jamk

Boss Of The SOC versio 3 Elastic Stack toteutus ja ohjeistus

Juho Pyykkinen

Opinnäytetyö, AMK
Toukokuu 2021
Tietojenkäsittely ja tietoliikenne
Insinööri (AMK), Tieto- ja viestintätekniikka

jamk | Jyväskylän ammattikorkeakoulu
University of Applied Sciences

jamk

Kuvailulehti

Pyykkinen, Juho

Boss of The SOC versio 3 Elastic Stack toteutus ja ohjeistus

Jyväskylä: Jyväskylän ammattikorkeakoulu. Toukokuu 2021, 81 sivua.

Tietojenkäsittely ja tietoliikenne. Tieto- ja viestintätekniikan tutkinto-ohjelma. Opinnäytetyö AMK.

Julkaisun kieli: suomi

Verkkojulkaisulupa myönnetty: Kyllä

Tiivistelmä

Toimeksiantaja JYVSECTEC tarvitsi Splunkin Boss Of The SOC tyylisen harjoituksen ja koulutusalustan Elastic Stackille. Toimeksiantaja halusi ohjeistuksen tämänkaltaiseen harjoitukseen valmiina olemassa olevasta realistisestä datasetistä, jolla toimeksiantaja voisi kouluttaa ja opettaa Elastic Stackillä data-analysointia. Tähän valittiin Boss Of The SOC versio 3 dataset ja siihen liittyvä kysymyssarja.

Ohjeistuksen teko jakaantui kolmeen osaan. Osassa yksi ulostuotiin Boss Of The SOC versio 3:n data Splunk Enterprisellä sellaiseen muotoon, jota Elastic Stack pystyi lukemaan. Osassa kaksi sisääntuontiin data Elastic Stackiin ja kolmannessa osassa tehtiin ohjeistus Boss Of The SOC version 3:n tekoon.

Boss Of The SOC versio 3 harjoituksesta tehtiin ohjeistus, jota seuraamalla pystyy Elastic Stackillä data-analyysiä opetteleva käyttäjä tekemää harjoituksen kokonaan ohjeistusta käyttäen tai katsomalla siitä apua. Ohjeistuksen lisäksi tehtiin vihjeitä jokaiseen kysymykseen ohjeistuksen pohjalta, jotka helpottaisivat Boss Of The SOC versio 3 harjoituksen tekoa, jos ei ole haluttua käyttää ohjeistusta. Sekä ohjeistuksekseen tehtiin kohta, jossa neuvotaan harjoituksen tekijää työkaluista ja hakutavoista Elastic Stackissä data-analysointi mielessä. Tämän lisäksi oli tarvetta muuttaa kahta kysymystä ja jättää pois yksi kysymys Boss Of The SOC versio 3 kysymyksistä

Ohjeistuksesta toteutettiin testi, jonka tarkoitus oli saada ohjeistuksesta ja ohjeistuksen laadusta ulkopuolisten henkilöiden mielipiteitä. Testissä testin tekijä teki Boss Of The SOC versio 3 harjoituksen seuraamalla suoraan ohjeistusta. Testiin saatiin vain yksi tekijä toimikeksiantajan kiireitten takia, mutta testistä saatiin silti hyvää palautetta ja tämän testin tuloksien perusteella pystyttiin tekemään ohjeistuksiin muutoksia ja parannuksia.

Tuloksena toimeksiantaja sai haluamanlaisensa realistisen Elastic Stack data-analysointi harjoituksen käyttäen Boss Of The SOC versio 3:a ja sivutuotteena käyttöjärjestelmä imagen, jossa on kaikki tarvittava valmiina harjoituksen tekoon.

Avainsanat (asiasanat)

Elastic Stack, Boss Of The SOC, harjoitukset, SIEM

Muut tiedot (salassa pidettävät liitteet)

Liite 4 on salassa pidettävä, ja ne on poistettu julkisesta työstä. Salassapidon peruste on Julkisuuslain 621/1999 24§, kohta 21. Salassapitoaika on kolme (3) vuotta, salassapito päättyy 01.06.2024.

jamk

Description

Pyykkinen, Juho

Boss Of The SOC version 3 Elastic Stack implementation and instructions

Jyväskylä: JAMK University of Applied Sciences, May 2021, 81 pages.

Information and Communication Technologies. Degree programme in Information and Communications Technology. Bachelor's thesis.

Permission for web publication: Yes

Language of publication: Finnish

Abstract

The customer JYVSECTEC needed exercise and training platform like Splunk's Boss Of The SOC for Elastic Stack. The client wanted instruction style walkthrough for this kind of exercise from an existing realistic dataset with which the customer could train and teach data analysis with Elastic Stack. For this, the Boss Of The SOC version 3 dataset and questions related to it were selected.

The creation of the instructions was divided into three parts. In first part, Boss Of The SOC version 3 dataset was outputted with Splunk Enterprise in a format that Elastic Stack could read it. In second part, data was imported into the Elastic Stack, and in third part, instruction style walkthrough were made for Boss Of The SOC version 3.

Instructions were made for the Boss Of The SOC version 3 exercise, which allows the user learning data analysis on the Elastic Stack to complete the exercise using the instruction style walkthrough or by looking that for help. In addition to the walkthrough, tips were made for each question based on the walkthrough, which makes Boss Of The SOC version 3 exercise easier to do the if it is not wanted to use the instruction. Also, a guide was made to advise and help the exercise maker on tools and search methods needed to make the exercise on Elastic Stack. In addition to this, two questions from Boss Of The SOC version 3 a needed changing and one question from Boss Of The SOC version 3 was removed form exercise.

Walkthrough was tested on the purpose of which to obtain the opinions of outsiders on the walkthrough and the quality of the walkthrough. In the test, the tester did the Boss Of The SOC version 3 exercise by directly following the instruction style walkthrough. Only one tester was obtained for the test due to the client's being busy, but good feedback was still received from the test, and based on the results of this test, changes and improvements were made to the instruction style walkthrough.

As a result, the customer received a realistic Elastic Stack data analysis exercise using Boss Of The SOC as the way they wanted and as by-product system image with everything that they need to do the exercise.

Keywords/tags (subjects)

Elastic Stack, Boss Of The SOC, exercises, SIEM

Miscellaneous (Confidential information)

Appendix 4 is secret and is removed from publicly available thesis. The basis for keeping it secret is Act on the Openness of Government Activities 621/1999 24§, paragraph 21. The period of secrecy is three (3) years, secrecy ends 01.06.2024.

Sisältö

Kuviot

Taulukot

Lyhenteet

ACL	Access-control List
AMI	Amazon Machine Image
API	Application Programming Interface
AWS	Amazon Web Service
BOTS	Boss Of The SOC
BOTSv3	Boss Of The SOC version 3
CIFS	Common Internet File System
CPU	Central processing unit
CSV	Comma-separated Values
DNS	Domain Name System
EC2	Elastic Compute Cloud
ECS	Elastic Container Service
ELK	Elasticsearch, Logstash and Kibana
FQDN	Fully Qualified Domain Name
HTTP	Hypertext Transfer Protocol
IAM	Identity and Access Management
ID	Identity Document
JSON	JavaScript Object Notation
KQL	Kibana Query Language
MFA	Multi-factor Authentication
NVM	Network Visibility Module
RGCE	Realistic Global Cyber Environment
REST	Representational State Transfer
RSS	Really Simple Syndication
S3	Simple Storage Service
SDK	Software Development Kit
SEM	Security Event Management
SEP	Symantec Endpoint Protection
SEPM	Symantec Endpoint Protection Manager
SIEM	Security Information and Event Management
SIM	Security Information Management
SLA	Service-level Agreement
SMTP	Simple Mail Transfer Protocol
SOC	Security Operations Center
XML	Extensive Markup Language

1 Johdanto

Nykyään on erilaisten kyberturvauhkien takia noussut tarvetta yhtiöille ottaa käyttöön Security Operations Center (SOC) tiimejä pienessä tai isossa muodossa. Nämä SOC tiimit sisältävät kyberturvallisuus ammattilaisia kuten analysoijia jotka, monitoroivat, ratkaisevat ja ennalta ehkäisevät isolla skaalalla kyberturvallisuus ongelmia organisaatioissa käyttäen erilaisia työkaluja. (WHAT IS A SECURITY OPERATIONS CENTER? WHY IS IT IMPORTANT? 2020.)

Tämä on nostanut tarvetta maailmalla sekä isoissa että pienissä organisaatioissa kouluttaa SOC henkilöstöä monitoroimaan, ratkaisemaan ja ennalta ehkäisemään kyberturvallisuustapauksia. Tähän kuuluu mm. koulutus työkalujen käyttöön ja kuinka niitä käytetään oikeassa maailmassa. (What is SOC Training? 2020.)

1.1 Toimeksiantajan esittely

Opinnäytetyön toimeksiantajana toimi JYVSECTEC (Jyväskylä Security Technology), joka on osa Jyväskylän ammattikorkeakoulun IT-instituuttia. JYVSECTEC on verkkosivujensa mukaan Suomen johtava puolueeton kyberturva tutkimus-, kehitys- ja harjoituskeskus. JYVSECTECin tarkoitus on auttaa asiakkaitaan valmistautumaan kyberuhkia vastaan sekä auttaa teknologista kehitystä. (JYVSECTEC overview n.d.)

JYVSECTEC tarjoaa erilaisia palveluita, jotka liittyvät kyberturvallisuuteen, näitä palveluita ovat FINCSC sertifikaatti, erilaiset kyberharjoitukset, koulutusta tieto- ja kyberturvallisuuteen, järjestelmien ja ohjelmien haavoittuvuuksiin ja turvallisuus puutteiden testausta sekä konsultointia eri osa-alueisiin tieto- ja kyberturvallisuus alalla. Tämän kaiken lisäksi he myös tutkivat ja kehittävät kyberturvallisuus projekteja. (JYVSECTEC n.d.)

JYVSECTECillä on myös kyberrata ympäristö nimeltä Realistic Global Cyber Environment (RGCE) tässä ympäristössä toteutetaan myös edellä mainittuja palveluita. RGCE:ssä on realistinen oikeasta maailmasta eristetty maailman laajuinen internet ympäristö ja tämä sisältää ”oikeita” organisaatio ympäristöjä. (JYVSECTEC cyber range n.d.)

1.2 Työn tausta ja tavoite

Opinnäytetyön alkuperäisenä ideana oli toimeksiantajan tarve saada Splunk Boss Of The SOC (BOTS) tyylisen harjoituksen ja koulutusalustan Elastic Stackille. Harjoituksen oli tarkoitus olla osana jonkinlaista kurssia tai harjoitusta tulevaisuudessa. Ideana oli käyttää valmista olemassa oleva realistista datasettiä, jossa olisi Elastic Stackillä data-analysointia mielessä pitäen hyödyllisiä tehtäviä. Ennen työn aloittamista selvisi toimeksiantajan kanssa, että BOTS oli ainoa hyvä olemassa oleva vaihtoehto

BOTS harjoituksesta haluttiin tehdä ohjeistus, jota seuraamalla pystyy Elastic Stackillä data-analyysiä opetteleva käyttäjä tekemää harjoituksen kokonaan ohjeistusta käyttäen tai katsomalla siitä apua. Ohjeistuksen lisäksi haluttiin vihjeitä jokaiseen kysymykseen, jotka helpottaisivat BOTS harjoituksen tekoa, jos ei ole haluttu käyttää ohjeistusta. Ja tapauksessa, jossa vihjeet eivät riitä, voidaan niistä tarvittaessa hypätä ohjeistukseen, tällä tavalla voitaisiin ohjeistuksen ja vihjeitten välillä tarvittaessa vaihdella.

Toimeksiantaja halusi myös lopputulos dokumentin muotoon, josta kurssia tai harjoitusta pitävän henkilön on helppoa muokata, ottaa kohtia tai rajata pois kohtia, jotta hän saisi harjoituksesta haluamanlaisensa.

Työn venymisen ja aikamäärän nousemisen takia päätettiin toimeksiantajan kanssa rajata ohjeistus 200-sarjaan ja jätettiin ohjeistus tekemättä 300-sarjaan. BOTSv3 200-sarja nähtiin tarpeeksi laajaksi yksinään harjoituksen kannalta.

1.3 Aiheeseen liittyvät aiemmat työt

Opinnäytetyöt

Theseusksesta löytyi opinnäytetöitä opinnäytetyön teon aikana, joissa tehtiin Elastic Stackiin erilaisia lokien hallinta- ja/tai monitorointijärjestelmiä. Sekä opinnäytetöitä, jossa Elastic Stackiä käytettiin tallennuspaikkana datalle, jota tuli työn tärkeämmästä osasta. Elastic Stackille ei löytynyt yhtään suoraa ohjeistusta ja vain yksi, jossa sivutaan lokianalyysiä. Tämä opinnäytetyö on Jouni Mikkolan (2020) tekemä ”Lokianalyysi ja valvonta ELK-järjestelmää hyödyntäen”.

Theseuksesta ei löydy harjoitus tyylistä tai suoraa harjoitusta Elastic Stackin käyttöä opettelevalle tai harjoittelevalle henkilölle.

Theseus haut näkyvät taulukossa 1.

Taulukko 1. Theseus opinnäytetyö haut

Haku	Tulosten määrä	Relevanttisuus
BOTS	2380	Ei mahdollista selvittää
BOTS AND splunk	3	Ei Boss Of The SOC liittyviä
"boss of the soc"	1	Vain maininta
BOTSv1 OR "BOTS version 1"	0	
BOTSv2 OR "BOTS version 2"	0	
BOTSv3 OR "BOTS version 3"	1	Ei Boss Of The SOC liittyvä
BOTSv4 OR "BOTS version 4"	0	
"Elastic Stack"	40	Ei Boss Of The SOC liittyviä. Ei ohjeistuksia tai harjoituksia Elastic Stackillä data-analysointiin.
”ELK Stack”	47	Ei Boss Of The SOC liittyviä. Ei ohjeistuksia tai harjoituksia Elastic Stackillä data-analysointiin.
"Elastic Stack" OR "ELK Stack" AND harjoitus	0	

Jyväskylän yliopiston opinnäytetyö haut näkyvät taulukossa 2.

Taulukko 2. Jyväskylän yliopisto opinnäytetyö haut

Haku	Tulosten määrä	Relevanttisuus
BOTS	0	
"boss of the soc"	0	
BOTSv1 OR "BOTS version 1"	0	
BOTSv2 OR "BOTS version 2"	0	
BOTSv3 OR "BOTS version 3"	0	
BOTSv4 OR "BOTS version 4"	0	
"Elastic Stack" OR "ELK Stack"	0	

Internet harjoitukset

Opinnäytetyön aikana internetistä ei löydy haluttavaa harjoitusta Elastic Stackille. Sieltä löytyy ohjeita, kuinka käyttää Elastic Stackiä ja myös, kuinka käyttää sitä data-analyysissä, mutta mitään niistä ei ole nähty sopivaksi toimeksiantajan käyttötarkoitukseen, joka on saada BOTS tyylinen harjoitus Elastic Stackille.

Esimerkiksi Dotan Horovits (2020) oli kirjoittanut pitkän ja perusteellisen käyttöohjeen Elastic Stackille logz.io sivustolle, mutta siinä ei ole harjoitusta vaan kerrotaan miten Elastic Stack toimii ja mitä sillä voi tehdä.

Toisena esimerkkinä Ben Andersen-Waine (2015) oli tehnyt pienen harjoituksen Elastic Stackin käyttöön. Tämä ei ole myöskään halutun tyylinen ja sen viimeisin versio on 25.6.2015 ja se käyttää vanhaa versiota Elastic Stackistä.

BOTSv3 olemassa olevat ohjeistukset

BOTSv3 oli olemassa jo valmiina muutamia ohjeistuksia ja teon aika tehtyjä blogi kirjoituksia, kun opinnäytetyötä tehtiin. Nämä tietenkin olivat kaikki Splunk Enterpriseä käyttäen tehtyjä. Tässä kohdassa käydään läpi niitä ja niitten muita ongelmia sen lisäksi että ne eivät ole tehty Elastic Stackillä.

Ellis Stannard (2020) on tehnyt ohjeistuksen, joka on tähän käyttötarkoitukseen hieman epäpätevä. Ohjeistuksesta on tehty vain osa, joka ovat kysymykset 200–208 kun 200-sarjan kysymykset menevät 225 asti. Ohjeistuksessa oli muutenkin haku sanat laitettu minimaalisilla selityksillä.

James Gibbinsin (2020) ohjeistus kattoi koko BOTSv3 harjoituksen myös 300-sarjan, mutta ohjeistuksena se oli aika suppea. Siinä oli käytetyt haut ja selitetty miksi tämä haku on tehty muutamalla sanalla.

Chris Longin (2020) ohjeistus tyyli oli hyvä, siinä selitetään, miten hakuihin päästiin, mutta nämä olivat silti aika suppeat. Tavat, jolla Chris pääsi vastauksiin olivat joissakin kohdissa aika huonoja tai

epäkäytännöllisiä sekä muutamassa kysymyksessä Chris ei päässyt vastaukseen tai vastaus oli väärin.

Chris oli tehnyt cwo1010 YouTube käyttäjätilillä ohjeistuksen BOTSv3:een (Splunk BOTS - Boss Of The SOC (v3) Walkthrough & Analysis 2020). Ohjeistuksesta puuttuivat kysymykset 216–218, 224 ja 225. Ohjeistuksessa tuli hyvin selville, miten hakuihin oli päästy ja tavat olivat pääasiassa hyviä, mutta Chris käytti paljon hyväkseen Splunk ominaisuuksia, joita ei voi käyttää Elastic Stackissä. Sekä koska ohjeistus oli videomuodossa se ei ole käytännöllinen toimeksiantajan tarkoitukseen.

2 SIEM

Security Information and Event Management (SIEM) on yhdistelmä Security Information Management (SIM) ja Security Event Management (SEM) järjestelmiä. SIEM on turvallisuus ratkaisu, joka auttaa organisaatiota havaitsemaan turvallisuusuhkia ja haavoittuvuuksia IT-infrastruktuurissa käyttäen automaatioituja prosesseja sekä joissakin tapauksissa tekoäly ratkaisuja. (What is SIEM? n.d.)

