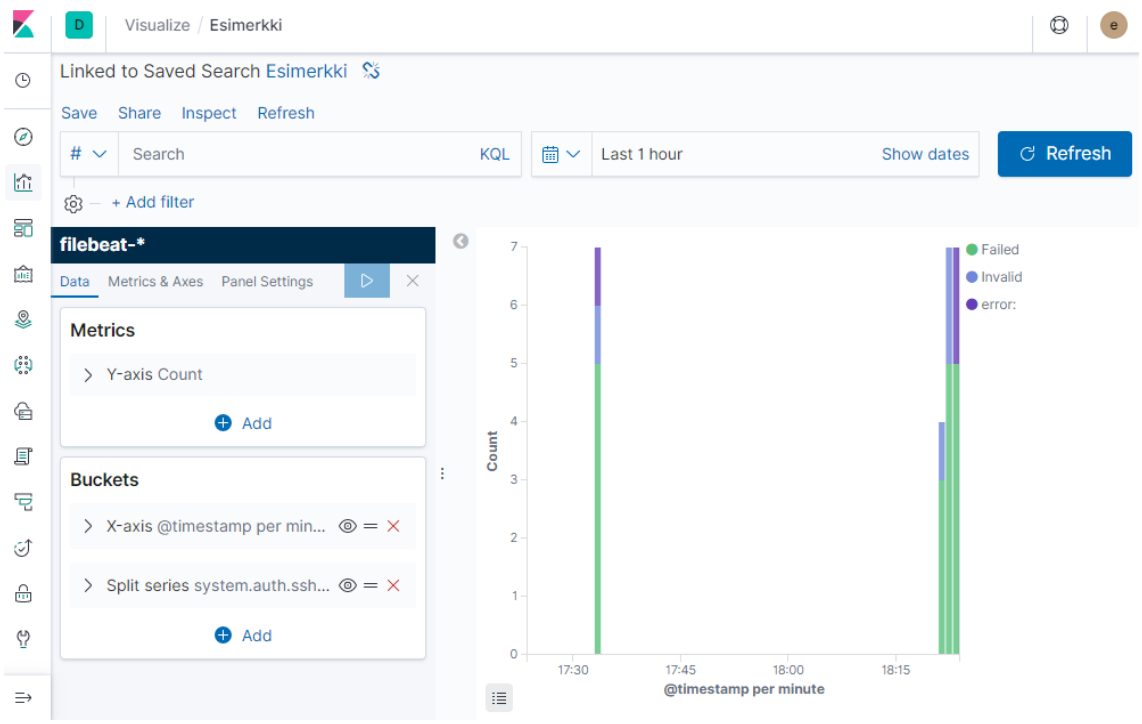
Visualize

Visualize välilehdellä voidaan luoda ja hallita kuvioita, kaavioita ja muita visualisointeja. Kibanalla voi luoda useita erilaisia visualisointeja (kuva 3). Visualisointeja voi luoda joko tallennetuista hauista tai tekemällä uuden haun. (Visualize n.d.)



KUVA 3. Visualisointien luonti

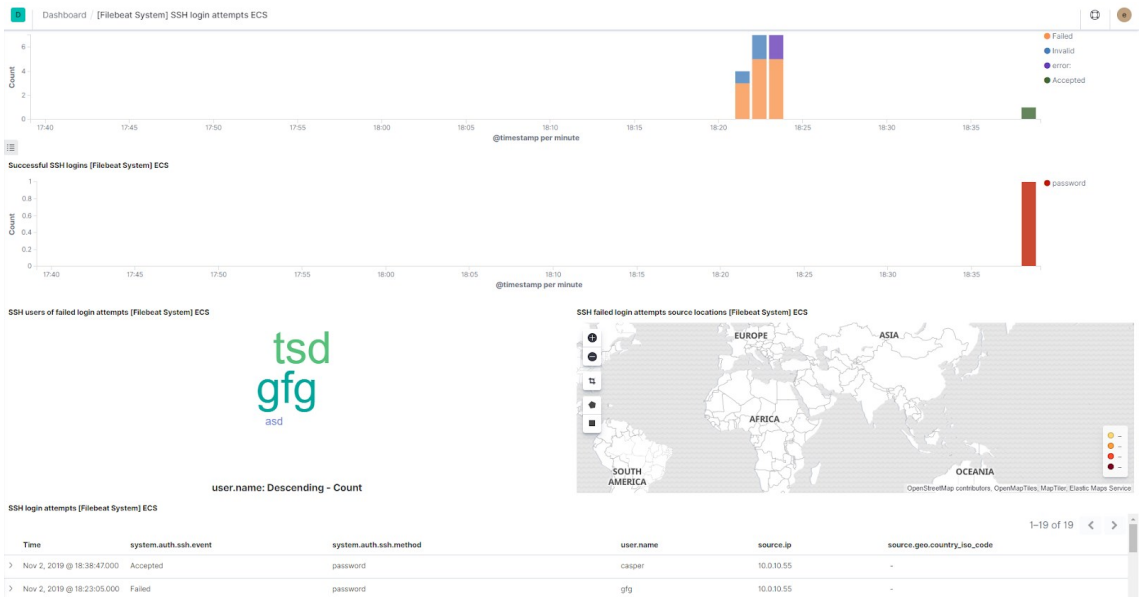
Valittua valmiin visualisoinnin tai valittua uuden visualisoinnin tyypin aukeaa visualisointi editori (kuva 4). Visualisointi editorissa voi discovery näkymää vastaavasti hakea, asettaa filttereitä ja valita aikavälin. Lisäksi voidaan asettaa kaavion asetukset esim. pylväskaavion akseleiden määreet. (Creating a Visualization n.d.)



KUVA 4. Visualisointi editori

Dashboard

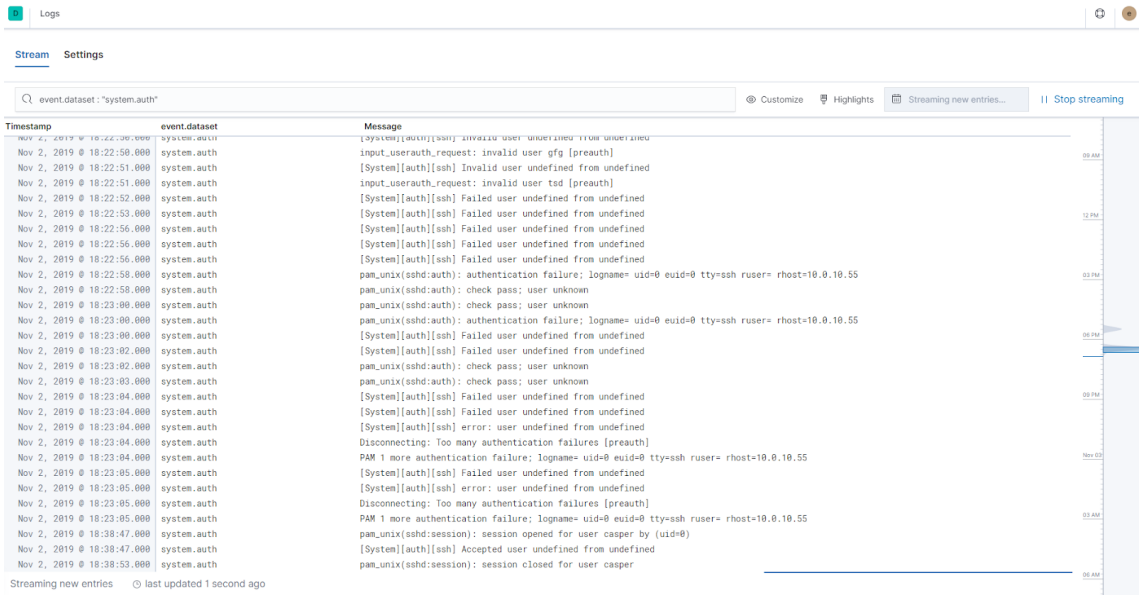
Dashboard näkymässä voidaan yhdistää useita tallennettuja visualisointeja ja tallennettuja hakuja (Dashboard n.d.). Dashboardit helpottavat järjestelmien kokonaiskuvan seuraamista (kuva 5). On mahdollista luoda useita dashboardeja seuraamaan eri asioita.