SIEM kerää ja tallentaa tietoja erilaisista tapahtumista yrityksen IT-ympäristön laitteista ja ohjelmistoista. Näitä laitteita ovat esimerkiksi päätelaitteet, verkkolaitteet, pilvipalvelut, palomuurit ja tietoturvajärjestelmät. SIEM analysoi tätä dataa verraten sitä asetettuihin turvallisuus sääntöihin ja turvallisuus säädöksiin. Tästä SIEM havaitsee mahdollisia riskejä ja ongelmia organisaation järjestelmissä. Kun SIEM on huomannut, analysoinut ja raportoinut ongelman, se voidaan asettaa ilmoittamaan tästä IT-henkilöstöllä halutulla tavalla. (What Is Security Information and Event Management (SIEM)? n.d.)

SIEMeissä on monia valmistaja kohtaisia ominaisuuksia ja lisäosia, mutta pääasiassa ne kaikki sisältävät samat pääominaisuudet.

Lokidatan hallinta. SIEMit keräävät tapahtuminen tietoja eri lähteistä organisaatioitten verkossa. Tämmöistä dataa voi olla, vaikka erilaiset loki- ja virtausdata (eng. flow data) käyttäjistä, laitteistoista, ohjelmista, päätelaitteista, erilaisista pilvipalveluista, verkoista ja verkkoliikenteestä. Nämä

kaikki datat tallennetaan ja analysoidaan reaaliajassa. SIEMeihin voidaan yleensä myös syöttää dataa muista kyberturvallisuus ohjelmistoista omasta verkosta tai muualta maailmasta. (What is SIEM? n.d.)

Tapahtumakorrelaatio ja -analyysi. Datan analysoinnissa ja tapahtumakorrelaatiossa yritetään ymmärtää ja löytää turvallisuusuhkia ja haavoittuvuuksia kaavojen ja mallien avulla mahdollisimman nopeasti, jotta ne voidaan hoitaa pois nopeasti organisaation IT-ympäristöstä. (What is SIEM? n.d.)

Tapahtumien valvonta ja turvallisuushälytykset. Koska SIEMit ovat keräävät data yhteen järjestelmään, joka analysoi kaikkia asioita IT-infrastruktuurissa. Tällä voidaan asettaa SIEM tunnistamaan erilaisia tapahtumia järjestelmässä. Näihin tapahtumiin voidaan sitten reagoida monella eri tavalla, kuten lähettämällä vastaavalle IT-henkilöstölle ilmoituksen, kerätä tiedot SIEMin kojelaudalle ja joissakin tapauksissa SIEM tai siinä kiinni oleva järjestelmä voi hoitaa ongelman automaattisesti itse. (What is SIEM? n.d.)

Vaatimustenmukaisuus raportit (eng. Complience reports). SIEM järjestelmiä käytetään myös erilaisiin laillisten tai muitten vaatimustenmukaisuusraporttien tarvitsemien tietojen keräämiseen ja varmistamiseen. SIEMeillä voidaan yleensä tuottaa reaaliaikaista vaatimuksenmukaisuusraportteja eri asioista esimerkiksi SOX (Sarbenes-OxleyAct) ja GDPR (General Data Protection Regulation). Joissakin SIEMeissä on mahdollisuus tuottaa automaattisia raportteja näistä. (What is SIEM? n.d.)

2.1 Elastic Stack

Elastic Stack vanhalta nimeltään ELK Stack, joka tulee sanoista Elasticsearch, Logstash ja Kibana. Elastic Stackistä käytetään myös SIEM käytössä nimiä Elastic SIEM ja ELK SIEM.

Elastic Stack on SIEM tapauksessa interaktiivinen työtila kyberturvallisuustiimeille, jossa he voivat luokitella ja tehdä tutkintaa erilaisiin tapahtumiin. Elastic Stack on ohjelmisto, joka mahdollistaa analysoijan kerätä ja tallentaa todistusaineistoa kyberhyökkäyksestä tai muista tapahtumista aikajananäkymän avulla. Se antaa myös analysoijan lisätä huomautuksia, kommentoida ja jakaa löydöksiä muitten analysoijien kanssa. (Introducing Elastic SIEM n.d.)

Elastic Stack järjestelmä koostuu kolmesta eri osasta, Elasticsearch, joka toimii tietokantana ja jossa kaikki haku indeksointi ja analysointi tapahtuu, Logstash ja Beats, joilla tehdään tietojen keräämistä, -kokoamista ja -rikastamista sekä niiden tallentamista Elasticsearchiin. Ja viimeiseksi Kibana käyttöliittymä ohjelma, joka antaa käyttäjän interaktiivisesti tutkia ja visualisoida kerättyä dataa sekä monitoroida Elastic Stackin toimintaa. (What is Elasticsearch? n.d.)

Kuviossa 1 näkyy Elastic Stackin osat ja niiden tarkoitukset.

Kuvio 1. ELK SIEMin osat ja sen arkkitehtuuri (Horovits 2020)

Elastic Stackistä saa kuukausimaksullisia versioita eri hintaluokkaan ja eri ominaisuuksilla, pilvi versioina ja itse hallittavana. Sekä ilmainen itse hallittava versio. (Elastic pricing n.d.)

2.1.1 Elasticsearch

Elasticsearch on Elastic Stackin pääkomponentti, joka on hajautettu avoimen lähdekoodin haku ja analysointi työkalu eli tietokanta. Tietokantana Elasticsearchin tarkoitus on olla nopea ja sellainen, että siihen voidaan tallentaa ja hakea monenlaista dataa, mukaan lukien tekstillistä, numeraalista, paikkatieteellistä, jäsennettyä ja jäsentämätöntä. (What is Elasticsearch? n.d.)

Kun Elasticsearchiin lisätään dataa erilaisista lähteistä, sitä jäsennetään, normalisoidaan ja rikastetaan ennen kuin se indeksoidaan. Kun data on indeksoitu tietokantaan, sitä vasten voidaan tehdä erilaisia hakuja. Elasticsearchin indeksi on kasa dokumentteja, jotka liittyvät toisiinsa jollain tavalla, tämä tapa on yleensä datan alkuperään tai lisäämistapaan liittyvää. Elasticsearch tallentaa datan JSON-tiedosto formaattiin ja data tallennetaan käännettyyn indeksi (eng. inverted index) tietueeseen. Haun tapahtuessa käännetty indeksi listaa jokaisen uniikki sanan jokaisessa dokumentissa

tunnistaa nämä dokumentit hakutulokseksi. Tämä tekee Elasticsearchin hausta erittäin nopeata. (What is Elasticsearch? n.d.)

Elasticsearchia käytetään monenlaiseen eri tarkoitukseen sen nopeuden ja skaalautuvuuden takia sekä mahdollisuuden indeksoida erilaisia datatyyppejä. Elasticin verkkosivulla on listattu muutamia eri käyttötarkoituksia, jotka ovat ohjelmistohaku, verkkosivuhaku, yhtiöhaku, lokaus ja lokausanalysointia, infrastruktuurin ja konttinen monitorointia, ohjelmistojen suorituskyky monitorointia, paikkatieteellistä data-analysointia ja visualisointia, turvallisuusanalysointia ja yritysanalysointia esim. myynti tarkoituksessa. (What is Elasticsearch? n.d.)

2.1.2 Logstash

Logstash on yhdessä Beatsien kanssa Elastic Stackin datan keräys osa. Logstash on avoimen lähdekoodin datan keräys työkalu, jossa on reaaliaikainen putkitus mahdollisuus datalle. Logstash voi dynaamisesti kerätä dataa eri lähteistä ja normalisoida sen haluttuun pääte kohteeseen. Kaiken tyylisiä tapahtumia voidaan rikastaa ja muuntaa laajalla valikoimalla sisääntuonti-, suodatus- ja ulostuonti lisäosia sekä datan sisäänottoa on helpotettu ja yksinkertaistettu useiden natiivi koodeksien avulla. (Logstash introduction n.d.)

Sisääntuonti lähde lisäosat mahdollistavat Logstashin mahdollisuuden lukea tapahtumia erilaisista lähteistä. Muutamia virallisesti tuettuja sisääntuonti lähteitä ovat Beats, AWS CloudWatch, Elasticsearch, File ja HTTP. (Input plugins n.d.)

Suodatus lisäosien tarkoitus on prosessoida tapahtumia. Suodattimien käyttö on tapahtuma ja data typpistä johtuen monen kaltaista. Suodatus ei myöskään ole pakollista. Suodatin lisäosiin kuuluu muun muassa cidr, clone, json ja mutate. (Filter plugins n.d.)

Koodeksi lisäosat muuttavat tapahtumadatan esitys tapaa. Näihin lisä osiin kuuluvat cloudtrail, csv ja json. (Codec plugins n.d.)

Ulostuonti lisäosat lähettävät dataa haluttuun päätekohteeseen. Ulostuonti lisäosa on putken viimeinen osa. Lisäosiin kuuluvat mm. elasticsearch, email ja http. (Output plugins n.d.)

2.1.3 Beats

Beatsit ovat Elastic Stackin neljäs osa, joka ei ole historiallisesti kuullut siihen. Beats on myös syy miksi ELK Stackin nimi on vaihdettu Elastic Stackiin. Beatsien tarve nousi Logstashin ongelmien takia, jotka alkoivat nousemaan esiin, kun dataputkista tuli monimutkaisempia ja prosessoitavan datan määrä nousi. (Berman 2020.)

Beatsit ovat kevyitä lokin lähetys työkaluja, ne käytännössä toimivat pieniä agentteina ympäristössä ja keräävät lokeja ja metriikkaa. Datan keräyksen jälkeen Beatsit lähettävät datan Logstashiin lisä prosessoitavaksi tai suoraan Elasticsearchiin. Beatseihin on olemassa valmiita moduuleita erilaisiin sisään- ja ulostuonti tarkoituksiin. (Berman 2020.)

Virallisista Beatsejä ovat:

Filebeat, jonka tarkoitus on lukea haluttuja tiedostoja järjestelmästä. Tämä Beat on kätevä, kun halutaan lukea muualta tuotuja tiedostoja, ja lokitiedostoja, joita ohjelmat tuottavat automaattisesti ja tiputtavat vaikka JSON-tiedostoon, tämänkaltaiset tiedot on tarkoitus kerätä Filebeatillä, jos Logstashiä ei ole haluttu käyttää. (What are Elasticsearch Beats? n.d.)

Metricbeatin tarkoitus on kerätä metriikkaa järjestelmistä ja ohjelmistoista. Metricbeatillä voidaan kerätä moduulien avulla tietoja Linux, Windows ja MacOS käyttöjärjestelmistä, esimerkiksi resurssien käyttöä ja tietoja ohjelmista kuten Apache, MongoDB ja MySQL. Metricbeatiin on suhteellisen helppoa tehdä ja muokata omia moduuleita. (What are Elasticsearch Beats? n.d.)

Packetbeat monitoroi verkko protokollia, jotta voidaan pitää yllä erilaisia tietoja verkosta ja verkkoliikenteestä. Esimerkiksi viiveitä, SLA tietoja ja käyttäjien verkkokäyttäytymistä. Packetbeat ymmärtää myös joitakin ohjelmisto tason protokollia kuten HTTP ja MySQL. (What are Elasticsearch Beats? n.d.)

Winlogbeatillä kerätään nimensä mukaan Windowsin tapahtumalokeja. Se voidaan asettaa suoraan lukemaan Windowsin tapahtumia, kuten kirjautumisia, USB-tallennusasemien käyttöä ja uusien ohjelmistojen asentamisia. (What are Elasticsearch Beats? n.d.)

Auditbeat on Winlogbeatin vastaava version Linux käyttöjärjestelmille. Sillä voidaan kerätä tietoja tapahtumista, kuten kirjautumisista. (What are Elasticsearch Beats? n.d.)

Heartbeatillä monitoroidaan järjestelmien käynnissä olemisaikaa. Se testaa haluttujen asioiden käynnissä olemista testaamalla yhteyttä niihin eri protokollia käyttäen. (What are Elasticsearch Beats? n.d.)

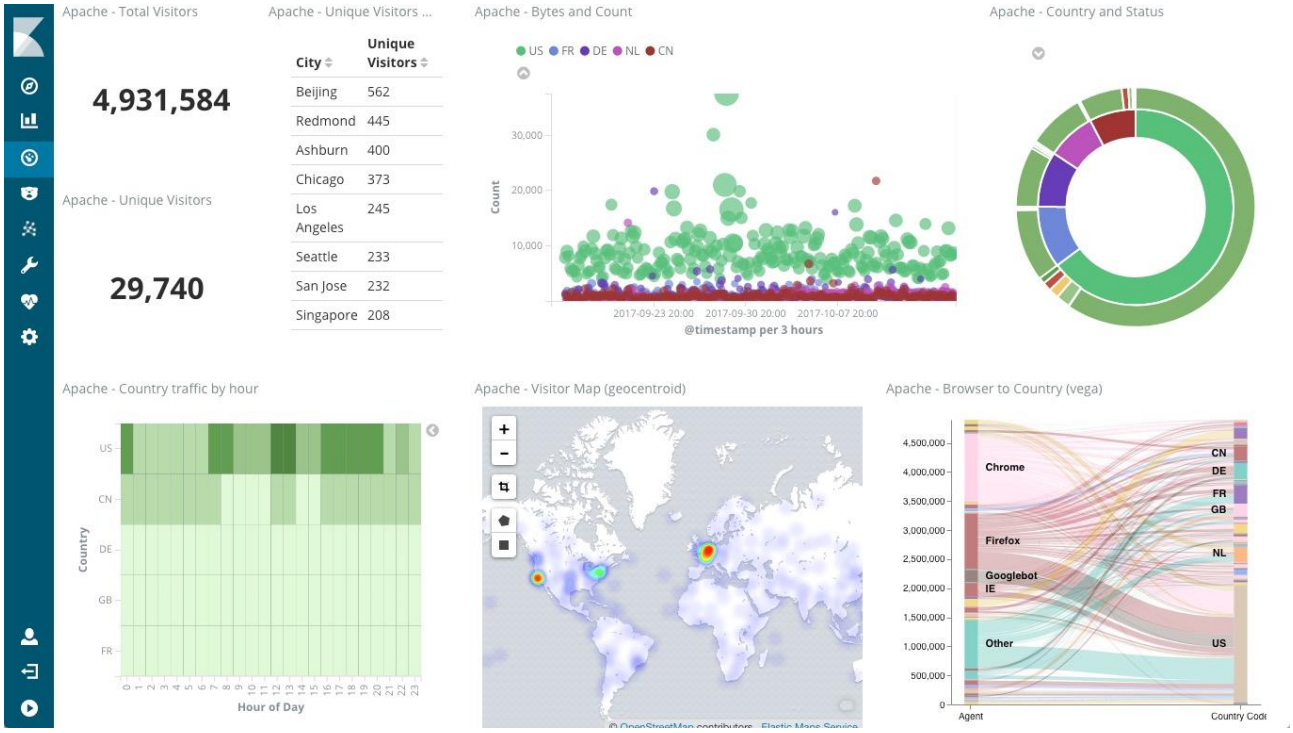
Functionbeat on tehty pilvipalveluiden monitorointia varten. Functionbeat on niin sanotusti palveliton ohjelma, se laitetaan palveluun esimerkiksi AWS Lambda sisälle toimimaan ja täten Functionbeat putkittaa data liikennettä suoraan Elastic Stackiin. (Functionbeat overview n.d.)

On olemassa myös erilaisia kommuunin tekemiä Beatsejä. Näitä on esimerkiksi Amazonbeat joka lukee tietoja halutusta Amazon palvelusta ja Apachebeat joka lukee HTTPD palvelimen tilaa. (Community Beats n.d.)

2.1.4 Kibana

Kibana on avoimen lähdekoodin selainpohjainen käyttöliittymä, jota käytetään datan visualisointiin ja tutkimiseen loki- ja aikajana-analyysissa, ohjelmisto monitoroinnissa ja tietoturva asioissa. Kibana tarjoaa monenlaisia visualisointityökaluja erilaisiin tarkoituksiin, myös hieman harvempiin tarkoituksiin kuten lämpökarttoja ja paikkatieteellisiä karttoja. Tietenkin Kibanasta löytyy normaaleita visualisointityökaluja kuten viivadiagrammeja ja ympyrädiagrammeja. (Kibana n.d.)

Esimerkki kuva Kibanan kojelaudasta, jossa on käytetty erilaisia visualisointeja kuviossa 2.



Kuvio 2. Kibana kojelauta (Kibana – More examples n.d.)

Kibana toimii Elastic Stackin käyttöliittymänä. Kibana mahdollistaa visuaalisen data-analysoinnin Elasticsearch indekseistä ja se antaa käyttäjän hakea Elasticsearch indeksejä ja visualisoida aikaisemmin mainituilla tavoilla. Kibana toimii myös käyttöliittymänä Elastic Stackin klusterin monitorointiin ja hallintaan. (What is Kibana? n.d.)

2.2 Splunk Enterprise

Splunkilla on useita tuotteita ja palveluita, jotka liittyvät datan tallentamiseen ja analytiikkaan sekä tietoturvallisuuteen (Splunk Products n.d.).

Splunk Enterprise on maksullinen Splunkin tuottama kyberturvallisuus alusta. Tällä alustalla voidaan kerätä, tallentaa, tutkia ja analysoida tekstipohjaista dataa, joka tulee minkälaisessa muodossa tahansa tai mistä lähteestä tahansa. Analysointi voidaan tehdä organisaation omasta IT-ympäristöstä tuotetusta datasta tai muualta maailmalta saadusta datasta (Why Splunk n.d.)

Splunk Enterprisellä dataa voidaan kerätä mm. verkkosivuilta, ohjelmistoista ja laitteista. Splunk Enterpriseä käytetään yleisesti selainpohjaisella käyttöliittymällä. (Splunk Enterprise overview 2021.)

Splunk Enterprisestä on olemassa myös 60 päivän ilmainen testiversio, joka muuttuu 60 päivän jälkeen rajoitettuun ilmaiseen versioon (Get Started With Your Free Trial n.d.).

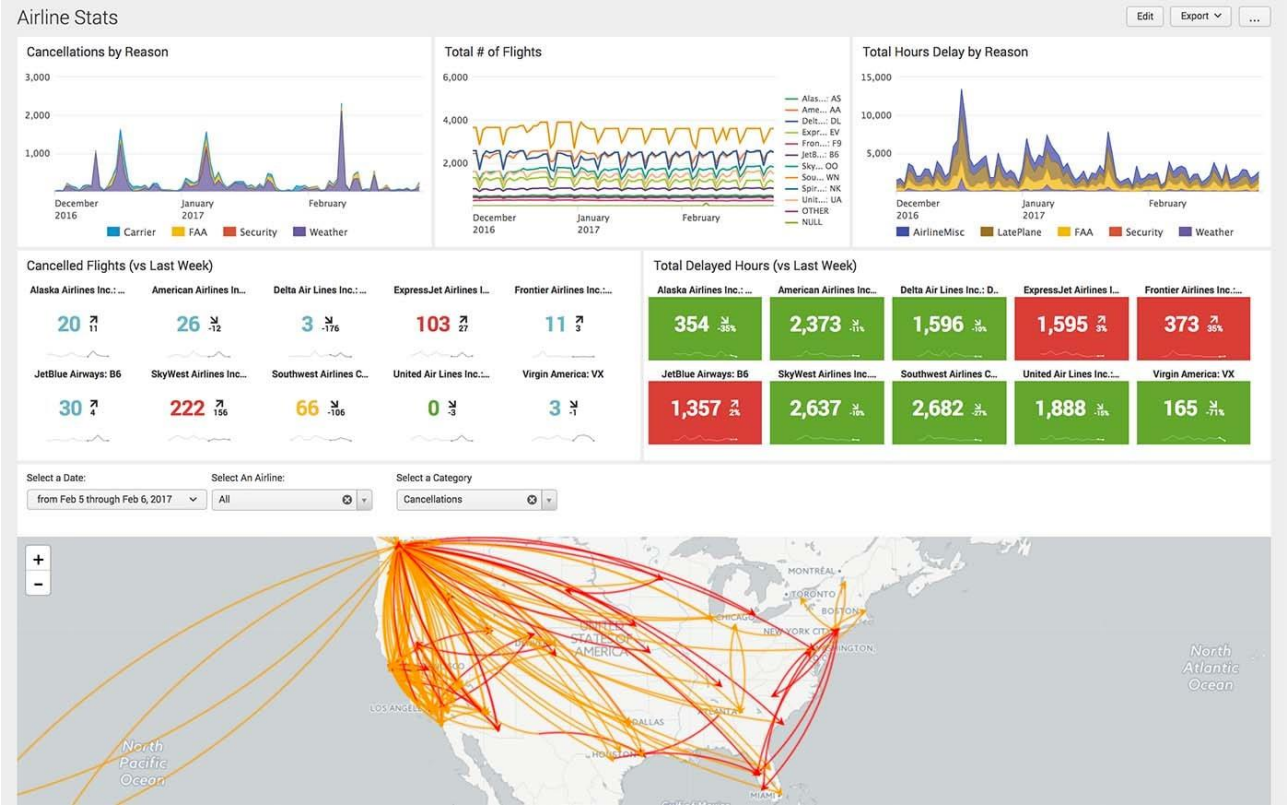
Splunk Enterprisen pääominaisuuksia ovat:

Search (haku), joka on käyttäjien päätapa navigoida dataa Splunk Enterprisessä. Hakuja voidaan tallentaa raporteiksi ja käyttää Dashboardissa erilaisissa paneeleissa. (Splunk Enterprise overview 2021.)

Alerts (hälytykset) on ominaisuus, joka antaa ilmoituksia, kun käyttäjän asetetut ehdot täyttyvät. Ilmoitukset voidaan laittaa lähettämään esimerkiksi sähköpostiin, RSS syötteeseen tai käyttäjän tekemään omaan päämäärään itse tehdyllä skriptillä. (Splunk Enterprise overview 2021.)

Dashboards (kojelauta) ominaisuus antaa käyttäjän asettaa kojelaudalle näkyviin erilaisia paneeleita esimerkiksi yllä mainittuja tallennettuja hakuja sekä ilmoituksia, näitten lisäksi normaaleita hakuikkunoita, diagrammeja. Diagrammeja ja muita visualisointeja voidaan tehdä reaaliaikaisesta datasta sekä historiallisesta. (Search Tutorial 2019.)

Esimerkki kojelaudasta erilaisilla visualisointeja käyttäen kuviossa 3.



Kuvio 3. Splunk Enterprise kojelauta (Luedtke 2017)

Pivot työkalun avulla voidaan raportoida tietystä data joukosta ilman, että käytetään Splunk Search Processing Languageta (SPL). Näitä raportteja voidaan sitten visualisoida eri tavoilla kojelaudalla. (Pivot Manual 2019.)

Reports (raportit) ovat tallennettuja hakuja sekä Pivot työkalulla tehtyjä asioita. Raportteja voidaan jakaa muitten käyttäjien kanssa sekä lisätä kojelaudalle. (Reporting Manual 2020.)

3 SIEM harjoitukset

3.1 Kyberharjoitus tyypit

Kyberharjoituksia on monenlaisia monenlaisiin eri tarpeisiin. Kyberharjoituksen tyyppi riippuu harjoituksen tekijöiden mukaan ja minkälaista laitteistoa on mahdollista saada mukaan kyberharjoitukseen. Harjoitustyyppejä voi sekoittaa keskenään tai ottaa osia toisesta harjoitustyypistä toiseen. (Instructions for organising cyber exercises 2020.)

3.1.1 Table Top harjoitus

Table Top (suom. lautapeli) harjoitus on kyberturvallisuus harjoitus, joka perustuu kokonaan kirjalliseen materiaaliin niin kuin normaalit lautapelit. Sen tarkoitus yleisesti on saada harjoituksen tekijät vastaamaan kysymyksiin erilaisista kyberturvallisuustilanteista kirjallisessa muodossa, hieman kuin kokeisessa koulussa. Eli tämän tyyppisessä harjoituksessa ei tarvita kyberturvallisuustyökaluja, kuten SIEM ratkaisuja. (Instructions for organising cyber exercises 2020.)

3.1.2 Toiminnallinen harjoitus

Toiminnallinen harjoitus (eng. functional exercise) on realistisempi ja jatketumpi versio Table Top harjoituksesta. Harjoituksessa harjoituksen tekijät saavat harjoitukset vetäjiltä lisäyksiä ja uusia tapahtumia harjoituksen aikana, joka tekee samalla harjoituksesta ajastetun, tällä tavalla yritetään simuloida realistisempaa tilannetta. Lisäys voi olla, vaikka haastattelu pyyntö uutistoimistolta koskien kyseisestä harjoitustilanteen tapahtumaa. (Instructions for organising cyber exercises 2020.)

3.1.3 Root cause harjoitus

Root cause (suom. pohjimmainen syy) harjoitus on kyberturvaharjoitus, jossa harjoituksen tekijöille esitetään oikea haitallinen kyberturvallisuustapahtuma, josta harjoituksen tekijöitten pitää selvittää ja jäljittää miten tähän lopputulokseen päästään. Tämmöisissä harjoituksissa voidaan käyttää jo tapahtuneita kyberturvallisuuden kannalta haitallisia tapahtumia organisaation sisällä. (Instructions for organising cyber exercises 2020.)

3.1.4 Tekninen harjoitus

Tekninen harjoitus (eng. technical exercise) on harjoitus, jossa IT-ympäristössä toteutetaan oikea tai oikeanalinen haitallinen kyberturvallisuustapahtuma. IT-ympäristö voi olla kuvitteellinen eli simulaatioympäristö tai vaikka kopioympäristö oikeasta ympäristöstä. (Instructions for organising cyber exercises 2020.)

Harjoituksessa on tarkoitus tehdä ongelmatilanne IT-ympäristöön ja testata harjoituksen tekijöitten ja organisaation valmiutta ja protokollaa ongelmatapauksissa. Tämmöisiä ongelmatilanteita

voi olla esimerkiksi havainto tuntemattomasta laitteesta verkossa, josta tulee erikoista liikennettä verkkoon. (Instructions for organising cyber exercises 2020.)

3.1.5 Capture the flag harjoitus

Capture The Flag (CTF) (suom. lipunryöstö) on harjoitus, joka on yleisesti enemmän tarkoitettu yksilöille tai ryhmille, organisaation sijaan. Näissä harjoituksissa tekijät etsivät ”lippuja” IT-ympäristöstä, nämä liput voivat olla merkkipätkiä jossain päin IT-järjestelmää mistä tekijän täytyy löytää ne. Useasti löydetty ”lippu” palautetaan, siitä saadaan pisteitä. Capture The Flag harjoituksista tehdään useasti erilaisia kilpailuita ja tapahtumia. (Instructions for organising cyber exercises 2020.)

3.2 Harjoitus tuotteet

3.2.1 Boss Of The SOC yleisesti

Boss Of The SOC (BOTS) on Splunkin tekemä CTF-tyylinen sinisen tiimin (eng. Blue Team) harjoitus. Harjoituksessa harjoituksen tekijät vastaavat kysymyksiin turvallisuus välikohtauksista hyödyntäen Splunk Enterpriseä sekä Splunkin tekemää realistista datasettiä fiktiivisestä yritys ympäristöstä. (Kovar 2020.)

BOTSien tarkoitus on emuloida miten oikeat turvallisuustapaukset näyttävät Splunkissa. BOTSin olemassaolon yksi syistä on se, että suurin osa CTF harjoituksista on punainen tiimi (eng. Red Team) pohjaisia. Sekä toinen syy on se että vaikka on muutamia sinisen tiimin harjoituksia, ne eivät Splunkin mielestä yritä luoda aidonlaista tilannetta kyberturva-analysoijalle. (Kovar 2020.)

BOTSistä on olemassa neljä eri versiota, joista versiot 1, 2 ja 3 ovat julkisesti saattavilla sekä versio 4 on opinnäytetyön tekohetkellä käytössä vain Splunkin harjoitus tapahtumissa ja turvallisuus konferensseissa. Splunk antaa julkiseen jakoon version, kun he ottavat itse seuraavan käyttöön. Splunk on myös kehittämässä versiota 5 opinnäytetyön tekohetkellä.

Scoreboard

BOTSeihin on olemassa pistetaulukko sovellus SA-ctf_scoreboard. Ohjelmalla voidaan pitää omaa jeoparty tyylistä tapahtumaa/kilpailua tai käyttää huvin ja urheilun vuoksi. Pistetaulukko sovellus toimii Splunk Enterprisen sisällä. Sovelluksessa on seuraavat ominaisuudet käyttäjä ja tiimi hallinta, pisteytys hallinta, kysymys, vastaus ja vihje hallinta sekä laaja pistetaulukko, kojelauta ja pistetaulukkoon liittyvä analysointityökalu. Sovellus ei ole pakollinen millään tavalla BOTS harjoitusten tekoon. Sovellusta voidaan käyttää opinnäytetyön teko hetkellä BOTS versioitten 1, 2 ja 3 kanssa. (SA-ctf_scoreboard 2020.)

3.2.2 Boss Of The SOC versio 1

Boss Of The SOC version 1 (BOTSv1) on Splunkin ensimmäinen BOTS harjoitus. Se sisältää tietoja kahdesta hyökkäyksestä organisaatioon. Se perustuu kuvitteelliseen po1s0n1vy aktivisti ryhmän hyökkäykseen kuvitteellista Batmanista tuttua Wayne Corp. yhtiötä vastaan. (Kovar 2018.)

BOTSv1:n kysymykset ja vastaukset saa halutessaan Splunkilta täyttämällä lomakkeen Splunkin verkkosivulla (BOTS 1.0 datasets n.d.).

3.2.3 Boss Of The SOC versio 2

Boss Of The SOC version 2 (BOTSv2) on Splunkin toinen BOTS harjoitus. Se on laajempi kuin BOTSv1 sillä se sisältää viisi turvallisuustilannetta, jotka ovat pitkäaikainen uhka, päätelaite turvallisuus, verkko ohjelmien turvallisuus, huijaus/petos yritys ja yrityksen sisäisen uhan tilanteen. (Herrald 2019.)

BOTSv2:n kysymykset ja vastaukset saa halutessaan Splunkilta laittamalla heille sähköpostia ja kysymällä niitä (BOTS 2.0 datasets n.d.).

3.2.4 Boss Of The SOC versio 3

Boss Of The SOC version 3 (BOTSv3) on Splunkin kolmas BOTS harjoitus. BOTSv3 sisältää kaksi kysymyssarjaa, 200-sarja ja 300-sarja. 200-sarja on enemmänkin yleisiä kysymyksiä ei toivotuista tapahtumista IT-infrastruktuurissa hyödyntäen muutamia työkaluja sekä tapahtumia ja ongelmia,

joita voi tulla vastaan yrityksillä siirtyessä käyttämään AWS palveluita. 300-sarja on monimutkaisempi ja joillekin hankalampi tehdä. 300-sarja on enemmän jatkumoja muutamasta turvallisuustapauksesta. (Herrald 2020.)

BOTSv3 200-sarjan kysymykset löytyvät liitteestä 1.

BOTSv3:n tapahtumapaikkana toimii kuvitteellinen panimo nimeltä Frothly. Harjoituksessa Frothly siirsi osan IT-infrastruktuuristaan pilveen Amazonin AWS palveluun, joka ei mennyt kaikella tavalla kunnolla. (CloudTrail - Digital Breadcrumbs for AWS 2018.)

BOTSv3:n kysymykset ja vastaukset saa halutessaan Splunkilta laittamalla heille sähköpostia ja kysymällä niitä (Herrald 2020).

Events

BOTSv3 sisällä on tapahtumia (eng. events). Erilaisten tapahtumien määrä BOTSv3 datasetissä on 2 842 010. Tapahtuma on nimensä mukaan yksi tapahtunut asia järjestelmässä, nämä tapahtumat syntyvät, kun jonkin asetetun datan keräys työkalun havaitsee haluttua tapahtumaa. Näitä voi olla esimerkiksi verkkoliikennettä seuraava asia tai ohjelma, joka kerää pienen aikavälin välein tietokoneen suorituskyky tietoja.

Esimerkki BOTSv3 tapahtumasta, joka on otettu Splunk Enterprisestä raakatekstinä.

{"endtime":"2018-08-20T15:27:09.296788Z","timestamp":"2018-08-20T15:27:09.296118Z","bytes":50637,"src_ip":"13.125.33.130","src_mac":"02:4F:D7:61:53:04","bytes_in":30,"packets_in":1,"protocol":"UDP","dest_ip":"172.16.0.178","dest_mac":"02:A0:38:1B:B3:70","bytes_out":50607,"packets_out":36,"flow_id":"c240fdab-997d-4285-a221-1080976a6b8a","protoid":17,"version":4,"tos":0,"fragment_count":0,"protocol_stack":"ip:udp:unknown"}

Fields

Fields eli kentät ovat kohtia tapahtumassa, jossa on kyseisen tapahtuman tiedot ovat. Kenttiä voi olla esimerkiksi kohde IP osoite, internet protokolla tai vaikka CPU käyttömäärän prosentti.

Kaksi tärkeää kenttää ovat host ja sourcetype. Host ja sourcetype ovat mukana kaikissa tapahtumissa.

Host

Host eli päätelaite kertoo tapahtumaan liittyvän päätelaitteen, tai jos ei ole päätelaitetta, tämä ilmoitetaan host kentässä arvolla ”serverless”.

Sourcetype

Datasetissä on useita eri data lähde tyyppejä, joita kutsutaan nimellä sourcetype, sourcetype on tallennettu datassa sourcetype kenttään. Sourcetype kenttä ilmoitta mistä tai minkälaista tietoa tapahtuman tieto on. Esimerkiksi sourcetype ”aws:cloudtrail” kertoo sen että kyseisen tapahtuman tiedot ovat AWS CloudTrail palvelusta kotoisin. Toinen esimerkki on sourcetype ”stream:smtp”, tämä sourcetypen tapahtumat sisältävät sähköposti liikennettä.

4 BOTSv3 200-sarjan ohjeistuksen ymmärrykseen vaadittavat teknologiat

Tässä kohdassa on käyty lyhyesti läpi teknologioita ja käsitteitä, joita tarvitaan BOTSv3 200-sarjan ohjeistuksen ymmärtämiseen.

4.1 Amazon Web Service

Amazon Web Services (AWS) on Amazonin alusta erilaisille pilvipalveluille (What is AWS n.d.).

4.1.1 AWS API Gateway

AWS API Gateway on AWS palvelu, joka on luotu luomista, julkaisua, ylläpitoa, monitorointia sekä REST, HTTP ja WebSocket API:en turvaamista varten. API-kehittäjät voivat tällä luota APIja joilla päästään käsiksi muihin AWS palveluihin. (What is Amazon API gateway? n.d.)

4.1.2 AWS Auto Scaling

AWS Auto Scaling monitoroi käyttäjien ohjelmistoja sekä järjestelmiä ja automaattisesti säätää resursseja, kuten laskentatehoa, jotta järjestelmiä voidaan pitää päällä ilman katkoksia pienimmällä mahdollisimmalla kuluilla käyttäjille. Auto Scalingia voidaan käyttää mm. Amazon EC2 instanssseissa. (Introducing AWS Auto Scaling 2018.)

4.1.3 AWS Bucket ja Amazon S3

Bucket on paikka, jossa kaikki pilvipalveluun tallennettu data säilytetään. Käytännössä Amazon S3 on palvelu, jolla hallitaan Bucketteja. Ohjeistuksen kannalta on hyvä tietää, että Amazon Buckettin oikeudet hoidetaan ACL avulla, jolla voidaan asettaa Bucket tai Bucketissa olevia objekteja avoimeksi kaikille internet käyttäjille. (Bucket overview n.d.)

Amazon S3 (Simple Storage Service) on Amazonin pilvitallennus palvelu. Sen suurimmat ominaisuudet ovat helppo käytettävyys ja konfigurointi sekä skaalautuvuus. (What is Amazon S3 n.d.)

4.1.4 AWS CloudTrail

AWS CloudTrail on AWS palvelu, joka mahdollistaa AWS käyttäjätilin hallinnan, vaatimuksenmukaisuus sekä operaatio- ja riski auditoinnin sallimisen (What is AWS CloudTrail? n.d.).