KUVA 5. SSH-autentikointi dashboard

Logs

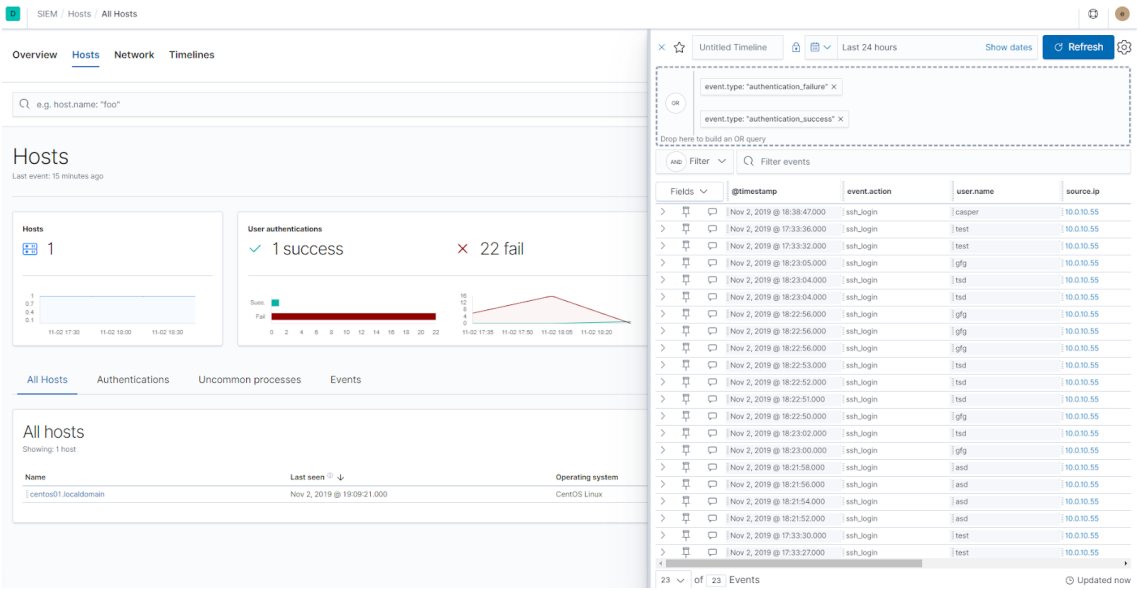
Logs välilehdellä voidaan seurata reaaliajassa lokeja, joita Elastic Search vastaanottaa (kuva 6). Näytettyjä lokeja voidaan rajata hakutermeillä ja lisäksi voidaan korostaa haluttuja termejä. (Logs n.d.)



KUVA 6. Lokien seuraaminen reaaliajassa

SIEM

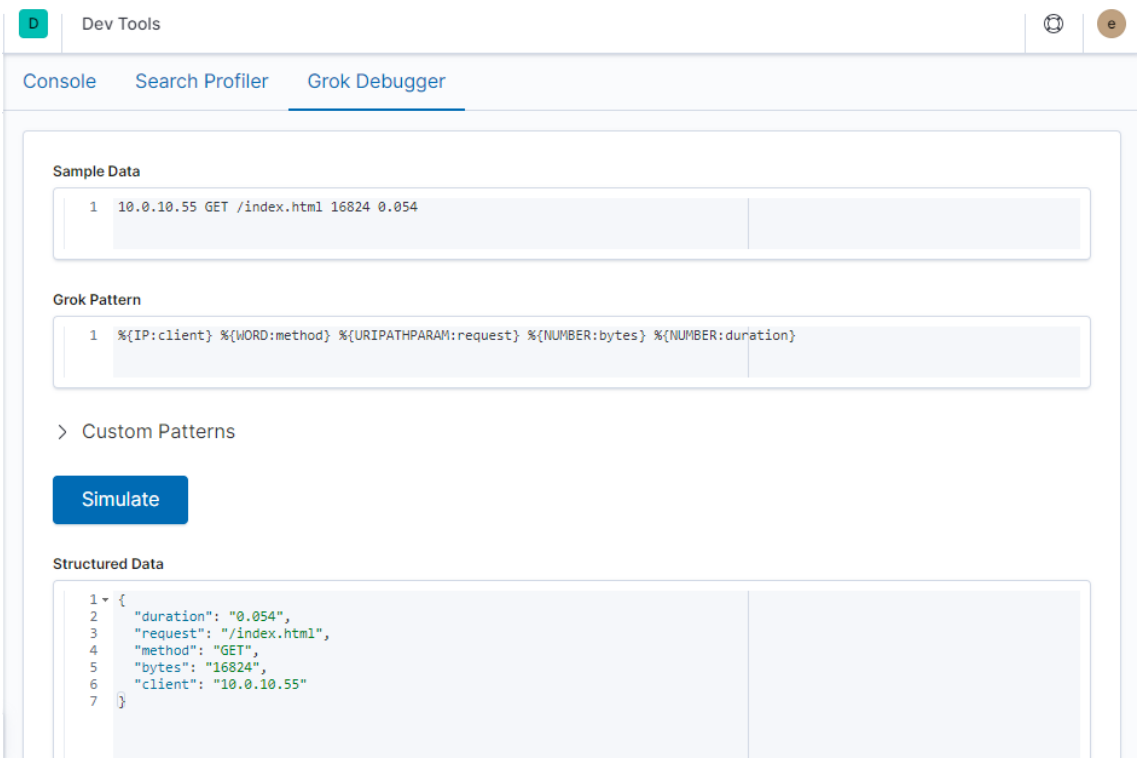
Elastic SIEM on beetavaiheessa oleva ominaisuus, jota voidaan käyttää SIEM-välilehdellä. Elastic SIEM:llä voidaan seurata ja tutkia hyökkäyksiin ja tietoturvaan liittyviä tapahtumia ja luoda tapahtumista aikajanoja, joilla selvitetään mitä on tapahtunut ja milloin on tapahtunut (kuva 7). Lisäksi Elastic SIEM tarjoaa Platinum lisenssillä koneoppia hyödyntävän poikkeamien seuranta palvelun. (SIEM UI n.d.)