Käyttäjän, roolin tai AWS palvelun toiminnot tallennetaan tapahtumina CloudTrailiin. AWS Cloudtrailiin sisältyvät tapahtumat sisältävät AWS-hallintakonsolin, AWS-komentoriviliittymän sekä AWS SDK:n ja AWS API:n toteuttamat asiat. (What is AWS CloudTrail? n.d.)

4.1.5 Amazon Machine Image JA AWS CloudImage

Amazon Machine Image (AMI) on AWS sisällä käytettävä instanssi/käyttöjärjestelmä image. Kun AWS palvelussa käynnistetään uusi instanssi saa se tiedot valitulta imagelta, kuten käyttöjärjestelmän ja konfiguraatiot. (Amazon Machine Images (AMI) n.d.)

Pilvipalveluissa käytettävistä imageista käytetään joskus myös nimeä ”Cloud Image”.

4.1.6 AWS EC2

Amazon EC2 (Elastic Compute Cloud) on Amazonin verkkopalvelu, joka tarjoaa turvallista ja koollisesti muutettavaa laskentakapasiteettiä pilvessä, toisin sanottuna tietokoneita. EC2 tarkoitus on tarjota helpompaa skaalautuvaa laskentatehoa hyödyntäen helppoa käyttäjäliittymää, josta voi konfiguroida ja hankkia kapasiteettiä. (Amazon EC2 n.d.)

4.1.7 AWS IAM

AWS IAM (Identity and Access Management) on Amazonin palvelu, joka antaa käyttäjän kontrolloida turvallisesti pääsyä AWS resursseihin. IAM:iä käytetään sen päättämiseen, kuka on autentikoitu ja kellä on lupa käyttää haluttuja resursseja. IAM käyttäjät ovat käyttäjiä, joilla hallinnoidaan AWS palveluita. (What is IAM? n.d.)

AWS MFA

AWS MFA (Multi-factor Authentication) lisää uuden turvallisuustason käyttäjä nimen ja salasanan päällä. Kun MFA on käytössä, käyttäjä joutuu kirjoitettuaan käyttäjänimen ja salasanan varmistamaan kirjautumisen vielä toisella tavalla. Tämä tapa on yleisimmin koodi, joka saadaan jostakin laitteesta, kun kirjautumisen ensimmäinen vaihe (käyttäjänime ja salasana) on tapahtunut, laitteena toimii yleensä sovellus älypuhelimessa tai erikseen ostettava laite. (Multi-factor Authentication n.d.)

AWS IAM avaimet

AWS palvelussa käytetään kahdenlaisia kirjautumisavaimia, access key ID ja secret access key. Ne ovat normaali käytössä kuin käyttäjätunnus ja salasana. (Managing access keys for IAM users n.d.)

AWS Account ID

AWS Account ID on tunnus, jolla IAM käyttäjä kirjautuu AWS käyttäjälle, ID on numero sarja kuten 123412341234. Account ID sijasta voidaan käyttää myös account aliasta joka osoittaa account ID:seen. (Your AWS account ID and its alias n.d.)

4.2 Kryptovaluuttojen louhintaan liittyvää terminologiaa

4.2.1 Mining

Krypto mining (suom. louhinta) tarkoittaa kryptovaluuttojen louhimista. Tämä tapahtuu käyttäjän tietokonetta tai vastaavaa laitetta käyttämällä. Ideana on se, että käyttäjän laite ratkaisee kryptologisia yhtälöitä, tästä saa sitten palkkioksi kryptovaluuttaa. (Crane 2020.)

4.2.2 Kryptovaluutta Monero

Kryptovaluutta on digitaalinen virtuaalivaluutta, joka on suojattu kryptologiaisilla teknolokioilla. Tämä tekee väärentämisestä erittäin hankalaa. Moni kryptovaluutta on perustettu jaetun Blockchain teknologian päälle. (Frankenfield 2021.)

Monero on yksi suosituista kryptovaluutoista, Moneron ero kahteen muuhun yksistä suosituimmista kryptovaluutoista Bitcoiniin ja Ethereumiin on se, että Moneron siirto tiedot eivät ole julkisija. Lähettäjä, vastaanottaja ja määrä on vakiona piilotettu. (What is Monero (XMR) n.d.)

4.2.3 Coin

Coin ja Crypto Coin ovat yleisiä termejä, joita käytetään kryptovaluutoista. Sana ”coin” esiintyy myös usean kryptovaluutan nimessä esimerkiksi Bitcoin, Litecoin ja Binancecoin sekä louhintaan liittyvissä sivustoissa.

4.2.4 Coinhive

Coinhive oli Kryptovaluutan louhinta sivusto. Coinhiveä käytettiin ennen sen sulkeutumista paljon malware tyyppisesti. Coinhiven koodi ujutettiin verkkosivun koodiin ja kun käyttäjä kävi kyseisellä verkkosivulla, Coinhiven louhinta ohjelma alkoi louhia Monero kryptovaluuttaa ilman käyttäjän lupaa. Kun käyttäjän tietokoneeseen on haitallinen ohjelma päässyt alkaa ohjelma louhia käyttäen prosessoria ja tai näytön ohjaimella. (Krebs 2018.)

Coinhiven selain pohjaisen louhinta ohjelman alkuperäinen tarkoitus ei ollut hakkereitten käyttöön, vaan sivustojen omistajille, jotka halusivat tienata hieman lisä rahaa sivuston käyttäjiltä. Tämä kumminkin muuttui nopeasti pahisten käyttämäksi tavaksi tienata lisä rahaa pienen selain

pohjaisen koodin takia. Coinhiven louhinta koodi nousi tästä syystä usean kyberturvallisuus yhtiön haittaohjelmien ja malwareitten kärkisijoille. (Krebs 2018.)

Coinhive lopetti toimintansa 8.3.2019 (Salát 2019).

4.2.5 JSCoinminer

JSCoinminer on haittaohjelma, joka louhii kryptovaluuttoja ilman käyttäjän lupaa (Pilici 2017).

4.3 Protokollat

4.3.1 DNS

Domain Name System (DNS) osa normaalia internetinfrastruktuuria. Kun internetin käyttäjä ottaa yhteyttä verkkosivuun esimerkiksi jamk.fi ottaa käyttäjä yhteyttä Domain nimen perusteella. DNS muuttaa tämän jamk.fi nimen IP osoitteeksi, jotta tietokone ymmärtää mihin sivuun yritetään ottaa yhteyttä. (What is DNS? N.d.)

4.3.2 FQDN ja Host

Fully Qualified Doamin Name (FQDN) on perusteellisin domain nimi, se sisältää päätelaitteen ja domainin ja ylimmän tason domainin. FQDN voi myös sisältää ali domaineja. Esimerkkinä www.jamk.fi, www on päätelaite, jamk on domain ja fi on ylimmän tason domain, tämä kokonaisuus tekee osoitteesta FQDN. Pelkkä jamk.fi ei on FQDN. (Christensson 2019.)

Host, suomeksi päätelaite joskus isäntälaite. Se on lyhyesti sanottuna tietokone tai muu laite esim. verkkotulostin, johon voidaan ottaa yhteys verkon ylitse. Palvelimet ovat myös päätelaitteita. (Christensson 2015.)

FQDN:ssä päätelaite laite on esimerkiksi www tai mail kohta (Christensson 2019).

4.3.3 PUT metodi

PUT metodi on HTTP protokollan metodi. PUT pyytää, että kohde resurssi luodaan tai korvataan annetulla arvolla tai asialla. Esimerkiksi JPEG-tiedoston lisääminen verkkosivulle voidaan tehdä PUT metodilla. (RFC7231 2014.)

4.3.4 Query

Query on informaation pyyntö tietokannasta (Query n.d.).

4.3.5 SMTP

Simple Mail Transfer Protocol (SMTP) on sähköpostiviestien siirto protokolla. Protokollan tarkoitus on lähettää/siirtää sähköposteja luotettavasta ja tehokkaasti. SMTP on riippumaton datan liikuttamis- alasysteemistä (esim. tcp) ja SMTP:llä voi liikuttaa dataa eri verkko. (RFC5321 2008.)

4.4 Teknologiat

4.4.1 ACL

Access Control List (ACL) on lista, jonne asetetaan laitteiden ja käyttäjien sisään tulevalle ja ulosmenevälle liikenteelle erilaisia oikeuksia erilaisiin resursseihin IT-infrastruktuurissa (Cox 2020).

Network ACL on AWS tapauksessa kuin palomuuri, joka kontrolloi liikennettä AWS aliverkkojen sisällä. Tätä kontrollointia pystyy muokkaamaan oman maun mukaan sekä sisään tulevaan ja ulos menevään liikenteeseen. (Network ACLs n.d.)

4.4.2 Cisco NVM

Cisco Network Visibility Module (NVM) kerää verkon virtaus (eng. flow) konteksti dataa päätelaitteilta. Tämä kun yhdistetään vaikka Splunkiin saadaan lisää näkyvyyttä ja tietoa verkko liikenteestä ja verkkolaitteista. (Network Visibility Module n.d.)

4.4.3 Cloud-init

Cloud-init on standardi ohjelma, jota käytettään pilvessä olevien instanssien alustamiseen. Kun instanssi käynnistetään pilvessä Cloud-init asentaa haluttuja ohjelmia ja tekee haluttuja konfiguraatioita kuten verkko asetuksia ja SSH avaimien asetuksen. (Cloud-init documentation n.d.)

4.4.4 Performance Monitor (perfmon)

Performance Monitor (suom. suorituskyky monitori) on Windowsin työkalu, joka mittaa tietokoneen erilaisia suorituskyky asioita kuten CPU käyttömäärää. Performance monitorilla voidaan myös katsoa historiallista tietoa ja ohjelma kohtaisia suorituskykyyn liittyviä asioita. (Marcho 2019.)

Performance Monitor voidaan lyhentää PerfMon ja näkyy myös tällä nimellä BOTSv3 datassa.

4.4.5 Symantec Endpoint Protection

Symentec Endpoint Protection (SEP) oli palvelin pohjainen kyberturva-alusta fyysisille- ja virtuaalilaitteille. SEPin ominaisuuksia olivat mm. antivirus ja antimalware ohjelmat, palomuuri, ohjelmisto- ja verkkoyhteyskontrollointia sekä verkkosivujen selailun turvaaminen. (Tittel 2017.)

SEPM on lyhennesanoista Symentec Endpoint Protection Manager. Tämä on palvelin, joka jutteli SEP asiakas ohjelman kanssa. (Symantec Endpoint Protection Manager n.d.)

Symentec Endpoint Protectionin Small Business Edition on lopettanut toimintansa 2.11.2020 (End of Life notice for Endpoint Protection Small Business Edition Cloud and Endpoint Protection Cloud 2021).

5 Toteutus

5.1 BOTS version valinta

BOTS versio valittiin BOTSV1, BOTSv2 ja BOTSv3 välillä julkaisu ajankohdan perusteella ja tiedosto kokojen takia, koska alustavissa testeissä Splunkilla oli ongelmia ulostuoda isoja data määriä.

BOTSv3:n tiedosto koko on ladatessa 320.1 MB ja BOTSv2:sta saa kaksi versiota “BOTS V2 Dataset” ja “BOTS V2 Dataset Attack Only” joista ensimmäinen on 16.4gb ladatessa ja jälkimmäinen 3.2 GB ladatessa. BOTSv1:sta saa myös ison koko version ”BOTSv1 Dataset”, joka on 6.1 GB ja pienen version ”BOTSv1 attack only” joka on 135MB.

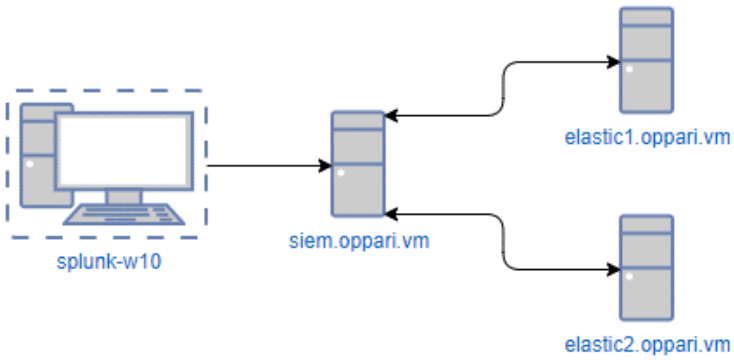
BOTSv1 pienempi versio, BOTSv2 pienempi versio ja BOTSv3 olivat käytännössä mahdollisia ulostuoda Splunk Enterprisestä pilkkomalla useampaan pienempään ulostuontiin, vaikka BOTSv2 koko olisi tuottanut isoja ongelmia. Tämän lisäksi BOTSv2 datan ulostuonti vain loppui kesken oudosti tai ulostuontitiedosto oli kokonaan tyhjä.

Tämän takia versioksi valittiin BOTSv3 pääasiassa sen takia että se on, uusin julkisesti saatavilla oleva BOTS versio ja pienesti sen takia että se on helpompi ulostuoda Splunk Enterprisestä kuin BOTSv2 oudon ongelmansa kanssa.

5.2 Topologia

Tietokoneet ja palvelimet olivat asennettuna toimeksiantajan VMware Cloud Director pilvipalveluun, tämä johtui siitä, että toimeksiantaja halusi järjestelmät, jotka syntyivät opinnäytetyön sivutuotteena heille. Tästä toimeksiantaja voi kopioida tai tehdä muuta haluamaansa suoraan eikä tarvitse alkaa tekemään kaikkia konfiguraatioita ja asennuksia erikseen. Kaikki tietokoneet ja palvelimet olivat virtuaalisia.

Kuviossa 4 olevissa laitteissa splunk-w10 on asennutettu Splunk Enterprice, siem.oppari.vm koneelle Kibana, Logstash ja Filebeats ja elastic1.oppari.vm ja elastic2.oppari.cm koneille kummallekin hajautettuna Elasticsearch.



Kuvio 4. Topologia kuva

5.3 Splunk osuus

Splunk osa toteutuksesta. Splunkilla tehtiin BOTSv3 datan muokkaaminen luettavaan muotoon Elastic Stackille.

5.3.1 Splunk asennus

Splunk Enterprise 8.1.0.1 asennettiin ilmainen 60 päivän kokeilu versio. Splunk Enterprisestä sai Windows 10, Windows Server 2016 ja 2019, Linux ja Mac OS OSX vesion.

Käyttöön valittiin Windows 10 käyttöjärjestelmä ja sen vastaava Splunk Enterprise versio. Valinta perustuin siihen että, toimeksiantajan pilvessä olevien valmiitten käyttöjärjestelmien joukossa Windows 10 oli ainoa tuttu käyttöjärjestelmä, joka sisälsi graafisen käyttöliittymän. Tämän lisäksi Splunk Enterprise asennusta käytettiin ainoastaan BOTSv3 sisääntuontiin sekä BOTSv3 ELK SIEM ohjeistuksen tekemisen avustukseen, täten stabiilisuus sekä muista asioista ei tarvitse huolehtia. Windows 10 käyttöjärjestelmä versio oli 1909.

Splunk Enterprise asennukseen ei tarvinnut tehdä muita muutoksia myöhemmin tulevien lisäosien lisäksi, kuin Session Timeoutin muuttaminen isompaan. Tämä muutos oli enemmänkin elämän laadun parantamista kuin käyttöön tarpeellista. Tämän muutoksen sai tehtyä Splunk Enterprisen selain pohjaisen käyttöliittymän kautta **Settings >> Server Settings >> Gerneral Settings** Splunk Web kohdasta muokkaamalla Session Timeout kenttää, tämä asetus muutettiin **60d**, joka muutti istunnon vanhenemisajaksi 60 päiväksi.

5.3.2 Splunkin valmistamien BOTSv3:sta varten

Splunk Enterprise vaati BOTSv3 kaikkien kenttien lukemiseen ison määrän lisäosia. Nämä lisäosat olivat listattuna Splunkin BOTSv3 GitHub sivustolla lataus linkkeineen, joka löytyi seuraavasta lähteestä (Frazier & Herrald 2020).

Splunk Enterprisen lisäosat piti ladata erikseen Splunkin tietokannasta ja sisääntuoda Splunk Enterpriseen manuaalisesti. Splunk Enterprise lisäosat ja lisäosien versiot, joita käytettiin, näkyvät taulukossa 3.

Taulukko 3. Splunk Enterprisessä käytetyt lisäosat ja niitten versiot

Lisäosa	Versio
Aws_guardduty	1.0.4
CiscoNVM	3.1.9
Code42 App For Splunk	3.0.12
Code42ForSplunk Technology Add-On	3.0.12
Splunk Add-on for Cisco ASA	4.1.0
Splunk Add-on for Microsoft Cloud Services	4.1.1
Splunk Add-on for Microsoft Office 365	2.0.3
Splunk Add-on for Microsoft Windows	8.1.1
Splunk Add-on for Symantec Endpoint Protection	3.2.0
Splunk Add-on for Tenable	5.1.4
Splunk Add-on for Unix and Linux	8.3.0
Splunk Common Information Model	4.18.1
Splunk Security Essentials	3.3.0
Splunk Stream Add-on	7.3.0
TA-VirusTotalActions	0.2.0
URL Toolbox	1.8
DecryptCommands	2.3.1
Microsoft Azure Active Directory Reporting Add-on for Splunk	3.1.1
Microsoft Cloud App for Splunk	3.2.2
Microsof Office 365 Reporting Add-on for Splunk	1.2.4
Microsoft Sysmon Add-on	10.6.2
OSquery App for Splunk	0.6.0
Splunk Add-on for AWS	5.0.3
ES Content Updates	3.16.0
SA-cim_vladiator	1.8.0

BOTSv3 tiedostot saatiin näkyviin hakemalla **index=botsv3 earliest=0** Splunk Enterprisen Search osasta.

5.3.3 Splunk ulostuonti

BOTSv3 oli GitHubista ladatessa Pre-indexed Splunk -formaatissa. Tämä formaatti täytyi muokata muotoon, jota Elastic Stack järjestelmä pystyi lukemaan ja jäsentämään. Splunk Enterprisestä pystyi ulostuomaan tiedostoja ja tapahtumia CSV, JSON ja XML muodossa. JSON valittiin formaatiksi koska, testailuiden jälkeen JSON osoittautui käytännöllisimmäksi tiedosto formaatiksi käyttää Elastic Stackin kanssa. Sekä Elasticsearch tallentaa itse tiedostonsa JSON formaattiin.

Ulostuonti tapahtui Splunk Enterprisen Search ikkunassa ”Export” painiketta käyttäen.

Elastic Stackissä oli vakiona rajoitus indeksien uniikki kenttä määriin, tämä rajoitus on 1000 kenttää indeksiä kohden. Oli haluttua pitää järjestelmät mahdollisimman lähellä vakio tilaa, joten tämä ongelma kierrettiin jakamalla ulostuonnin kuuteen eri osaan. Jako tapahtui keinotekoisesti päätelaitteiden perusteella.

Taulukossa 4 näkyy että, osassa yksi oli 548 uniikki kenttää ja 547 132 tapahtumaa. Osassa kaksi a oli 11 507 ja kaksi b oli 6 042 tapahtumaa. Kolmannessa osassa oli 919 uniikki kenttää sekä tapahtumien määrä oli 910 854. Neljännessä osassa uniikki kenttiä oli 804 ja tapahtumia 875 296. Viidennessä osassa uniikki kenttiä oli 890 ja tapahtumia 491 179. Kokonaistapahtumien määrä oli 2 842 010.

Taulukko 4. BOTSv3 ulostuonnit

Osa	Haku	Uniikki kenttä määrä	Tapahtuma määrä
1	index=botsv3 earliest=0 host="gacrux*"	548	547 132
2a	index=botsv3 earliest=0 host="splunk.froth.ly" product=*	a ja b yht. 919	11 507
2b	index=botsv3 earliest=0 host="splunk.froth.ly" NOT product=*	a ja b yht. 919	6 042
3	index=botsv3 earliest=0 host="mars*" OR host="*-L"	919	910 854
4	index=botsv3 earliest=0 host="ip*" OR host="hoth" OR host="SEPM"	804	875 296

5	index=botsv3 earliest=0 host!="gacrux*" host!="hoth" host!="mars*" host!="*-L" host!="splunk.froth.ly" host!="ip*" host!="SEPM"	890	491 179

Nämä ulostuontitiedostot toimitettiin toimeksiantajalle, jottei heidän tarvitse tulevaisuudessa itse tehdä niitä uudestaan.

Tämän jälkeen kaikki tiedot olivat tallessa JSON-tiedostoissa valmiina lataukseen Elastic Stackiin.

Splunk koneella asetettiin jaettu kansio, josta voitiin Elastic Stackin käyttöliittymä koneella ottaa käsittelyyn tiedostoja, jottei tarvitsisi siirtää tiedostoja järjestelmästä toiseen.

5.4 Elastic Stack osuus

5.4.1 Elastic Stack asennus

Elastic Stackistä saatiin toimeksiantajalta valmis asennus opinnäytetyötä varten. Asennus oli tehty toimeksiantajan pilvipalveluun ja palvelimet olivat virtuaalisia. Asennus oli jaettuna kolmeen osaan, siem.oppari.vm kone, joka jossa oli asennettuna Kibana ja Logstash tämä toimi niin sanotusti ”frontendinä”. Sekä elastic1.oppari.vm ja elastic2.oppari.vm joissa oli asennettu hajautettuna Elasticsearch kumpaankin. Ohjelmat olivat valmiina yhdistettyinä ja konfiguroituna.

Elastic Stack asennus ja konfiguraatio haluttiin pitää mahdollisimman vakiona toimeksiantajan pohjasta.

Elastic Stackiin käytetyt koneet olivat Red Hat Enterprice Linux 7 versio 7.7.1908 käyttöjärjestelmän päällä.

Splunk koneen jaettu kansio saatiin asetettua käyttäen cifs-utils työkalua ja asettamalla Elastic Stackin käyttöliittymä koneelle se komennolla **mount -t cifs -o username=administrator //10.110.3.189/botsv3 /mnt**.

Siem.oppari.vm palvelimella asennettiin jälkikäteen lisäksi Filebeat. Filebeatillä yritettiin korjata ongelma sisääntuonnissa. Elastic Stackin sisääntuonti ongelmista on kerrottu myöhemmin.

5.4.2 Elastic Stack sisääntuonti

BOTSv3:n JSON-tiedostojen sisääntuonti tehtiin hyväksikäyttäen Logstashiä. Logstash toimii konfiguraatiotiedostoa ajamalla Logstash ohjelman kanssa. Konfiguraatiotiedosto on jaettu kolmeen osaa input (sisääntuonti), filter (suodatus) ja output (ulostuonti).

Input kohtaan määritettiin ensiksi Input Plugin, tähän tarkoitukseen valittiin ”File” syystä, että haluttiin prosessoida JSON-tiedostoa.

Codec kohta purkaa halutun tyypin koodauksen. Tässä tapauksessa, kun haluttiin lukea JSON-tiedostoja, käytettiin JSON asetusta

Start_position kohdalla päätetään, luotaanko tiedosto alku vai loppu päästä aloittamalla, tässä tarkoituksessa alkupää oli parempi vaihtoehto.

Path kohtaan tulee tiedoston absoluuttinen sijainti.

Sincedb_path kohdalla osoitetaan database tiedosto, joka pitää yllä missä kohdassa luettua kohtaa luettavasta tiedostosta. Tämä ominaisuus on hyödyllinen, jos tiedostoon tulee lisää tietoja, mutta tässä tapauksessa niin ei tapahtunut, joten se laitettiin osoittamaan tiedostoon, jota ei ole olemassa. Tällä tavalla saatiin lukeminen aloittamaan aina alusta.

Työssä käytetyn konfiguraatio tiedoston Input kohta

```
input {
  file {
    codec => "json"
    start_position=> "beginning"
    path => "/mnt/botsv3part1.json"
    sincedb_path => "/dev/null"
  }
}
```

Filter kohdassa voidaan muuttaa ja filteröidä dataa. Filter kohta myös hoitaa datan jäsentämisen. Tähän kohtaan asetettiin lisäosaksi JSON, jotta saatiin Logstash jäsentämään JSON dataa.

Source kohtaan asetetaan kenttä, josta jäsentäminen tapahtuu. Logstash tallensi JSON datan tapahtumat message kenttään, joten se otettiin käyttöön.

```
filter {
  json {
    source => "message"
  }
}
```

Output pluginina käytettiin tietenkin Elasticsearch pluginia, koska data on tarkoitus saada Elastic Stackin Elasticsearch palvelimille. Kyseisessä Elastic Stack asennuksessa käytettiin SSL salausta, tähän liittyvät asetukset tuli laittaa Output kohtaan.

Hosts kohtaan laitettaan Elasticsearch palvelimien osoitteet sekä portit tarvittaessa. SSL kohdassa asetetaan SSL päälle tai pois päältä. Cacert kohdassa laitetaan SSL sertifikaatio tiedoston absoluuttinen sijainti. Ssl_certificate_verification kohdassa voidaan ottaa pois päältä palvelimen puolen sertifikaatin validointi. User kohtaan tulee Elasticsearchin käyttäjänimi. Password kohtaan tulee Elasticsearch salasana. Index kohta asettaa läpi kulkeviin tapahtumiin indeksin.

Ilm_enabled kohdassa päätetään, onko Index Lifecycle Management ominaisuus päällä vai ei. Index Lifecycle Management on Elastic Stackin ominaisuus, jolla automaattisesti hallitaan indeksejä. Tätä ominaisuutta ei tarvittu tässä tapauksessa, joten se oli otettu pois päältä.

```
output {
  elasticsearch {
    hosts => ["10.110.3.156","10.110.3.162"]
    ssl => true
    cacert => "/etc/pki/tls/certs/elastic-ca.crt"
    ssl_certificate_verification => false
    user => "elastic"
    password => "root66"
    index => "botsv3part1"
    ilm_enabled => false
       manage_template => false
  }
}
```

Kun Logstashin läpi kulkee dataa Logstash jäsentää niitä Elasticsearhille sopivaan muotoon ja lisää samalla metadata kenttiä. Tämän takia Elastic Stack lisää automaattisesti jäsennetyn JSON datan kenttien alkuun sisennyksen ”result” tämä tekee esimerkiksi JSONin sisällä olevasta ”host” kentästä ”result.host” kentän. Tämä johtuu siitä, että Elastic Stackin lisäämät kentät voivat käyttää samaa kenttänimeä, ”host” kentän tapauksessa Elastic Stack lisää kentän ”host” itse, jossa näkyy laite, jossa sisääntuonti tapahtui, tässä tapauksessa ”siem.oppari.vm”.

Lyhyesti siis BOTSv3 ”host:BSTOLL-L” kenttä muuttui ”result.host:BSTOLL-L” ja uusi kenttä ”host:siem.oppari.vm” ilmestyi mukaan.

5.4.3 Ongelmat

Sisään tuonnissa Elastic Stackiin todettiin ongelma BOTSv3 datassa. Tapahtumista tippui noin 1/3–2/3 riippuen tapahtuminen määrästä ja koosta sekä järjestelmästä missä sisääntuonti tehtiin. Tämä vielä tarkennettiin siihen, kun tapahtumien määrä lähenee tuhatta tapahtumaa, alkaa tapahtumia tippumaan. BOTSv3 tapahtuma määrä on 2 842 010, joten sisääntuontien jakaminen kokonaan eri osiin ei ole käytännöllistä.

Ongelma jäljitettiin todennäköisemmin Logstashin välimuistiin rajoitukseen. Joka alkaa tiputtamaan tapahtumia, kun se täyttyi. Välimuistin täyttyminen tapahtuu ennen kuin Logstash alkaa käsittelemään dataa.

Alla on kerrottu lyhyesti ongelman korjaus yrityksen päätavoista ja tärkeimmistä tavoista, tulevaisuuden kannalle.

Logstash lisäosat

Ongelmaa testattiin ja yritettiin korjata sillä, että lisättiin konfiguraatioon Filter (suodatin) kohtaan ”Sleep” suodatus lisäosa, se antoi komennoilla ”time” lisätä ajan, jonka se odottaa sekunteina ennen kuin jatkaa toimintaa ja ”every”, jolla valittiin, kuinka monen tapahtuman välein odottaminen tapahtui. Tämä ei nostanut läpi menevien tapahtumien määrää.

Toinen testi oli käyttää Logstashin Fingerprint lisäosaa. Tällä lisäosalla voitiin lisätä uniikkileima tapahtumaan, jotta Logstash päivittää eikä lisää uutta tapahtumaa Elasticsearchiin. Tällä yritettiin testata, voidaanko sisääntuonnin puuttuvat tapahtumat lisästä toisella sisääntuonnilla. Tämäkään ei toiminut.

Tässä vaiheessa tuli selväksi, että Logstash ei ollut yksinään tarpeeksi hyvä näin ison tapahtuma määrän sisääntuontiin.

Filebeat

Filebeat ohjelmaa suositellaan käytettäväksi, kun lisätään isoja määriä dataa Elastic Stackiin. Filebeat huomaa, jos Logstash ei pysty prosessoimaan dataa tarpeeksi nopeasti ja hidastaan datan lähettämistä.

Filebeatillä sisääntuonti suoraan Elasticsearchiin ilman Logstashiä tuotti ongelmia SSL:n kanssa, joka antoi alla olevan virhe ilmoituksen.

> ERROR instance/beat.go:971 Exiting: error initializing publisher: (assert) value of type 'bool' not convertible into unsupported go type 'tlscommon.Config' accessing 'output.elasticsearch.ssl' (source:'filebeat.yml')

Sekä Filebeatistä Logstashin kautta Elasticsearchiin sisääntuonti epäonnistui oudon yhteys ongelman takia.

Ongelman lopputulos

Toimeksiantajan kanssa tehtiin päätös, että sisääntuonti ongelmaa ei ole kannattavaa alkaa ratkaisemaan, koska siihen ja muihin järjestelmä ongelmiin oli käytetty tässä vaiheessa jo niin paljon aikaa ja vaivaa ja se ei estänyt ohjeistuksen tekemistä, joka kumminkin oli opinnäytetyön päätarkoitus. Myös puuttuvat tapahtumat eivät olleet pääasiassa 200-sarjaan liittyvistä tapahtumista, joten ongelma ei ollut niin suuri miltä se aluksi vaikutti.

Filebeat ongelmat todennäköisesti johtuivat ohjelmiston versiosta tai ohjelmien versioitten yhteensopivuudesta. Tämän lisäksi kokemus Elastic Stackistä ja etenkin toimeksiantajan Elastic Stack pohjasta oli todella rajallinen, joka olisi hankaloittanut korjaamista vielä enemmän.

Lyhykäisyydessä normaalia tilanteessa Filebeat sisääntuonnissa ei pitäisi olla mitään ongelmaa.

Ongelman ”kierto”

Ongelma kierrettiin ohjeistusta tehdessä tekemällä haut myös Splunk Enterprisessä. Nämä haut olisi tehty joka tapauksessa ohjeistusta tehdessä mutta tällä tavalla nähtiin, jos tuloksista puuttui tapahtumia. Kun hakutuloksista puuttui tapahtuma tai tapahtumia, ulostuotiin ne puuttuvat tapahtumat Splunk Enterprisellä erikseen.

Taulukossa 5 on lisäulostuonnit Splunk Enterprisestä, joita tarvittiin ohjeistuksen tekemiseen.

Taulukko 5. Lisäulostuonnit

Ulostuonin nimi	Haku komento Splunk Enterpisessä
botsv3part6	index=botsv3 earliest=0 sourcetype="hardware"
botsv3part7	index=botsv3 earliest=0 frothlywebcode AND PUT
botsv3part8	index=botsv3 earliest=0 frothlywebcode AND *txt
botsv3part9	index=botsv3 earliest=0 coin OR coins OR coinhive
botsv3part10	index=botsv3 earliest=0 coinhive
botsv3part11	index=botsv3 earliest=0 sourcetype=aws:cloudtrail AND user_type=IAMuser AND errorMessage=*
botsv3part12	
botsv3part13	
botsv3part14	index=botsv3 earliest=0 *.jpeg*
botsv3part15	index=botsv3 earliest=0 sourcetype="aws:cloudtrail" AND additionalEventData.MFAUsed=*

5.4.4 Jälkikäteen korjattu sisääntuonti

Sisääntuoti korjattiin jälkikäteen, kun aikaa jäi, vaikka toimeksiantajan kanssa keskusteltiin siitä, että tätä ei tarvitse korjata. Tämä tehtiin eri järjestelmässä omalla tietokoneella Oracle VM VirtualBoxissa Ubuntu 20.04 käyttöjärjestellä käyttäen Filebeatiä. Tämän järjestelmän image toimitettiin toimeksiantajalle.

Sisääntuonti tehtiin suoraan Elasticsearchiin. Filebeatin konfiguraatioon tarvitsi laittaa muutama komento, jotta se lukee JSONia kunnolla. Kohta **json.keys_under_root: true** käytännössä jättää JSONissa json sisennyksen pois esimerkiksi kentästä json.result.host tulee result.host. Sekä **json.override_keys: true** kohta laittaa mahdolliset kentät tiedostosta yli kirjoittamaan Filebeatin lisäämät kentät konflikti tilanteessa. Muuten konfiguraatio kohdat toimivat samalla tavalla kuin aikaisemmat. **Type** kertoi datan tyypin, **path** kertoi tiedoston sijainnin, **index** asetti tapahtumille indeksin ja **output.elasticsearch: hosts: ["localhost:9200"]** kertoo mihin data lähetettiin.

```
filebeat.inputs:
- type: log
  paths:
    - /home/elk/Desktop/botsv3/full/botsv3part1.json
  json.keys_under_root: true
  json.message.key: "message"
  json.override_keys: true
  index: "bots1"

output.elasticsearch:
  hosts: ["localhost:9200"]
```

Sisääntuontia tehdessä huomattiin, että Filebeatsillä oli ongelmia jäsentää ja lukea muutamia kenttiä datasta. Onneksi nämä kentät todettiin olevan turhia harjoituksen kannalta ja pääasiassa muutenkin. Nämä tiputettiin Filebeatissä jäsentämisen aikana processors kohdan drop_fields ominaisuudella. Drop_fields kohtaan tehtiin asetus, jossa Filebeats tiputtaa tapahtumista listan ongelma kenttiä, jotka näkyvät konfiguraatiossa alla.

```
processors:
  - drop_fields:
      fields: ["result.timestamp","result.endtime","result.backupUsage{}.lastCompletedBackupDate","result.res-
ponseElements","result.requestParameters","result.lastLoginDate","result.instance_profile","result.location"]
ignore_missing: false
```

Huomioita korjatusta sisääntuonnista liitteessä 2.

6 BOTSv3 ohjeistus

Ohjeistuksesta tehtiin neljä osainen. Ensimmäinen osa sisälsi johdannon, ohjeita hakuun ja ohjeistuksen Data Tablen käyttöön. Toinen osa sisälsi kysymykset ja vastaukset. Kolmas osa sisälsi kysymykset ja vihjeitä niihin ja Neljäs osa sisälsi kysymykset ja ohjeistuksen niitten ratkaisemiseen.

6.1 Kibanan käyttöohjeita BOTSv3 tekoon

Ohjeistuksen tekeminen alkoi käymällä peruskäyttöohjeita Elastic Stackiin. Tarkoituksena on antaa harjoituksen tekijälle valmiit ohjeet, miten harjoitusta tehdään, jotta harjoituksen tekijä ei kuluta aikaa miettimään esimerkiksi minkälainen hakusyntaksi on kyseessä tai miten saadaan Elastic Stack luottelemaan kenttien eri arvot.

6.1.1 Haku

Ohjeistuksessa on kerrottu missä muodossa ja millä tavalla hakuja voi suorittaa.

Ohjeistuksen tekeminen alkoi käymällä peruskäyttöohjeita Elastic Stackiin. Elastic Stack käyttää hakusyntaksina Kibana Query Languageta (KQL) jonka hakusyntaksi on hieman erilainen mitä jois-sakin muissa hakukoneissa. Tämän ohje näkyy kuviossa 5.