KUVA 7. SIEM Hosts näkymä ja aikajana

Dev Tools

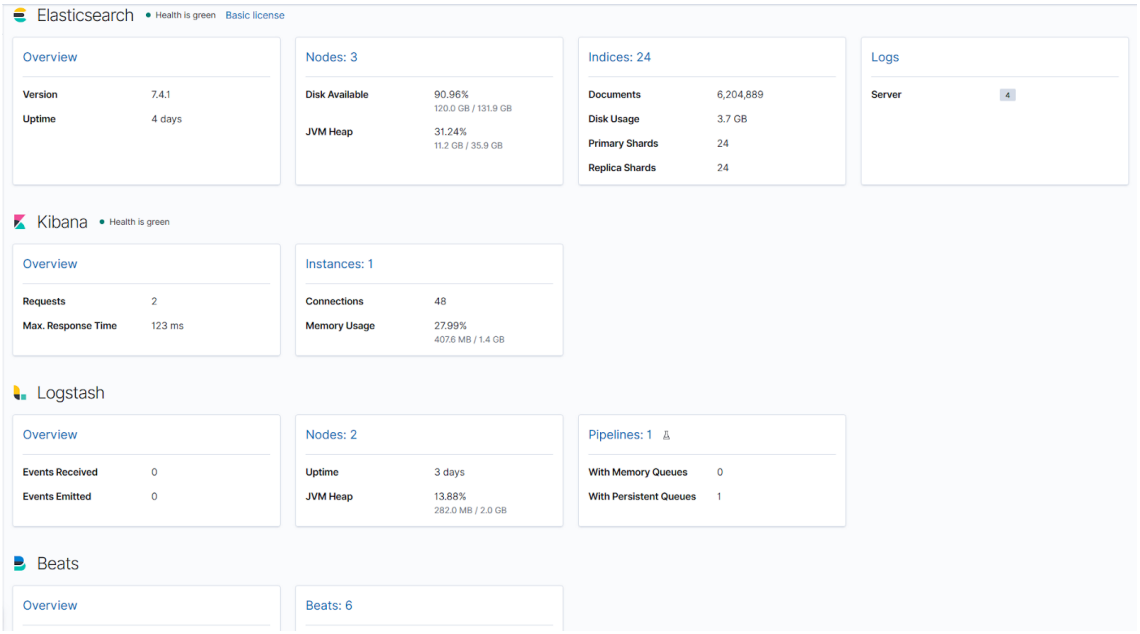
Dev Tools -välilehti sisältää työkaluja, joilla voi kokeilla Elasticsearchin API:n komentoja, analysoida haku kyselyitä, sekä testata Grok suodattimien toimintaa ennen sen käyttöönottoa (kuva 8) (Dev Tools n.d.).



KUVA 8. Grok Debugger Dev Tools -välilehdellä.

Stack Monitoring

Stack Monitoring -välilehdeltä voidaan valvoa kaikkia Elastic Stack -klusterin komponentteja (kuva 9). Stack Monitoring sisältää klusterin tilan, suorituskykytiedot sekä klusteriin liittyvät lokit. (Stack Monitoring n.d.)



KUVA 9. Stack Monitoring -välilehti

2.5 X-Pack

X-Pack on Elasticin tarjoama lisäominaisuuskirjasto. X-Packin lähdekoodi on julkista, mutta käyttää Elastic lisenssiä Apache 2.0 lisenssin sijasta (We opened X-Pack n.d.). X-Packista on ilmainen Basic versio, sekä maksulliset Gold ja Platinum versiot, jotka tuovat Elastic Stackiin lukuisia lisäominaisuuksia (taulukko 2) (Elastic Stack subscriptions n.d).

TAULUKKO 2. Elastic Stack lisenssit tiivistetysti

Avoin lähdekoodi	**Basic**	**Gold**	**Platinum**
Vain avoimen lähdekoodin ominaisuudet	Avoimen lähdekoodin ominaisuuksien lisäksi	Alemman lisenssin ominaisuuksien lisäksi	Alemman lisenssin ominaisuuksien lisäksi
Klusterointi	Keskeisimmät turvallisuus ominaisuudet	LDAP-, PKI-, AD -todennus	SAML kertakirjautuminen
Hakutyökalut	Rooleihin perustuva käyttäjien hallinta	Hälytykset	Tallennetun datan salaaminen
Analysointityökalut	Sovellusten valvonta	Keskitetty hallinta	Koneoppiminen
Datan visualisointi	SIEM	Raporttien generointi	Klustereiden välinen replikointi
	Kartat	Tukipalvelu	24/7/365 tukipalvelu

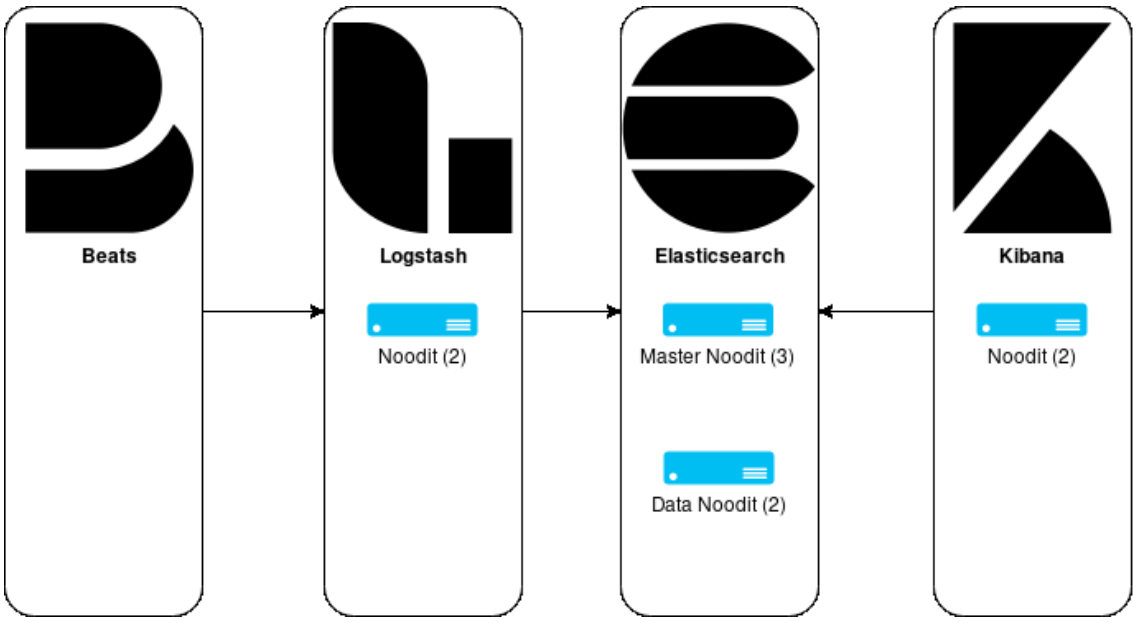
3 LOKIEN ANALYSOINTIYMPÄRISTÖN SUUNNITTELU

“Lokitietojen tehokas analysointi on usein lokitietojen hallinnan ja käsittelyn vaativin mutta toisaalta usein myös tärkein osa-alue. Vaikka lokitietojen analysointi nähdään työläänä ja tehottomana toimenpiteenä, voidaan hyvillä lokien hallintatyökaluilla ja ympäristöllä automatisoida lokitietojen käsittely ja analysointi, jolloin se vie vähemmän aikaa.” (Valtionhallinnon tietoturvallisuuden johtoryhmä 2009)

3.1 Infrastruktuuri

Vikasietoisuus

Lokien keräys- ja tallennusjärjestelmälle on tärkeää, että siinä ei ole käyttökatkoja ja, että järjestelmällä on korkea saatavuus (engl. high availability) (kuvio 6). Joten järjestelmän toteutuksen tulee olla vikasietoinen ja minimoida järjestelmän käyttökatkot. Elasticsearch vaatii kyseiseen toteutukseen vähintään kolme master-noodia, joista vähintään kaksi pitävät toimia aktiivisena master-noodina (Node n.d.). Mikäli klusterissa on vain yksi Elasticsearch data-noodi, ei replikasirpaleita voida luoda, joten vikasietoisuuden saamiseksi Elasticsearch vaatii ainakin kaksi data-noodia (Adding nodes to your… n.d.). Logstash vaatii toteutukseen kaksi noodia, jolloin Logstash-noodit voivat tasata kuormituksen ja toisen noodin katkoksessa toinen voi vastaanottaa kaiken datan (Deploying and Scaling Logstash n.d.). Lisäksi Logstash voidaan asettaa käyttämään palvelimen tallennustilaa lokien puskurina muistin sijasta, jolloin esim. järjestelmän kaatumisen yhteydessä ei menetetä puskurissa olevaa dataa (Execution Model n.d.).