Haussa on hyvä käyttää AND komentoa. Tämä tekee sen, että jokaisen hakuehdon tulee täyttyä. Tällä voidaan putkittaa useampi hakuehto putkeen esimerkiksi **result.field1 : foo AND result.field2 : fee AND result.field3 : faa**

Haussa voi myös käyttää OR komentoa. OR komento tekee sen, että jonkin hakuehdon tulee täyttyä. Esimerkiksi **foo OR fee** antaa hakutulokset, joissa on jompikumpi. OR merkki ei ole pakollinen käyttää saman tuloksen saa hakemalla **foo fee**.

Haku **foo AND fee faa AND fuu** hakee tulokset, joissa on foo ja fuu sekä fee ja/tai faa.

OR merkkiä käyttäessä haussa **foo AND fee faa AND fuu** pitää hakuun ottaa OR arvojen ympärille sulut () **foo AND (fee OR faa) AND fuu.**

Kuvio 5. Kibana Query Language AND ja OR syntaksi ohje

Kibana Query Language sisältää outouksia, kun käytetään erikoismerkkejä kuten wildcard merkkiä ”*” ja kaksois pistettä ”:”. Erikoismerkki syntaksista tehty ohje näkyy kuviossa 6.

Haussa voi käyttää wildcard merkkiä *. Esimerkiksi *txt. On hyvä tietää että, jos hakuun laitetaan heittomerkit “” wildcard haun päällä, haku muuttuu. Esimerkiksi jos haetaan txt tiedostoja haku sanat *txt antaa enemmän tuloksia kuin ”*txt”. Wildcard merkkiä voi myös käyttää fieldissä esim. **result.field1 : foo***, mutta silloin ei saa käyttää heittomerkkejä.

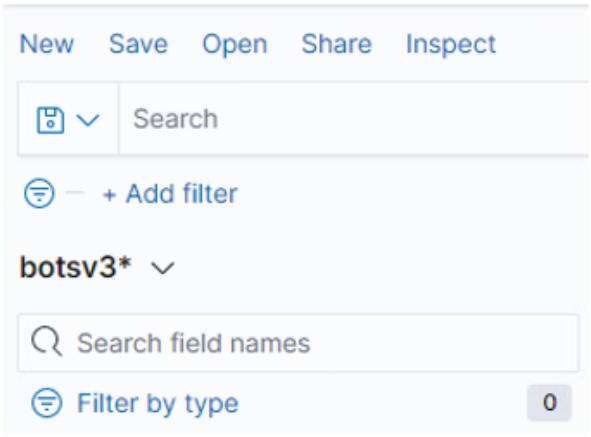
Jos haetaan fieldiä jonka arvossa on erikoismerkkejä tulee arvo laittaa heittomerkkien sisälle. Esimerkiksi **result.field1 : ”foo:fee”**

Kuvio 6. Kibana Query Language erikoismerkki syntaksi ohje

6.1.2 Kenttähaku

Kenttien haulle on Kibanassa oma kohta, johon tehtiin oma lyhyt ohje, joka näkyy kuviossa 7.

Fieldejä voidaan hakea Discover ikkunan vasemmasta reunasta niille varatulla hakukentällä. Huomaa että kenttänimet välittävät isoista ja pienistä kirjaimista, FOO ja foo eivät anna samaa tulosta.



Kuvio 7. Kenttähaku ohje

6.1.3 Visualisointityökalu Data Table

BOTSv3 teossa tarvitaan mahdollisuus nähdä kaikki kentän mahdolliset arvot. Ohjeistuksessa ohjattiin käyttämään tässä Data Table visualisointityökalua. Data Tablen käyttöönotto ja löytäminen on tarpeeksi monimutkaista, että siihen tehtiin erikseen ohje. Kirjallinen osa ohjeesta kuviossa 8.

1. Navigoi itsesi Dashboardin Vizualize valikkoon
2. Valitse luo uusi visuaalisaatio
3. Valitse uudeksi visuaalisaatioksi Data Table (Uusissa Elastic Stack versiossa Aggregation based alla)
4. Valitse BOTSv3:n Index Pattern

Kuvio 8. Data Tablen löytämisohje

Data Tablen käyttö itsessään on myös hieman monimutkaista ja siihen voi harjoituksen tekijä kuluttaa turhaa aikaa, jotta sen saa näyttämään mitä halutaan. Tämän takia siitäkin tehtiin ohje, miten saadaan Data Table näyttämään kentässä olevat eri arvot ja miten saadaan näkyviin uniikki arvojen määrä kentässä per haluttu eri kenttä. Kuviossa 9 on kirjallinen ohje, miten saadaan Data Table näyttämään mahdolliset eri sourcetype arvot.

Tässä kohtaa näytetään miten saa Data Tablen listaamaan eri arvot fieldissä. Esimerkkinä käytettään miten saadaan listattua eri sourcetypet.

Listan alla on kuvat ikkunoista

1. Vallitse Metrics ikkunaan Aggregation kohtaan Count.
2. Klikkaa Buckets kohdasta Add ja valitse Splitrows
3. Valitse Buckets ikkunan Aggregation kohtaan Terms
4. Valitse Buckets ikkunan Field kohtaan haluttu field. Tässä taupauksessa result.sourcetype.keyword.
5. Vaihda Size määrä isompaan. Size rajoittaa tulosten määrän
6. Vaihda halutessasi Order by kohdasta toinen järjestys
7. Painamalla Update nappulaa saat listan kaikista sourcetypeistä

Kuvio 9. Kirjallinen Data Table ohje

6.1.4 Tarpeettomien kenttien suodatus

Seuraavaksi tarpeelliseksi nähty ohje oli BOTSv3:n kannalta turhien kenttien suodattaminen pois hakutuloksista, näihin kenttiin kuului Splunkin jäsentäessä tekemiä kenttiä ja Elastic Stackin lisäämiä metadata kenttiä. Mukana on myös lista kentistä, jotka ovat nähty tarpeettomiksi.

Taulukko 6 on kentistä, jotka voidaan suodattaa Kibana 7.9.2 versiossa.

Taulukko 6. Suodatettavaksi ohjeistetut kentät

Kentän nimi	Kentän sisältö
@timestamp	Elastic Stack sisään tuonti aikaleima
@version	Elastic Stackin lisäämä versio numero tapahtumasta
host	Päätelaite, joka sisään tuonnin teki.
offset	Tavu etäisyys riville, jossa tapahtuma oli
path	Sijainti, josta Logstash on ottanut tiedoston käsittelyyn
result.linecount	Rivien määrä tapahtumassa, kun se on tiedostossa
result.punct	Tapahtumassa kenttien välissä olevat merkit kuten / tai (, jotka on jäsennetty omaan kenttään
result.splunk_server	Splunk serveri, josta tapahtumien ulostuonti tehtiin

Taulukko 7 on kentistä, jotka halutaan suodattaa pois mutta niitä ei voi suodattaa Kibanan 7.9.2 versiossa.

Taulukko 7. Suodatettavaksi ohjeistetut kentät, joita ei voitu suodattaa

Kentän nimi	Kentän sisältö
_id	Elasticsearchin tekemä uniikki ID
_index	Indeksi, joka on asetettu Logstash sisääntuonnissa
_score	Elasticsearchin antama arvo tiedostelun relevanttisuudelle
_type	Dokumentin data tyyppi, jonka Elasticsearch antaa

Kuviossa 10 on ohje kenttien suodattamiseen pois hausta.

1. Navigoi itsesi Management >> Stack Management
2. Kibana >> Index Patterns
3. Valitse Index Pattern jossa BOTSv3 tiedot ovat
4. Valitse ikkuna Source Filters
5. Kirjoita ei haluttu field ja paina Add

Kuvio 10. Kenttien suodatus ohje

6.2 Kysymys muutokset

BOTSv3 200-sarjassa on 25 kysymystä, joista 224 ei ollut toteutettavissa Elastic Stackillä ja kysymyksiin 212 ja 217 muokattiin mukaan paremmin sopiva versio Elastic Stackille. Kaikki kysymykset löytyvät liitteestä 1.

6.2.1 Kysymys 212 ja siihen tehty muutos

> Using Splunk's event order functions, what is the first seen signature ID of the coin miner threat according to Frothly's Symantec Endpoint Protection (SEP) data?

Kibanassa ei ollut mahdollisuutta käyttää Splunkin tapahtuma järjestys funktiota tai vastaavanlaista, joka toisi samalaisen tuloksen. Splunkin tapahtuma järjestys funktio, jota tässä on tarkoitus käyttää, on ”first”. Tämä funktio järjestää hakutulokset sen mukaan, miten Splunk löytää ne. Tämä on oikeassa maailmassa aika turha, joten se vain korvattiin ajallisesti ensimmäisellä tapahtumalla. Tämän kysymyksen vastaus tapahtuma oli tapahtunut samaan aikaan kuin toinen tapahtuma, jossa oli eri Signature ID. Tästä syystä hyväksytään ohjeistuksessa kummastakin tapahtumasta saatu vastaus. Alla ohjeistukseen lisätty korjaus.

> Kysymyksessä haetaan Splunkin First komennolla ensimmäisenä saatua vastausta, tätä komentoa ei ole Kibanassa mahdollista käyttää, joten Kibanan kummankin kahden ensimmäisenä ajallisesti listaaman eventin signature ID on hyväksyttävä vastaus.

6.2.2 Kysymys 217 ja siihen tehty muutos

> What kind of Splunk visualization was in the first file attachment that Bud emails to Frothly employees to illustrate the coin miner issue? Answer guidance: Two words. (Example: choropleth map)

Elastic Stack ei osannut jäsentää sähköpostiviestejä kunnolla BOTSv3 datasta. Tämän takia sähköposti ja siihen liittyvät tiedot olivat jäsennetty yhden kentän sisällä, joka oli result.content{}. Tämä kenttä ei sisältänyt myöskään kaikkea dataa. Kysymyksessä kysytty tiedoston kuva oli muutettu Base64 koodi kielelle, tätä koodi pätkää ei löytynyt siitä kentästä vaan se oli tapahtuman raaka data kentässä. Elastic Stackin jäsennys ongelman takia koodi pätkän kaivaminen ja sen koodauksen muuttaminen kuvaksi voisi olla harjoituksen tekijälle erittäin turhauttavaa ja aikaa vievää sekä tämän lisäksi oikeassa maailmassa tämänlaista jäsennys ongelmaa ei pitäisi olla hyvin toteutetussa järjestelmässä.

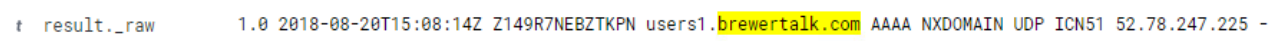
Kysymyksen muutetussa versiossa halutaan tietää mikä on kyseisen tiedoston koko nimi. Kysymyksen muotoilu jouduttiin tekemään hieman huonolta kuulostavalla tavalla, jotta harjoituksen tekijät eivät voisi oikaista vaiheita ja tehdä kysymystä eri tavalla kuin BOTSv3 tekijät ovat alun perin suunnitelleet. Alkuperäinen vastaus oli ”Column Chart” joka on tiedostossa nimeltä image003.jpeg. Kysymys haluttiin kirjoittaa tavalla, jolla tiedoston pääte olisi hankalampi arvailla ja hakea vastausta sitä kautta. Alla muutettu kysymys.

> What is the name of Splunk visualization file that was in the first file attachment that Bud emails to Frothly employees to illustrate the coin miner issue?

6.2.3 Kysymys 224

> Frothly uses Amazon Route 53 for their DNS web service. What is the average length of the distinct third-level subdomains in the queries to brewertalk.com? Answer guidance: Round to two decimal places. (Example: The third-level subdomain for my.example.company.com is example.)

Kysymyksessä 224 haluttiin tietää keskivertopituus uniikeille kolmannen tason aliverkkotunnuksille. BOTSv3 datassa ei ole jäsennetty verkkotunnusta tämänlaisessa tapauksessa lainkaan, vaan ne on jätetty raaka data kenttään, joka näkyy kuviossa 11. Tästä syystä Kibanalla ei voida käsitellä domain nimiä muulla tavalla kuin haulla.



Kuvio 11. Raaka data kenttä, jossa verkkotunnus on

Tästä kentästä ei pystytä erottamaa kolmannen tason aliverkkotunnusta taikka ottamaan niitten uniikki määriä verkkotunnuksesta. Tästä ei voi myöskään laskea edes verkkotunnusten määrää, joissa on kolmas osa.

Datan tyyli tekee tästä kysymyksestä mahdottoman tehdä Elastic Stackillä ja muutokset kysymykseen eivät ole mahdollisia, jotta kysymys pysyisi edes lähellä alkuperäistä. Joten tämä kysymys on jätetty tyhjäksi ohjeistuksessa ja tarkoitus ohittaa harjoitusta tehdessä.

6.3 Ohjeistus yleisesti

Opinnäytetyön pää tarkoitus oli tehdä ohjeistus BOTSv3 200-sarjaan. Tämän ohjeistuksen tarkoitus oli antaa kokemattomallekin harjoituksen tekijälle mahdollisuus tehdä harjoitus seuraten ohjetta ja silti päästä itse miettimään mitä tekee. Ja sammalla opetella Elastic Stackillä data-analyysiä.

Ohjeistus tehtiin ratkaisemalla kysymys monella eri tavalla ja valitsemalla niistä sopivin tapa. Sopivimmaksi tavaksi valittiin tapa, joka vastaa eniten tapaa, jota voisi käyttää oikeassa maailmassa kumminkin olematta tekemättä ylimääräisiä hakuja, joita ei voi toteuttaa tässä datasetissä. Esimerkiksi joissakin kysymyksissä vastaukseen päästiin pienemmällä määrällä rajauksia mitä todennäköisesti vaatisi oikeanmaailman tapauksessa, mutta ohjeistukseen ei olisi ollut pätevää ottaa mukaan rajauksia, jotka eivät rajaa tuloksia. Tämä ongelma ratkaistiin joissakin kysymyksissä etsimällä ohjeistuksiin kohta missä kyseisen kysymyksen päätapahtuma on tapahtunut, joten harjoituksen tekijä voi katsoa halutessaan mahdollisia lisä rajauksia.

Ohjeistusta tehdessä otettiin myös huomioon se, että kysymysten vastauksiin pystyi välillä oikomaan tavalla, joka ei ole mahdollista oikean maailman isommalla data määrällä. Tämmöisiä tapoja ohjeistuksessa vältettiin parhaaksi katsotun tavan mukaan. Kuitenkin ottaen huomioon edellä mainitun ei turhia rajauksia asian.

Ohjeistus tehtiin käyttäen ensin Elastic Stackiä ja tekemällä samat tai saman kaltaiset haut Splunkilla. Tällä tavalla voitiin varmistaa, että Elastic Stackin tuloksissa ei ollut minkäänlaisia outoja tuloksia tai puutu tapahtumia.

Ohjeistuksessa ei ole selitetty teknologioita ja käsitteitä muuten kuin kysymys 200:ssa selitettiin lyhyesti AWS CloudTrail. Tämä johtuu siitä, että se helpottaisi tekijälle harjoitusta liikaa ja tällä tavalla rohkaisee internetillä tutkimaan eri teknologioita ja käsitteitä itse.

6.4 Ohjeistuksen teko

Tässä kohdassa on käyty ohjeistuksen tekoprosessia lyhyesti mutta ei itse ohjeistusta. Ohjeistus lopputulos oli haluttu pitää ainoastaan toimeksiantajalla ja sen takia ei ole haluttu käydä ohjeistusta kokonaan läpi tai tarkasti läpi tässä. Vaan käydään läpi tekoprosessia ja mitä oli ohjeistuksessa ohjeistettu harjoituksen tekijä tekemään sekä yksi ohjeistus esimerkki.

6.4.1 Aloitus

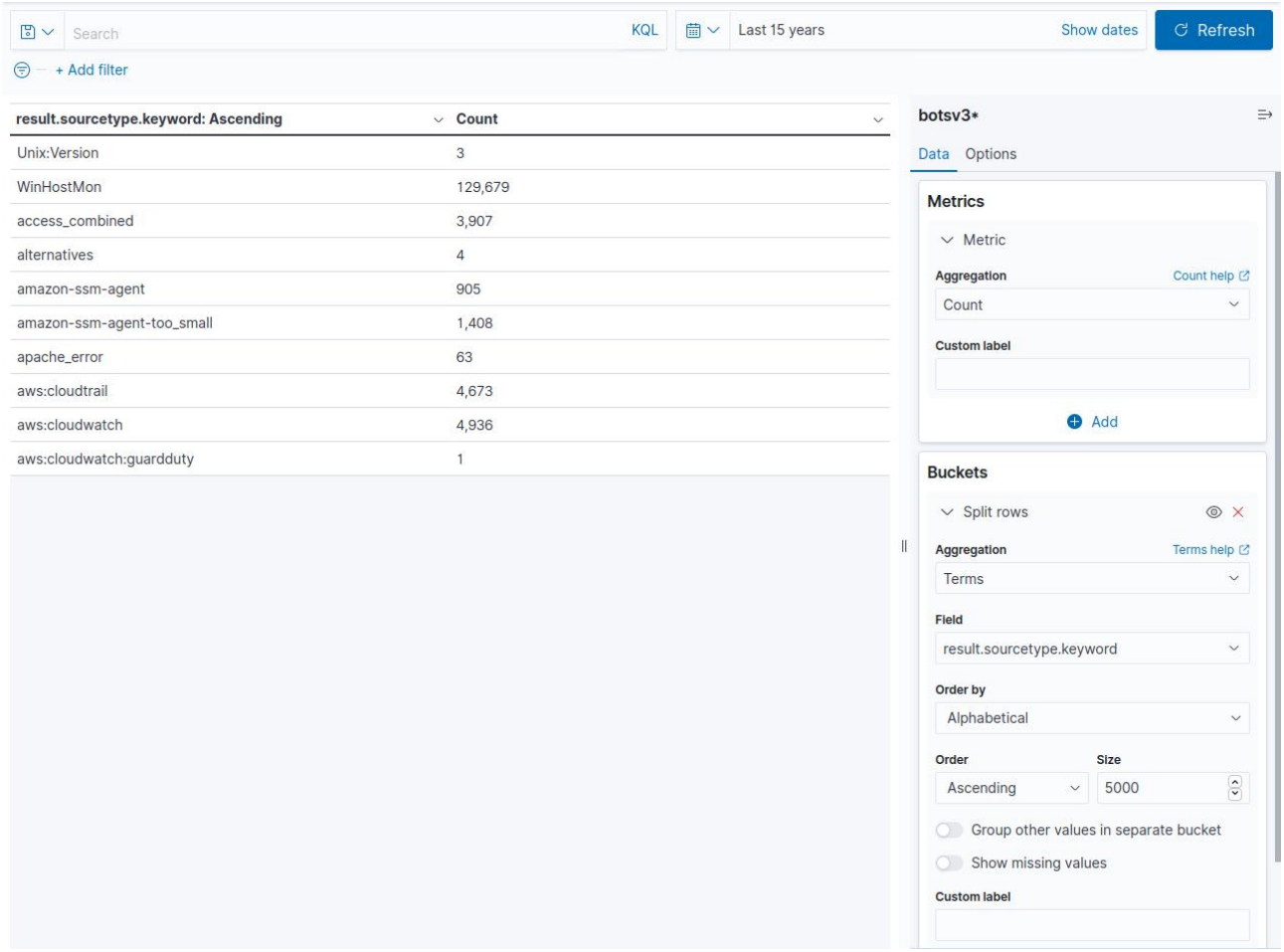
Kysymyksien ohjeistusta ennen tekemistä kysymys tietenkin selvitettiin itse. Tästä kun saatiin kysymykseen oikea vastaus, selvitettiin muita ratkaisuja ja kysymyksen vastauksen selviytyessä ”reverse engineerattiin” muita vaihtoehtoja. Mahdollisista vaihtoehdoista valittiin paras, parhaimmaksi tavaksi valittiin tapa, joka nähtiin olevan yksinkertainen ja looginen harjoitusta tekevälle henkilölle.

Ohjeistuksen tarkoitus on opastaa harjoituksen tekijä harjoituksen läpi niin sanotusti kädestä pitäen. Mutta kohdissa, joissa on mahdollista jättää pieniä kohtia, jossa harjoituksen tekijä pääsee miettimään ratkaisua itse, vaikka tekisi harjoitusta suoraan ohjeistuksesta.

6.4.2 Mitä ohjeistetaan tekemään ”normaalissa” kysymyksessä

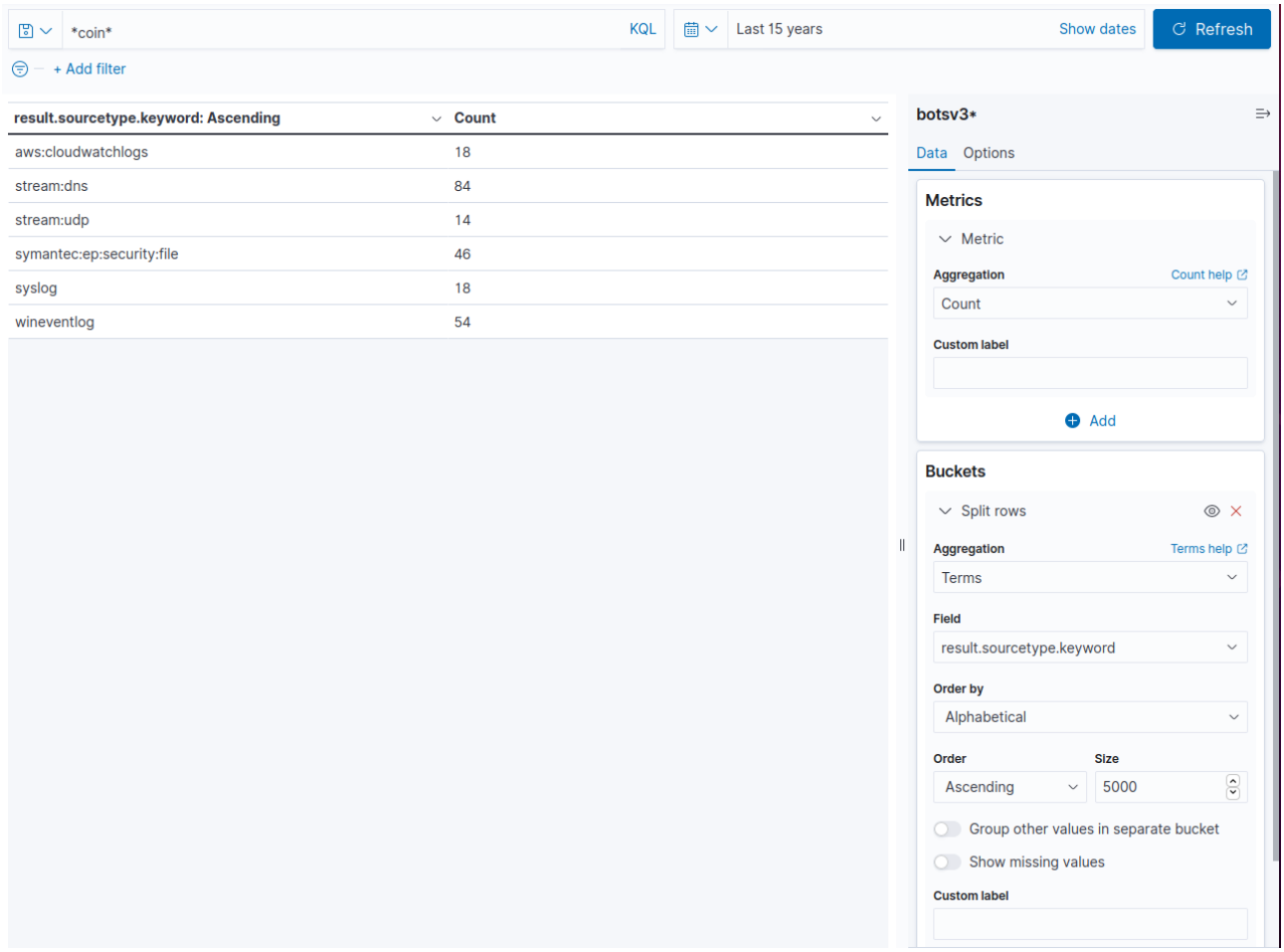
Jokainen BOTSv3 200-sarjan kysymys kuuluu johonkin sourcetypeen josta kerrottiin kappaleessa 3.2.4. Oheistuksessa aina ensimmäisenä haetaan sourcetype jos tämä on mahdollista selvittää.

Esimerkiksi kysymys 200 haluttaan tietää ketkä kaikki IAM (IAM kappaleessa 4.1.7) käyttäjät ovat ottaneet yhteyttä yhtiön AWS ympäristöön. Elastic Stackissä ei ollut mahdollisuutta Discover työkalulla (jolla haut tehdään) esittää mahdollisia eri arvoja kentässä paitsi viidestä arvosta, joita on eniten 500 ensimmäisessä tapahtumassa. Ohjeistuksessa päädyttiin harjoituksen tekijää ohjeistamaan käyttämään Data Tablea joka näkyy kuviossa 12.



Kuvio 12. Kysymys 200 sourcetypet Data Tablessa

Kysymys 210 halutaan tietää mikä päätelaite louhi Monero kryptovaluuttaa. Tämmöisessä tapauksessa kun ei ole suoraan selvää minkälainen sourcetype on haluttu, tämä ohjeistuksessa ohjeistettu rajaamaan avoinsanoilla sourcetypejen määrää. Ottamalla avainsana ”*coin*” hakuun saadaan paljon pienempi määrä mahdollisia sourcetypejä joista on helpompi selvittää haluttu. Kuviossa 13 näkyy pienempi määrä sourcetypejä verrattuna kuvioon 12, kuviosta 13 puuttuu muutama merkityksetön sourcetype aikaisemmin mainitun sisääntuonti ongelman takia.



Kuvio 13. Kysymys 210 sourcetypet Data Tablessa

Seuraavaksi yleensä ohjattiin ohjeistuksessa käyttämää avainsanoja kysymyksestä. Tällä tarkoitetaan, että käydään kysymystä läpi ja otetaan sieltä sanoja, jotka auttavat rajaamaa hakutuloksi. Esimerkiksi kysymys 209 joka on alla, hyviä avainsanoja, joita ohjeistetaan käyttämään ovat: auto scaling, packages, cloud initialization ja installed.

> When a Frothly web server EC2 instance is launched via auto scaling, it performs automated configuration tasks after the instance starts. How many packages and dependent packages are installed by the cloud initialization script?

Tämän jälkeen pitäisi harjoituksen tekijällä olla tuloksia, jotka liittyvät jollakin tavalla kysymykseen. Joissakin kohdissa ohjeistuksessa neuvotaan harjoituksen tekijää etsimään hyviä kenttiä joko selaamalla tapahtumia tai hakemalla kenttähausta. Esimerkiksi aiemmin mainitusta 200 kysymyksestä sen jälkeen, kun sourcetype ja avainsana IAM on otettu hakuun, ohjeistetaan hakemaan IAM

liittyviä kenttiä kenttähausta ja kun niitä ei löydy, ohjeistetaan selaamaan tapahtumia läpi ja etsimään kenttiä, on jotain IAM käyttäjään liittyvää tai jossa IAM käyttäjä on ilmoitettu.

Tässä kohtaa joissakin kysymyksissä saa tekijä jo selvitettyä oikean varman vastauksen. Joissakin kohdissa kuten aikaisemmin mainitussa kysymys 210, on ohjeistettu harjoituksen tekijää etsimään varmistus, että kyseinen vastaus on oikein. Esimerkiksi kysymys 210 tilanteessa on kerrottu, että mikä päätelaite on louhinut kryptovaluutta Moneroa katsomalla onko vastaus päätelaite ottanut itse yhteyttä louhinta sivustoon. Ja lisäksi on näytetty miten ja missä kohtaa tämä näkyy tapahtumassa. Kuviossa 14 näkyy se, että päätelaite ottaa yhteyttä louhinta sivuun.

Kuvio 14. Päätelaitteen yhdeyden oton louhinta sivuun todennus

6.4.3 Esimerkki ohjeistus

Alla kysymys 200 ja sen ohjeistus kokonaisuudessaan. Kysymys 200 tehtiin hieman enemmän auttavasti/selittäen kuin muut kysymykset, jotta harjoituksen tekijän on helpompi ymmärtää alkuun harjoitusta ja ohjeistusta

> **List out the IAM users that accessed an AWS service (successfully or unsuccessfully) in Frothly's AWS environment? Answer guidance: Comma separated without spaces, in alphabetical order. (Example: ajackson,mjones,tmiller)**
>
> Tämän kysymyksen vastauksen löytämiseen on monta tapaa. Helppo tapa on aluksi selvittää sourcetype jos, sitä ei valmiiksi tiedä.
>
> Data Tablella aggregoimalla **result.sourcetype.keyword** ja saa näkyviin kaikki sourcetypet. Listaamalla aakkosjärjestyksessä tulokset saa näkymää Amazon ja AWS tulokset heti alkuun, joten listaa ei tarvitse alkaa käymään läpi. Näistä vaihtoehdoista voi googlaamalla tai manuaalisesti testaamalla selvittää halutun sourcetypen
>
> Tarvittava sourcetype on aws:cloudtrail. Aws:cloudtrail sisältää API toiminta AWS käyttäjällä.

Seuraavaksi haetaan **result.sourcetype : "aws:cloudtrail"** Discover toiminnolla. Kysymyksessä haluttiin hakea kaikki ”IAM users” joten sivun vasemmassa reunasta on kannattavaa hakea user fieldejä. Toinen vaihtoehto on hakea IAM useria suoraan hausta ja etsiä sitä kautta sopiva field **result.sourcetype : "aws:cloudtrail" AND iam ”iam user” iamuser**.

Kaksi fieldiä joista löytyy IAMuser arvo on result.user_type, result.useridentity.type.

Seuraavaksi voi hakea joko fieldejä vasemmasta reunasta tai tutkimalla eventtiä ja etsiä fieldin jossa näkyy käyttäjä nimi. Result.userIdentity.userName ja result.userName siältävät käyttäjä nimen.

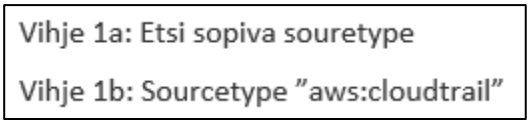
Kun on selvillä kaikki tarvittavat fieldit ja arvot voidaan ne syöttää Data Tableen. Hakemalla **result.sourcetype : "aws:cloudtrail" AND result.user_type: "IAMUser"** ja aggregoimalla **result.userName.keyword** saadaan näkyviin kaikki IAM käyttäjät jotka ovat yhdistäneet AWS palveluun.

Etsimällä yllä olevalla hakusanalla normaali haussa eventin voidaan ottaa hakuun mukaan AWS account ID. Account ID löytyy sekä result.aws_account_id , result.recipientAccountId ja result.recipientAccountId fieldeistä. Account ID voidaan ottaa hakuun mukaan ja tarkistaa että ei ole otettu yhteyttä kuin yhteen AWS enviromenttiin.

6.4.4 Vihjeet

Vihjeitten teko suoritettiin käymällä läpi kysymys, johon se kuului ja ottamalla tärkeimmät kohdat, jossa harjoituksen tekijä voi jäädä jumiin. Vihjeitten tarkoitus oli siis yksinomaan tuupata harjoituksen tekijää eteenpäin eikä kertoa, miten vastaus saadaan.

Esimerkkinä kun kysymyksen 200 ohjeistuksessa ohjataan sopivan sourcetypen löytämiseen, tehtiin siitä kaksiosainen vihje kuviossa 15.

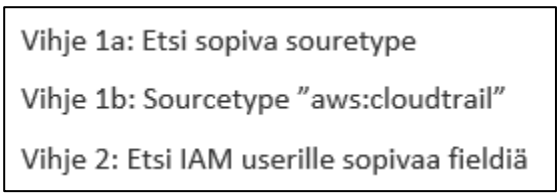


Kuvio 15. Kysymys 200 vihje 1a ja 1b

Ja toisena vihjeenä on käytetty ohjeistuksen kohtaa, jossa kerrotaan hakea kenttää, jossa IAM käyttäjä voisi olla. Tämä näkyy kuviossa 16.

Kuvio 16. Kysymys 200 vihje 2

Tämän enempää ei katsottu kysymykseen 200 tarvittavan vihjeitä. Eli koko vihje osuus kysymys 200 on kuviossa 17.



Kuvio 17. Kysymys 200 kaikki vihjeet

7 Testi

7.1 Taustatietoja

Toimeksiantajalta ei ollut mahdollista saada testaajia ohjeistukseen joten, ohjeistuksesta toteutettiin yhden henkilön kokoinen testi.

Testihenkilönä toimi Jyväkylän ammattikorkeakoulun tieto- ja viestintätekniikkaa neljättä vuotta opiskeleva henkilö, jolla ei ollut aikaisempaa kokemusta SIEMeistä, Elastic Stackistä tai BOTSeistä, mutta hänellä oli kokemusta muista kyberharjoituksista.

Testissä testihenkilö teki BOTSv3 200-sarjan kysymykset seuraamalla ohjeistusta. Tämän lisäksi testihenkilö antoi palautetta suullisesti ongelmista ohjeistuksesta ja täytti lomaketta, jossa oli muutama kysymys ohjeistuksen laadusta.

Testin tarkoitus oli saata muitten henkilöiden mielipiteitä ohjeistuksesta ja löytää ohjeistuksessa olevia ongelmia mitä ei ohjeistuksen tekijä ei ole itse huomannut. Mahdolliset ongelmat voitiin täten vielä korjata ohjeistukseen. Samalla oli tarkoitus saada pientä tutkimus tulosta ohjeistuksen laadusta ja vaikeustasosta, vertaus pohjana pitäen IT-alan henkilöitä.

7.2 Testin tulokset

Testihenkilön täyttämä lomake löytyy kokonaisuudessa liitteestä 3. Taulukossa 8 on testihenkilön mielipiteet ohjeistuksekseen tärkeimmistä asioista, jotka kerättiin lomakkeeseen. Taulukossa 8 näkee sen, että testihenkilön kokemukset ohjeistuksesta olivat pääasiassa positiivisia. Parannettavaa testihenkilö näki kysymyksissä 201, 202, 203 ja 212.

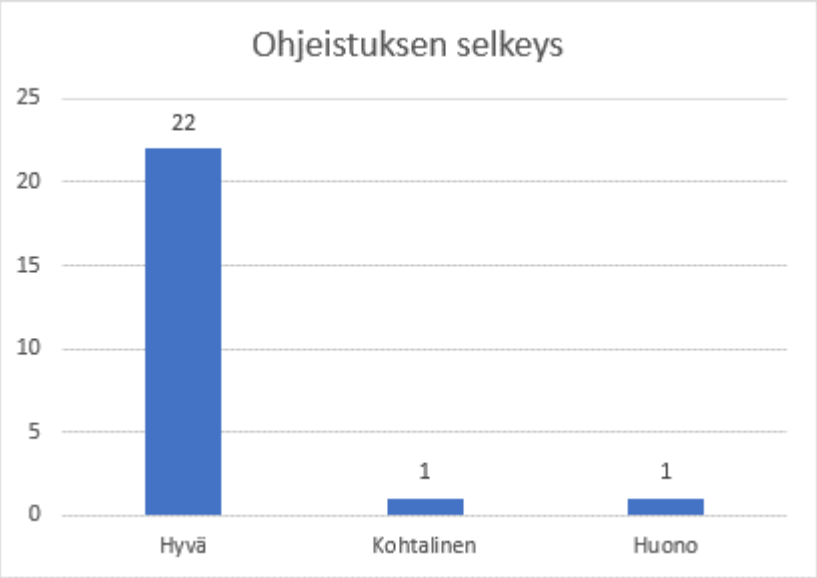
Taulukko 8. Testihenkilön mielipiteet tärkeimmistä asioista ohjeistuksessa

Kysymys	Vastaus oikein ensimmäisellä yrityksellä	Ohjeistuksen selkeys	Ratkaisu logiikka oli selvä	Yleinen laatu
200	Kyllä	Hyvä	Kyllä	Hyvä
201	Kyllä	Hyvä	Jotenkin	Hyvä
202	Kyllä	Kohtalainen	Kyllä	Kohtalainen
203	Ei	Huono	Jotenkin	Kohtalainen
204	Kyllä	Hyvä	Kyllä	Hyvä
205	Kyllä	Hyvä	Kyllä	Hyvä
206	Kyllä	Hyvä	Kyllä	Hyvä
208	Kyllä	Hyvä	Kyllä	Hyvä
209	Kyllä	Hyvä	Kyllä	Hyvä
210	Kyllä	Hyvä	Kyllä	Hyvä
211	Kyllä	Hyvä	Kyllä	Hyvä
212	Ei	Hyvä	Kyllä	Kohtalainen
213	Kyllä	Hyvä	Kyllä	Hyvä
214	Kyllä	Hyvä	Kyllä	Hyvä
215	Ei	Hyvä	Kyllä	Hyvä
216	Kyllä	Hyvä	Kyllä	Hyvä
217	Kyllä	Hyvä	Kyllä	Hyvä
218	Kyllä	Hyvä	Kyllä	Hyvä
219	Kyllä	Hyvä	Kyllä	Hyvä
220	Kyllä	Hyvä	Kyllä	Hyvä
221	Kyllä	Hyvä	Kyllä	Hyvä
222	Kyllä	Hyvä	Kyllä	Hyvä
223	Kyllä	Hyvä	Kyllä	Hyvä
225	Kyllä	Hyvä	Kyllä	Hyvä

Kuviossa 18 näkee koottuna että, tekijä sai oikean vastauksen ensimmäisellä yrityksellä kaikissa muissa tapauksissa paitsi kolmessa. Sisääntuonti ongelma kysymyksessä 215, esti oikean vastauksen saamisen siinä kysymyksessä. Kysymyksessä 203 ohjeistus oli ongelmallinen/huono. Ja kysymyksessä 212 oli tekijällä ongelmia syntaksin kanssa, jonka takia hän sai väärän vastauksen aluksi.