KUVIO 6. Esimerkki Elastic Stackin vikasietoisuudelle

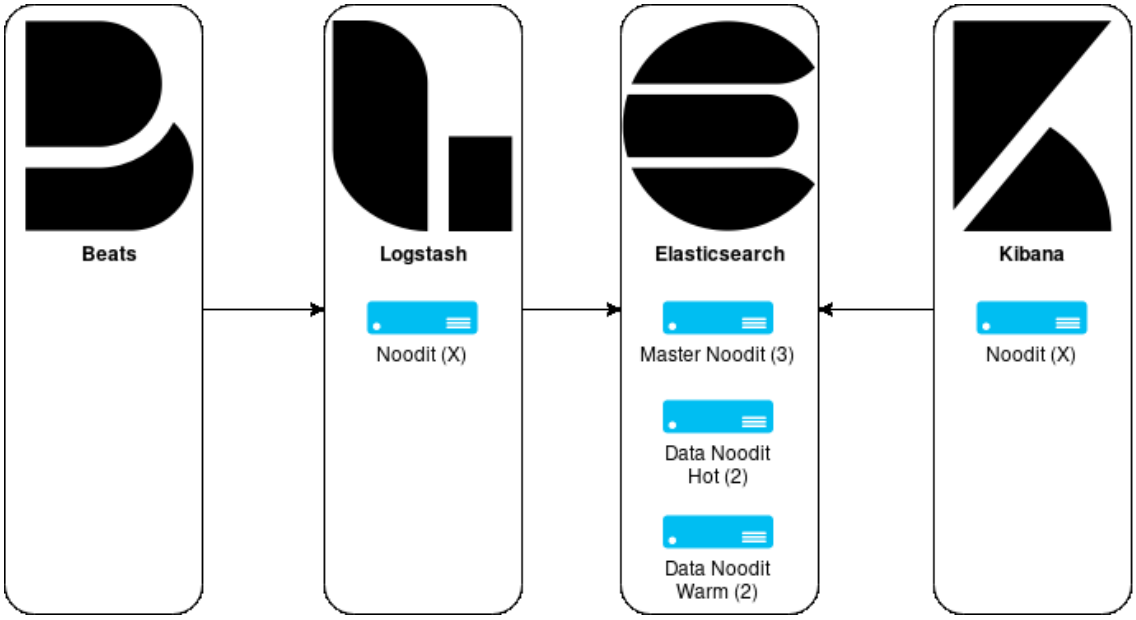
Resurssit

Elastic Stack -klusterin resurssien käyttö ja topologia riippuu tarpeista ja prosessoitavan datan määrästä. Esimerkiksi jos master- + data Elasticsearch-noodilla indeksoidaan ja analysoidaan suuria määriä dataa saattaa data instanssin resurssien käyttö haitata samalla noodilla olevaa master-instanssia (Node n.d.). Tällöin voidaan siirtyä jaetusta master- ja data-noodista topologiaan, jossa jokainen master- ja data-noodi on erillisellä koneella. Tarvittaessa datan prosessointi tehoa voidaan myös kasvattaa lisäämällä klusteriin data-noodeja, tai lisäämällä noodeille resursseja. Elasticsearchin käyttämä tallennustila vaikuttaa paljon Elasticsearchin nopeuteen, joten Elasticsearch hyötyy paljon siitä, että klusteri käyttää tallennustilana SSD:tä (Hardware n.d.). Lisäksi tulee huomioida palvelimien muistin määrä ja kuinka paljon muistia asetetaan JVM:n heapille. Elastic (n.d.) suosittelee, että JVM:n heapille asetetaan 50% palvelimen muistista. On kuitenkin huomioitava, että JVM käyttää pakattua oop:ia, jos JVM:lle on asetettu alle 32GB muistia. Tämä tarkoittaa sitä, että Elasticsearchille annettaessa alle 32GB muistilla on Elasticsearchin muistinkäyttö optimoidumpi. (Setting the heap size n.d.)

Pitkäaikainen tallennustila

Elastic Stack -klusterissa voidaan myös käyttää niin sanottua hot-warm toteutusta (kuvio 7), jolloin voidaan tallentaa dataa pidemmältä aikaväliltä. Tällöin

data noodeja on kahta tyyppiä: Hot ja Warm. Hot data-noodi käyttävät nopeaa SSD:tä ja toimii datan vastaanottajana ja indeksoijana. Datan iän ylittäessä määritetyn ajan se siirretään warm-noodeille, jotka käyttävät hidasta, mutta suuren tallennustilan kovalevyjä. Warm-nodeilla data säilytetään määritetyn säilytysajan loppuun. (Samir Bennacer 2015)



KUVIO 7. Esimerkki hot-warm toteutuksesta

Muut datalähteet

Elastic Stackiin voidaan liittää myös muita datan lähteitä ja palveluita tarpeiden vaatiessa, kuten esimerkiksi kappaleessa 2 esitetyissä Blizzardin ja Verizonin ratkaisuissa käytetyt valmiit ja itse rakennetut palvelut. Logstash voidaan esimerkiksi asettaa keräämään dataa lähteistä, joihin ei voida asentaa Beatseja, käyttämällä joitain Logstashin sisääntulo-lisäosista (Input plugins n.d.). Mikäli lähde ei itse tue kuorman tasausta voidaan myös lisätä näiden muiden lähteiden ja Logstashin väliin kuormantasaaja, joka jakaa datan Logstash-noodien välillä. Tilanteita varten, jolloin lokeja luodaan paljon, voidaan Elastic Stackiin lisätä datan puskurina toimiva palvelu Verizonin ratkaisun mukaisesti. Kuitenkin tulee huomioida, että Logstash sisältää oman puskurinsa (Execution model n.d.). Tästä syystä ei erillinen puskuri välttämättä ole tarpeellinen.

Asentaminen

Elastic Stack on mahdollista asentaa useilla eri asennustavoilla, jotka vaihtelevat Elastic Stackin komponenteista riippuen. Elasticin tarjoaman Elasticsearch palvelun lisäksi Elasticsearch voidaan asentaa pakatusta tar.gz tai .zip arkistosta, tai käyttäen valmiita sovelluspaketteja, esimerkiksi RPM- tai Docker-pakettia, sekä hyödyntäen hallintatyökaluja, esim. Puppet tai Ansible (Installing Elasticsearch n.d.). Kibana myös sisältyy Elasticin Elasticsearch palveluun ja voidaan asentaa samoilla tavoilla kuin Elasticsearch (Installing Kibana n.d.). Logstash taas ei sisälly Elasticin Elasticsearch palveluun, ja tarvitsee erikseen asennetun Java 8- tai Java 11-version toimiakseen, mutta muuten Logstash on mahdollista asentaa käyttäen samoja menetelmiä, kuten Elasticsearchin ja Kibanan kanssa (Installing Logstash n.d.). Beatsit voidaan asentaa Linux-käyttöjärjestelmään samalla tavalla, kuin muutkin Elastic Stackin komponentit, mutta Windows-käyttöjärjestelmälle ainut asentamisvaihtoehto on käyttää .zip arkistoja, jotka sisältävät Beatsin, sekä asennus- ja poisto -skriptit. (Install Filebeat n.d.; Install Winlogbeat n.d.)

Asetukset

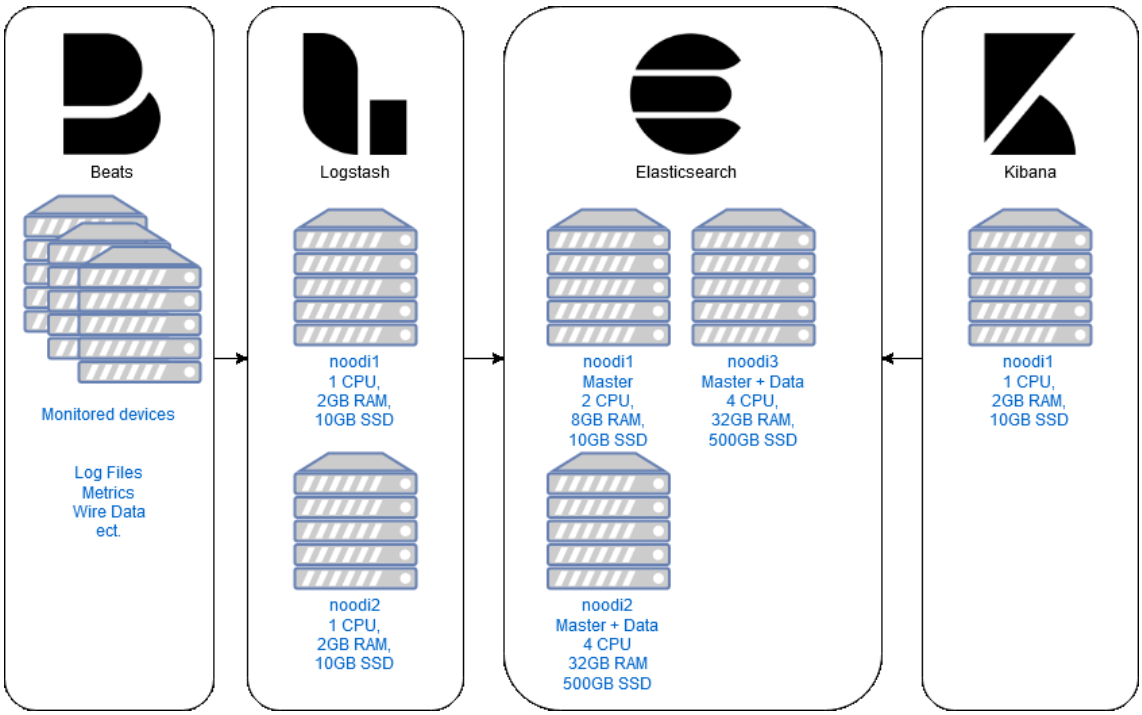
Elastic Stack käyttää pääsääntöisesti YAML-konfiguraatiotiedostoja asetuksien hallintaan, mutta poikkeuksellisesti Logstashin kanavat käyttävät JSON muotoa. Elastic Stack sisältää valmiit oletusasetukset, jotka vaativat vain vähän muokkausta. Konfiguraatiotiedostot sijaitsevat komponentin asennuskansiossa ja ovat nimetty komponentin nimellä, esimerkiksi ”elasticsearch.yml”. Lisäksi on erillinen konfiguraatiotiedosto JVM:n asetuksille: ”jvm.options”, josta voidaan määrittää mm. aikaisemmin mainittu JVM:n heapin koko. (Configure Filebeat n.d.; Configuring Elasticsearch n.d.; Configuring Kibana n.d.; Logstash Configuration Files n.d.)

4 LOKIEN ANALYSOINTIYMPÄRISTÖN TOTEUTUS

Capnovalle esitettiin suunnitelma ja jatkokehityssuunnitelma lokien tallennus ja analysointiympäristön rakentamisesta ja vaatimukset vikasietoisuudelle. Samalla määriteltiin klusterille asetettavat alustavat resurssit (taulukko 3). Klusterin resurssien käyttöä seurataan aktiivisesti ratkaisua rakennettaessa ja siihen siirtyessä, jolloin voidaan käyttötarpeen mukaan säätää resurssit vaaditulle tasolle tai muokata klusteria jatkokehityssuunnitelman mukaisesti lopullisiin tarpeisiin. Elastic Stack hyötyy paljon SSD:stä tallennustilana, joten jokainen noodi käyttää tallennustilana SSD:tä. Muistin kooksi määriteltiin 32GB, josta muistia skaalataan tarvittaessa. On kuitenkin huomioitava, että JVM:n heapin kooksi ei aseteta yli 32GB muistia, sillä yli 32GB:n jälkeen JVM ei voi enää käyttää pakattua oop:ia. Lisäksi päätettiin, että aluksi käytetään vain yhtä Kibana noodia, sillä Kibanan käyttökatkot eivät haittaa muun klusterin toimintaa. Täten topologia sisältää 2 Logstash noodia, 1 Elasticsearch master noodin, 2 Elasticsearch master ja data noodia ja yhden Kibana noodin (kuvio 8).

TAULUKKO 3. Klusterin alustavat resurssit

	CPU (säikeet)	**RAM (GB)**	**SSD (GB)**
Logstash 2x	1	2	10
Elasticsearch Master	2	8	10
Elasticsearch Data & Master 2x	4	32	500
Kibana	1	2	10



KUVIO 8. Elastic Stack -klusteri

Elastic Stack -klusteri ja lokien keräämisessä käytettävät Beatsit asennettiin suunnitelman pohjalta Capnovan järjestämiin palvelimiin. Klusterin palvelimet käyttävät Centos 8 käyttöjärjestelmää ja palvelimet, joilta lokeja kerätään käyttävät Centos, CloudLinux tai Windows käyttöjärjestelmää. Lisäksi hyödynnettiin Elastic Stackin valmiita oletuskonfiguraatioita, joita muokattiin tarpeiden mukaan.

4.1 Valmistelut

Elastic Stack -klusteri asennettiin Centos 8 käyttöjärjestelmälle, joten voitiin käyttää suoraan Elasticin tarjoamaa RPM pakettivarastoa (engl. repository). Pakettivarasto sisältää kaikki Elasticin tarjoamat RPM paketit. Pakettivaraston paketit ovat allekirjoitettu Elasticsearchin PGP-avaimella, joten lisättiin kyseinen avain käyttäen seuraavaa komentoa:

```
rpm --import https://artifacts.elastic.co/GPG-KEY-elasticsearch
```

Avaimen lisäämisen jälkeen järjestelmään lisättiin RPM pakettivarasto luomalla /etc/yum.repos.d/ kansioon elasticsearch.repo tiedosto (liite 1). Tämän jälkeen palvelimille siirrettiin Capnovan tarjoamat SSL sertifikaatit.

4.2 Elasticsearch

RPM-pakettivaraston lisäämisen jälkeen järjestelmille asennettiin Elasticsearch. Tämän jälkeen Elasticsearchiin voitiin konfiguroida alustava konfiguraatio muokkaamalla Elasticsearchin oletuskonfiguraatiota /etc/elasticsearch/elasticsearch.yml liitteessä 2 määritettyjen alustavien asetuksien mukaisesti. Lisäksi muutettiin jvm.options tiedostosta -Xms ja -Xmx optiot käyttämään 16GB järjestelmän muistia JVM:n heappiin. Tämän jälkeen Elasticsearch käynnistettiin. Elasticsearchin käynnistymisen jälkeen asetettiin muille Elasticsearch noodeille vastaava konfiguraatio. Lisäksi elastic-1 noodin konfiguraatiotiedostoon lisättiin noodin roolit määrittävät asetukset, ja määriteltiin noodi toimimaan ainoastaan master-noodina liitteessä 2 määritetyillä parametreilla. Master noodin Java optioiksi asetettiin Xms4g ja Xmx4g, master-noodin pienemmän muistin johdosta. Tämän jälkeen kaikki klusterin noodit käynnistettiin.

Elasticsearch asetettiin käyttämään salattua yhteyttä kommunikointiin. Ensin siirrettiin palvelimen SSL-sertifikaatti Elasticsearchin konfiguraatio kansioon. Elasticsearchin konfiguraatiotiedostoon asetettiin liitteessä 2 määritellyt SSL-konfiguraatio parametrit. Tämän jälkeen Elasticsearch uudelleen käynnistettiin ja asetettiin Elasticsearchin käyttäjille salasanat komennolla:

```
/usr/share/elasticsearch/bin/elasticsearch-setup-passwords interactive
```

Jonka jälkeen voitiin lisätä HTTPS konfiguraatio Elasticsearchiin liitteessä 2 määritetyillä xpack.security.http.ssl-parametreilla. Tämän jälkeen Elasticsearch uudelleen käynnistettiin ja asetettiin käynnistymään järjestelmän käynnistyessä.

4.3 Kibana

Kibana asennettiin palvelimelle käyttäen Elasticin RPM-pakettivarastoa. Tämän jälkeen siirrettiin Kibanan konfiguraatio kansioon palvelimen SSL sertifikaatti. Kibanan konfiguraatiotiedostoon: /etc/kibana/kibana.yml asetettiin liitteessä 3 määritetyt alustavat asetukset. Tämän jälkeen lisättiin Kibanaan Elasticsearchin käyttäjätunnus ja salasana seuraavilla komennoilla:

```
/usr/share/kibana/bin/kibana-keystore create
/usr/share/kibana/bin/kibana-keystore add elasticsearch.username
/usr/share/kibana/bin/kibana-keystore add elasticsearch.password
```

Tämän jälkeen Kibana käynnistettiin ja asetettiin käynnistymään järjestelmän käynnistymisen yhteydessä.

4.4 Logstash

Sillä Logstash paketti ei sisällä Javaa, piti Java asentaa Logstashin lisäksi. Noodin nimi ja prosessointijonon tyyppi määritettiin Logstashille asettamalla tiedostoon /etc/logstash/logstash.yml liitteessä 4 olevat alustavat asetukset. Tämän jälkeen luotiin /etc/sysconfig/logstash tiedosto, johon asetetaan “LOGSTASH_KEYSTORE_PASS” muuttuja, jota käytetään Logstashin käyttämän Elasticsearchin käyttäjänimen ja salasanan salaamiseen. Tämän jälkeen Logstashille annettiin Elasticsearchin Logstash käyttäjän tunnukset:

```
/usr/share/logstash/bin/logstash-keystore --path.settings /etc/logstash
create
/usr/share/logstash/bin/logstash-keystore --path.settings /etc/logstash
add ES_USER
/usr/share/logstash/bin/logstash-keystore --path.settings /etc/logstash
add ES_PWD
```

Logstashin pääkanava sijaitsee oletuksena /etc/logstash/conf.d/ kansiossa josta, Logstash prosessoi järjestyksessä tiedostot. Luotiin 01-input.conf tiedosto, johon

asetettiin TCP sisääntulo liitteen 5 mukaisesti ja 30-elasticsearch-output.conf, johon asetettiin Elasticsearch ulostulo liitteen 6 mukaisesti. Tämän jälkeen Logstash käynnistettiin ja asetettiin käynnistymään järjestelmän käynnistyksen yhteydessä.

4.5 Beats

Beatsit sisältyvät samaan pakettivarastoon kuin muut Elastic Stackin komponentit, joten kappaleen 4.1 pakettivarastoa voitiin suoraan hyödyntää asentamaan kaikki Linux järjestelmän tarvitsemat Beatsit. Luotiin valmiiksi Elasticsearchiin Beatseille käyttäjät. Esimerkki käyttäjäroolien oikeuksista taulukossa 4.

Taulukko 4. Roolien oikeudet

	Klusterin oikeudet	Indeksin oikeudet
setup	monitor, manage_ilm, manage_ml, manage_index_templates	manage, read
monitor	monitor	
writer	monitor, read_ilm, manage_pipeline	view_index_metadata, index, create_index
reader		read

4.5.1 Filebeat

Filebeat asennettiin valvottavalle laitteelle aikaisemman mukaisesti RPM pakettivarastosta. Testiympäristössä huomattiin, että Beatsien moduulit toimivat parhaiten lähettämällä datan suoraan Elasticsearchiin, eikä Logstashin kautta. Muokataan Filebeatin konfiguraatio tiedostoon /etc/filebeat/filebeat.yml palvelimet, johon dataa lähetetään, protokolla, jota käytetään lähettämiseen, sekä käyttäjäni-

melle ja salasanalle muuttuja liitteen 7 asetuksien mukaisesti. Tämän jälkeen luotiin Filebeatille keystore ja lisättiin Elasticsearchin Filebeat käyttäjä ja salasana siihen:

```
filebeat keystore create
filebeat keystore add ES_USER
filebeat keystore add ES_PWD
```

Lisäksi otettiin käyttöön halutut Filebeatin moduulit, esimerkiksi elasticsearch ja system moduuli:

```
filebeat modules enable elasticsearch system
```

Tämän jälkeen Filebeat käynnistettiin ja asetettiin käynnistymään järjestelmän käynnistymisen yhteydessä.

4.5.2 Winlogbeat

Sillä Winlogbeat on Windowsille kohdennettu, sen asentaminen poikkeaa muista Beatseista. Winlogbeat asennettiin koneelle lataamalla asennustiedosto osoitteesta: https://artifacts.elastic.co/downloads/beats/winlogbeat/winlogbeat-7.4.1-windows-x86_64.zip. Tämän jälkeen purettiin asennustiedoston sisältö C:/Program Files/Winlogbeat/ kansioon. Asetettiin Winlogbeatin asetukset puretusta kansiosta muokkaamalla winlogbeat.yml tiedosto aikaisemman Filebeat -konfiguraation mukaan (liite 7). Tämän jälkeen luotiin Winlogbeatille keystore ja asetettiin käyttäjänimi ja salasana:

```
.\winlogbeat.exe keystore create
.\winlogbeat.exe keystore add ES_USER
.\winlogbeat.exe keystore add ES_PWD
```

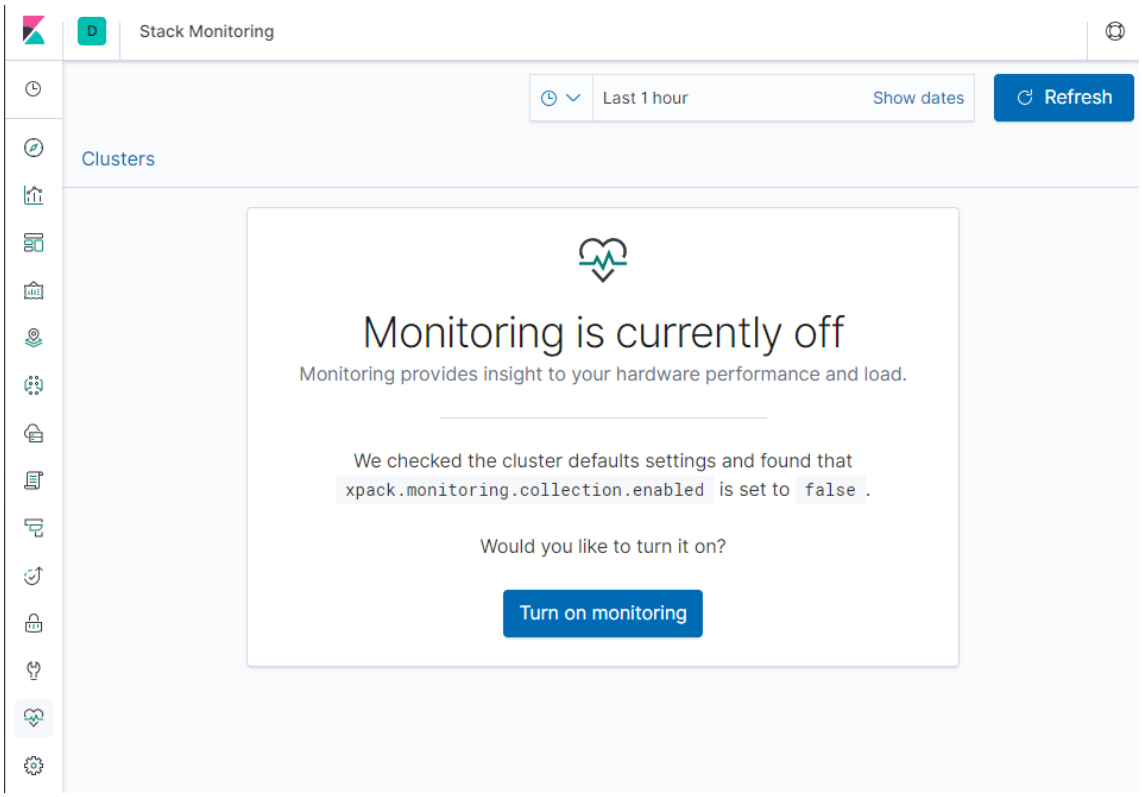
Winlogbeat asennettiin Windows palveluksi ajamalla install-service-winlogbeat.ps1 skripti järjestelmänvalvojana. Tämän jälkeen käynnistettiin Winlogbeat prosessit ajamalla seuraava PowerShell komento:

Start-Service winlogbeat

Winlogbeat palvelu käynnistyy oletuksena järjestelmän käynnistymisen jälkeen.

4.6 Elastic Stackin valvonta

Sillä on tärkeä seurata Elastic Stack -klusterin tilaa, otettiin käyttöön klusterin valvonta. Elasticsearchin ja Kibanan valvonta aktivoitiin Kibanan kautta “Stack Monitoring” välilehdestä painamalla “Turn on monitoring” -nappia (kuva 10).

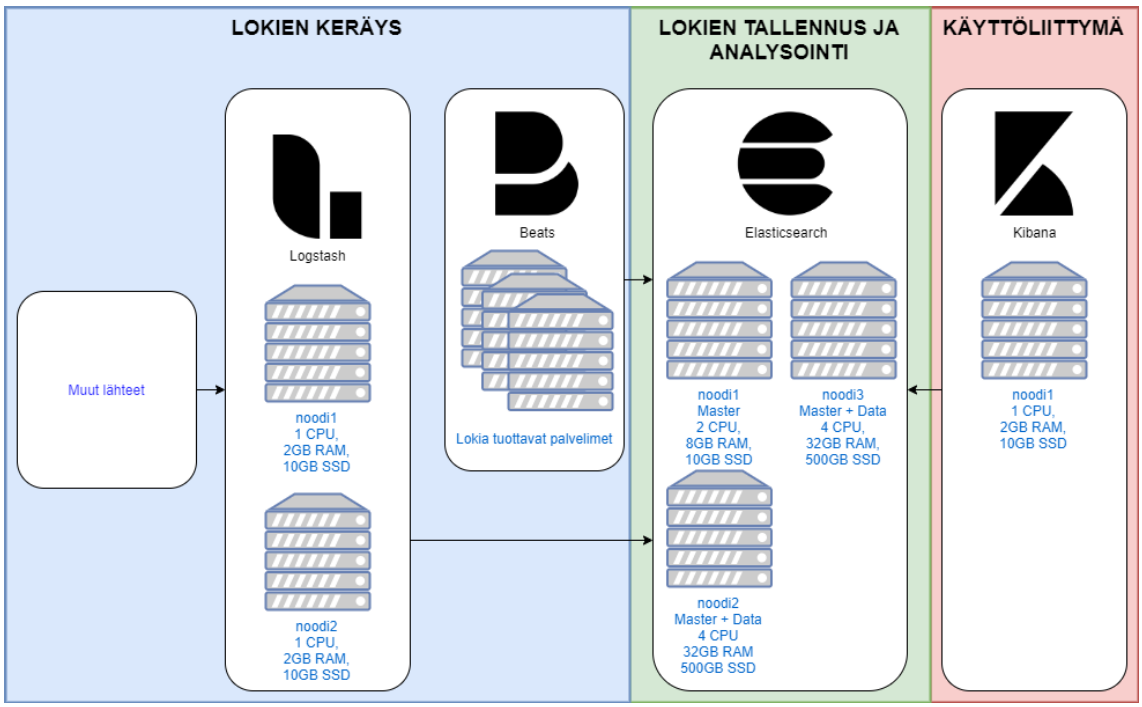


KUVA 10. Valvonnan käyttöönotto Kibanasta

Logstashin valvonta otettiin käyttöön lisäämällä Logstashin konfiguraatiotiedostoon liitteen 4 valvonnan parametrit. Lisäksi Beatsien valvonta otettiin käyttöön lisäämällä liitteessä 7 asetetut valvonnan asetukset beatsien konfiguraatiotiedostoihin. Tämän jälkeen koko klusterin toimintaa voitiin seurata “Stack Monitoring” -välilehdestä.

5 POHDINTA

Työn tarkoitus oli rakentaa Capnovan tarjoamaan ympäristöön Elastic Stackillä toteutettu keskitetty lokien keräys-, analysointi- ja tallennusjärjestelmä. Järjestelmän tarkoitus oli tehostaa lokien analysointia ja mahdollistaa lokien visualisoiminen. Järjestelmäkokonaisuus saatiin toteutettua ja vastaa määritettyjä tavoitteita, vaikka alkuperäisestä suunnitelmasta poikettiin ja määritettiin Beatsit lähettämään kerätyt lokit suoraan Elasticsearchiin. Järjestelmällä voidaan tallentaa valmiiksi hakuja, visualisointeja ja dashboardeja, jotka nopeuttavat datan analysointia verrattuna aikaisempaan syslogin kanssa käytettyyn manuaaliseen prosessiin. Lisäksi visualisoinnit ja dashboardit mahdollistavat tehokkaan yleiskatsauksen järjestelmän kokonaiskuvasta ilman manuaalista hakemista. Esimerkiksi järjestelmällä voidaan asettaa kartalle palvelimille kohdentuva liikenne. Tämä mahdollistaa sen, että järjestelmänvalvojat näkevät nopeasti mistä maasta yrityksen järjestelmiin kohdentuva liikenne tulee ja onko siinä poikkeamia, jotka tarvitsevat mahdollisesti toimenpiteitä. Lisäksi lokijärjestelmä mahdollistaa useiden järjestelmien lokien yhdistäminen ja analysoimisen yhtäaikaisesti, joka ei ollut mahdollista aikaisemmin käytössä olevalla järjestelmällä. Kuviossa 9 on esitettynä lopullinen järjestelmäkokonaisuus. Sillä Capnova oli määritellyt ennen työn alkua, että järjestelmä toteutetaan Elastic Stackillä, en nähnyt tarpeelliseksi vertailla Elastic Stackiä muiden lokien tallennusjärjestelmien kanssa.



KUVIO 9. Toteutettu lokienhallinta ratkaisu

Vaikka työ täyttää halutut vaatimukset, on siinä vielä mielestäni paljon kehitettävää. Esimerkiksi suunnittelun yhteydessä esitetty Hot-Warm infrastruktuurin toteutus, jolla voitaisiin tallentaa lokeja paljon pidempään. Lisäksi työssä ei otettu kantaa mitä lokeja kerätään, kauanko niitä tallennetaan ja miten lokien automaattinen elinkaaren hallinta toteutetaan. Mielestäni myös järjestelmän tietoturvaa kannattaa tarkastella tarkemmin, sillä esimerkiksi järjestelmänvalvojien ja käyttäjien oikeuksia lokijärjestelmään ei määritelty esimerkiksi lokien luvattoman poistamisen estämiseksi. Ehdottaisin jatkotutkimuksen kohteeksi, kuinka Elastic Stack lokienhallinta ratkaisu integroituu valvonta- ja hälytysjärjestelmiin, kuten Icinga ja Nagios tai voiko Elastic Stack korvata kyseiset järjestelmät keskittäen infrastruktuurin valvonnan yhteen palveluun.

LÄHTEET

Adding nodes to your cluster. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/elasticsearch/reference/current/add-elasticsearch-nodes.html

Backing, Joshua. 2016. Every Shard deserves a home. Luettu 11.10.2019. https://www.elastic.co/blog/every-shard-deserves-a-home

Basic Concepts. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/elasticsearch/reference/6.2/_basic_concepts.html

Beats overview n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/beats/libbeat/current/beats-reference.html

Bennacer, Samir. 2015. “Hot-Warm” architecture. Luettu 13.10.2019. https://www.elastic.co/blog/hot-warm-architecture

Burkhart & Warneck. 2018. Watching Overwatch at Activision Blizzard. Luettu 09.10.2019. https://www.elastic.co/elasticon/conf/2018/sf/watching-overwatch-at-activision-blizzard

Configure Filebeat. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-configuration.html

Configuring Elasticsearch. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/elasticsearch/reference/current/settings.html

Configuring Kibana. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/kibana/current/settings.html

Creating a Visualization. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/createvis.html

Dashboard. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/dashboard.html

Deploying and Scaling Logstash. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/logstash/current/deploying-and-scaling.html

Dev Tools. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/devtools-kibana.html

Discover. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/discover.html

Elastic Stack subscriptions. n.d. Luettu 11.10.2019. https://www.elastic.co/subscriptions

Execution model. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/logstash/current/execution-model.html

Filebeat overview. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/beats/filebeat/current/filebeat-overview.html

Hardware. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/elasticsearch/guide/current/hardware.html

How Filebeat works. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/beats/filebeat/current/how-filebeat-works.html

Input plugins. n.d. Luettu 14.10.2019. https://www.elastic.co/guide/en/logstash/current/input-plugins.html

Install Filebeat. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/beats/filebeat/7.4/filebeat-installation.html

Install Winlogbeat. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/beats/winlogbeat/7.4/winlogbeat-installation.html

Installing Elasticsearch. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/elasticsearch/reference/7.4/install-elasticsearch.html

Installing Kibana. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/kibana/7.4/install.html

Installing Logstash. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/logstash/7.4/installing-logstash.html

Introduction. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/introduction.html

Kibana Guide. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/index.html

Logs. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/xpack-logs.html

Logstash Configuration Files. n.d. Luettu 12.10.2019. https://www.elastic.co/guide/en/logstash/current/config-setting-files.html

Logstash Introduction n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/logstash/current/introduction.html

Node. n.d. Luettu 14.10.2019. https://www.elastic.co/guide/en/elasticsearch/reference/current/modules-node.html

Saving searches. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/save-open-search.html

Setting the heap size. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/elasticsearch/reference/current/heap-size.html

SIEM UI. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/siem/guide/7.4/siem-ui-overview.html

Stack Monitoring. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/xpack-monitoring.html

The Elastic Stack-powered evolution of Verizon Wireless for a better log analytics solution. n.d. Luettu 09.10.2019. https://www.elastic.co/customers/verizon-wireless

Valtionhallinnon tietoturvallisuuden johtoryhmä. 2009. Lokiohje. Luettu 24.10.2019. https://www.vahtiohje.fi/c/document_library/get_file?uuid=775179cb-6c54-4dfb-b65d-e925d47c61d2&groupId=10229

Visualize. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/kibana/current/visualize.html

We Opened X-Pack. n.d. Luettu 11.10.2019. https://www.elastic.co/what-is/open-x-pack

Winlogbeat Overview. n.d. Luettu 11.10.2019. https://www.elastic.co/guide/en/beats/winlogbeat/current/_winlogbeat_overview.html

LIITTEET

Liite 1. elasticsearch.repo

```
[elasticsearch-7.x]
name=Elasticsearch repository for 7.x packages
baseurl=https://artifacts.elastic.co/packages/7.x/yum
gpgcheck=1
gpgkey=https://artifacts.elastic.co/GPG-KEY-elasticsearch
enabled=1
autorefresh=1
type=rpm-md
```

Liite 2. Elasticsearchin oletuskonfiguraation muutokset

```
# Klusterin alustavat asetukset
cluster.name: elastic-logs
node.name: elastic2
network.host: 0.0.0.0
http.port: 9200
discovery.seed_hosts: [“elastic2.esimerkki.com”, “elas-tic3.esimerkki.com”]
cluster.initial_master_nodes: ["elastic-2"]
http.cors.enabled: true

# Master-noodin roolien määrittely
node.master: true
node.voting_only: false
node.data: false
node.ingest: false
node.ml: false
cluster.remote.connect: false

# Klusterin välisen liikenteen salaus
xpack.security.enabled: true
xpack.security.transport.ssl.enabled: true
xpack.security.transport.ssl.key: /etc/elasticsearch/elastic.key
# Palvelimen koko sertifikaatti
xpack.security.transport.ssl.certificate: /etc/elasticsearch/elastic.crt
xpack.security.transport.ssl.certificate_authorities: [ "/etc/elasticsearch/ca.crt" ]
# Jos ei käytetä sertifikaattia, joka sisältää palvelimen IP:n tarkistetaan vain ser-
tifikaatti.
xpack.security.transport.ssl.verification_mode: certificate

# HTTPS käyttöönotto
xpack.security.http.ssl.enabled: true
xpack.security.http.ssl.key: /etc/elasticsearch/elastic.key
xpack.security.http.ssl.certificate: /etc/elasticsearch/elastic.crt
```

Liite 3. Kibanan oletuskonfiguraation muutokset

```
# Alustavat Kibanan asetukset
server.port: 5601
server.host: "0.0.0.0"
server.name: "kibana1"

# Yhdistetään Kibana Elasticsearchin data-noodeihin
elasticsearch.hosts: ["https://elastic2.esimerkki.com:9200", "https://elastic3.esi-
merkki.com:9200"]

# Otetaan HTTPS käyttöön
server.ssl.enabled: true
server.ssl.certificate: /etc/kibana/kibana.crt
server.ssl.key: /etc/kibana/kibana.key
```

Liite 4. Logstashin oletuskonfiguraation muutokset

```
# Alustavat Logstashin asetukset
node.name: logstash1
# Käytetään palvelimen levyä puskurina muistin sijasta, jotta dataa ei menetettäisi kaatumisien tms. johdosta.
queue.type: persisted

# Valvonnan käyttöönotto
xpack.monitoring.enabled: true
xpack.monitoring.elasticsearch.username: ${ES_USER}
xpack.monitoring.elasticsearch.password: ${ES_PWD}
xpack.monitoring.elasticsearch.hosts: ["https://elastic2.esimerkki.com:9200",
"https://elastic3.esimerkki.com:9200"]
```

Liite 5. 01-input.conf

```
input {
  tcp {
    port => 5000
  }
}
```

Liite 6. 30-elasticsearch-output.conf

```
output {
  elasticsearch {
    hosts => ["https://elastic2.esimerkki.com:9200", "https://elastic3.esi-
merkki.com:9200"]
    ssl => true
    user => "${ES_USER}"
    password => "${ES_PWD}"
  }
```

Liite 7. Beatsien oletuskonfiguraatioiden muutokset

```
# Alustavat beatsien asetukset
hosts: [“https://elastic2.esimerkki.com:9200”, “https://elastic3.esi-
merkki.com:9200”]
protocol: “https”
username: ${ES_USER}
password: ${ES_PWD}

# Valvonnan käyttöönotto
monitoring.enabled: true
monitoring.elasticsearch:
```