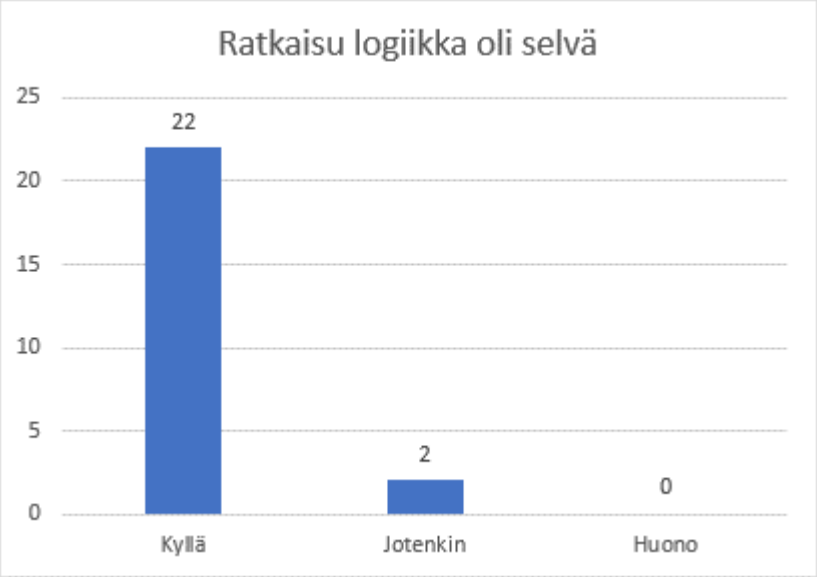
Kuvio 18. Testihenkilön vastausten oikeellisuus

Kuviossa 19 näkee koottuna ohjeistuksen selkeyden. Kohtalainen arvosanan sai kysymys 202 jossa oli huonosti kirjoitettu kohta ja puuttuva kohta. Huono arvosanan sai kysymys 203 joka oli huonosti tehty yhdessä kohdassa.



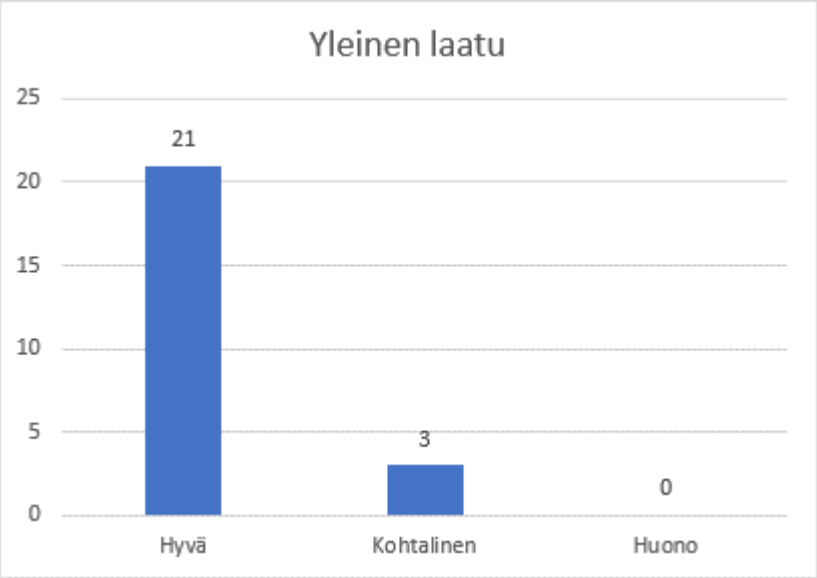
Kuvio 19. Testihenkilön mielipide ohjeistuksen selkeydestä

Kuviossa 20 näkee koottuna testihenkilön mielipiteen kysymysten ratkaisu logiikan selkeydestä. Jotenkin arvosanan sai myös aikaisemmin mainittu kysymys 203 ja kysymys 201 jossa oli huonosti kirjoitettu kohta.



Kuvio 20. Testihenkilön mielipide kysymysten ratkaisu logiikan selkeydestä

Kuviossa 21 näkee koottuna testihenkilön mielipiteen kysymysten ohjeistusten yleisen laadusta. Kohtalainen arvosanat saivat aikaisemmin mainitut kysymykset 202, 203 ja 212



Kuvio 21. Testihenkilön mielipide kysymysten ohjeistusten yleisestä laadusta

7.3 Muutokset

Tässä kohdassa on listattu muutoksia, jotka tehtiin ohjeistus kohtaan testihenkilön kirjallisen lomakkeen ja suullisten kommenttien perusteella.

Kysymys 201 vastaukseen lisättiin toinen vaihtoehto.

Kysymys 202 selvennetty ja lisätty kohta helpottamaan kysymyksen vastauksen löytämistä.

Kysymys 203 ohjeistusta paranneltu ja korjattu siinä oleva ongelma, joka teki siitä huonon.

Kysymys 203 oli muotoilu virhe. Tämä korjattiin.

Kysymys 203 oli epäselvä kohta, kun ohjattiin tekijää oikeaa tapahtumaa tutkimaan.

Kysymys 212 selvennettiin hieman.

Kysymys 213 muutettiin muotoilua selvemmäksi.

Kysymys 215 korjattiin viallinen haku.

Kysymys 220 ohjeistukseen lisätty kohta helpottamaan kysymyksen vastauksen löytämistä.

Ohjeistuksen alkuun KQL syntaksi kohtaan lisättiin kohta selventämään syntaksia enemmän.

8 Loppupohdinta

Opinnäytetyön tavoitteena oli saada toimeksiantajalla ohjeistus BOTS tyyliselle harjoitukselle Elastic Stackiin, jolla halutut harjoituksen tekijät voisivat harjoitella Elastic Stackin käyttöä ja sillä data-analysointia. Tämä saatiin toteutettua käyttäen BOTSin kolmatta versiota. Ohjeistukseen saatiin haluttu ”kädestä pidetty” versio, jossa harjoituksen tekijä voi seurata suoraan ohjeistusta ja silti itse miettiä halutessaan ratkaisuja, kun tekee harjoitusta. Sekä saatiin tehtyä haluttu vihje versio, josta halukas harjoituksen tekijä voi lukea vinkkejä, jotta pääsisi eteenpäin harjoituksessa. Tämän lisäksi tehtiin Elastic Stackistä käyttöohjeita kohdista, joita tarvitsee BOTSv3:n teossa.

Ohjeistuksen tekeminen itsessään oli suoraviivaisempaa mitä aluksi luultiin, alkuun tarvitsi opetella Kibanan haku ominaisuuksia ja tehdä BOTSva3 harjoitus itse. Tämän jälkeen tarvitsi etsi ja tehdä mahdollisia tapoja ratkaista kysymys ja valita niistä paras. Tämä tapa tehdä ohjeistus oli hidas mutta nähtiin hyväksi teko vaiheessa ja sen jälkeenkin.

Itse ohjeistuksesta tuli opinnäytetyön tekijän mielestä hyvä ja sellainen mitä alun perin suunniteltiin toimeksiantajan kanssa. Vaikka toimeksiantajalta ei saatu testaajia ohjeistuksen kanssa tekoon pidettiin opinnäytetyö projektin aikana palavereita, jossa kysyttiin mielipidettä ohjeistuksesta ja tuotoksesta yleisesti ja nämä mielipiteet olivat positiivisia. Tämän lisäksi testihenkilön mielipiteet ohjeistuksesta olivat pääasiassa positiivisia ja testihenkilön ilmoittavat ongelma kohdat korjattiin tai parannettiin palautteen perusteella.

Ohjelmistojen kanssa oli paljon ongelmia, jotka veivät paljon aikaa. Pää ongelmia olivat Splunk Enterprisen outo ongelma BOTSv2 ulostuonnin kanssa, jonka kanssa tapeltiin kauan ja Elastic Stackin

sisääntuonti ongelma, jonka kanssa tapeltiin vielä kauemmin. Elastic Stack sisääntuonti ongelmakin saatiin ratkaistua loppujen lopuksi eri järjestelmää käyttäen ja tämän järjestelmän image saatiin myös toimitettua toimeksiantajalle, jos he haluavat ottaa sen käyttöön jossakin muotoa.

Jatkokehityksenä tulevaisuuteen voi tulla se että, jos toimeksiantaja näkee tarvetta sen jälkeen, kun on testannut harjoitusta, tehdä ohjeistus myös BOTSv3 300-sarjaan. Tämän lisäksi, vaikka hieman opinnäytetyöhön liittymätön, olisi hyvä idea tehdä sisääntuonnista yhdellä komennolla toimiva versio, joka toimisi sekä uniikki kenttä rajoituksen kanssa että shard (suom. siru) rajoituksen kanssa.

Lähteet

Amazon EC2. N.d. Amazon EC2 tuote esittely sivu. Viitattu 23.4.2021. https://aws.amazon.com/ec2/?ec2-whats-new.sort-by=item.additionalFields.postDateTime&ec2-whats-new.sort-order=desc.

Amazon Machine Images (AMI). N.d. Amazon Machine Images dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/AMIs.html.

Andersen-Waine, B. 2015. Getting-started-with-the-elk-stack. GitHub readme dokumentti. Viitattu 22.4.2021. https://github.com/LoveSoftware/getting-started-with-the-elk-stack/blob/master/docs/Exercises.md.

Berman, D. 2020. A Beats Tutorial: Getting Started. Logz.io sivun Beats ohjeistus artikkeli. Viitattu 22.4.2021. https://logz.io/blog/beats-tutorial/.

BOTS 1.0 datasets. N.d. Lomake BOTSv1 kysymyksille ja vastauksille Splunkin verkkosivulla. Viitattu 23.4.2021. http://explore.splunk.com/BOTS_1_0_datasets.

BOTS 2.0 datasets. N.d. Lomake BOTSv2 kysymyksille ja vastauksille Splunkin verkkosivulla. Viitattu 23.4.2021. https://events.splunk.com/BOTS_2_0_datasets.

Bucket overview. N.d. AWS Buckets dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/AmazonS3/latest/userguide/UsingBucket.html.

Christensson, P. 2015. Host. TechTermsin Host määrittely sivu. Viitattu 23.4.2021. https://techterms.com/definition/host.

Christensson, P. 2019. FQDN. TechTermsin FQDN määrittely sivu. Viitattu 23.4.2021. https://techterms.com/definition/fqdn.

Codec plugins. N.d. Elasticin koodekki lisäosien listaus sivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/logstash/current/codec-plugins.html.

Community Beats. N.d. Elasticin kommuuni beatsien listaus sivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/beats/libbeat/current/community-beats.html.

Cox, J. 2020. Access Control List (ACL) – What are They and How to Configure Them! Ittsystems sivun artikkeli. Viitattu 23.4.2021. https://www.ittsystems.com/access-control-list-acl/.

Cloud-init Documentation. N.d. Cloud-initin documentointi sivu. Viitattu 23.4.2021. https://cloudinit.readthedocs.io/en/latest/#cloud-init-documentation.

CloudTrail - Digital Breadcrumbs for AWS. 2018. Blogi kirjoistus Splunkin verkkosivulla. Viitattu 23.4.2021. https://www.splunk.com/en_us/blog/security/cloudtrail-digital-breadcrumbs-for-aws.html.

Crane, C. 2020. What Is Crypto Mining? How Cryptocurrency Mining Works. Louhinta artikkeli Sectigon verkkosivulla. Viitattu 23.4.2021. https://sectigostore.com/blog/what-is-crypto-mining-how-cryptocurrency-mining-works/.

Elastic pricing. N.d. Elastic Stackin hinnoittelu sivu. Viitattu 22.4.2021. https://www.elastic.co/pricing/.

End of Life notice for Endpoint Protection Small Business Edition Cloud and Endpoint Protection Cloud. 2021. Broadcomin End of Life ilmoitus. Viitattu 23.4.2021. https://knowledge.broadcom.com/external/article/198499/end-of-life-notice-for-endpoint-protecti.html.

Filter plugins. N.d. Elasticin suodatus lisäosien listaus sivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/logstash/current/filter-plugins.html.

Frankenfield, J. 2021. Cryptocurrency. Investopedian termin selitys sivu. Viitattu 23.4.2021. https://www.investopedia.com/terms/c/cryptocurrency.asp.

Frazier, T. & Herrald, D. 2020. Boss of the SOC (BOTS) Dataset Version 3. Spunkin BOTSv3 GitHub repositorio. Viitattu 3.5.2021. https://github.com/splunk/botsv3.

Functionbeat overview. N.d. Tuotekuvaus Elasticin verkkosivulla. Viitattu 22.4.2021. https://www.elastic.co/beats/functionbeat.

Get started with your free trial. N.d. Splunkin ilmaisen testiversion rekisteröinti sivu. Viitattu 22.4.2021. https://www.splunk.com/en_us/download/get-started-with-your-free-trial.html.

Gibbins, J. 2020. Splunk BOTSv3 Write-Up. BOTSv3 läpikäynti. Viitattu 22.4.2021. https://www.jamesgibbins.com/articles/digital-forensics/botsv3/.

Herrald, D. 2019. Boss of the SOC 2.0 Dataset, Questions and Answers Open-Sourced and Ready for Download. Blogi kirjoitus BOTSista Splunkin verkkosivulla. Viitattu 23.4.2021.

https://www.splunk.com/en_us/blog/security/boss-of-the-soc-2-0-dataset-questions-and-answers-open-sourced-and-ready-for-download.html.

Herrald, D. 2020. Boss of the SOC v3 Dataset Released! Blogi kirjoitus BOTSista Splunkin verkkosivulla. Viitattu 23.4.2021. https://www.splunk.com/en_us/blog/security/botsv3-dataset-released.html.

Horovits, D. 2020. THE COMPLETE GUIDE TO THE ELK STACK. Logzin Elastic Stack ohje artikkeli. Viitattu. 22.4.2021. https://logz.io/learn/complete-guide-elk-stack/.

Input plugins. N.d. Elasticin sisääntuonti lisäosien listaus sivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/logstash/current/input-plugins.html.

Instructions for organising cyber exercises. 2020. Traficom – Kyberturvallisuuskeskuksen kyberharjoitus ohje. Viitattu 23.4.2021. https://www.kyberturvallisuuskeskus.fi/sites/default/files/media/file/Instructions%20for%20organising%20cyber%20exercises.pdf.

Introducing AWS Auto Scaling. 2018. AWS Auto Scaling palvelun esittely sivu. Viitattu 23.4.2021. https://aws.amazon.com/about-aws/whats-new/2018/01/introducing-aws-auto-scaling/.

Introducing Elastic SIEM. N.d. Elastic Stackin esittely blogisivu. Viitattu 26.3.2021. https://www.elastic.co/blog/introducing-elastic-siem.

JYVSECTEC. N.d. JYVCETEC verkkosivujen etusivu. Viitattu 22.4.2021. https://jyvsectec.fi/.

JYVSECTEC cyber range. N.d. RGCE esittely PDF. Viitattu 22.4.2021. https://jyvsectec.fi/wp-content/uploads/2018/10/JYVSECTEC-cyber-range.pdf.

JYVSECTEC Overview. N.d. JYVSECTEC esittely sivu. Viitattu 26.3.2021. https://jyvsectec.fi/about/overview/.

Kibana. N.d. Kibanan AWS esittely sivu. Viitattu 22.4.2021. https://aws.amazon.com/elasticsearch-service/the-elk-stack/kibana/.

Kibana – More examples. N.d. Kibanan dokumentointi sivu. Viitattu 22.4.2021. https://manualkibanaocds.readthedocs.io/en/latest/C2/Seccion1/3_Kibana.html.

Kovar, R. 2018. Boss of the SOC Scoring Server, Questions and Answers, and Dataset! Open-Sourced and Ready for Download. Blogi kirjoitus BOTSista Splunkin verkkosivulla. Viitattu 23.4.2021. https://www.splunk.com/en_us/blog/security/boss-of-the-soc-scoring-server-questions-and-answers-and-dataset-open-sourced-and-ready-for-download.html.

Kovar, R. 2020. Boss of the SOC V at .conf20. Blogi kirjoitus BOTSista Splunkin verkkosivulla. Viitattu 23.4.2021. https://www.splunk.com/en_us/blog/conf-splunklive/bots-v-at-conf20.html.

Krebs, B. 2018. Who and What Is Coinhive? Blogi kirjoitus Coinhivestä. Viittattu 23.4.2021. https://krebsonsecurity.com/2018/03/who-and-what-is-coinhive/.

Logstash introduction. N.d. Elasticin Logstash esittelysivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/logstash/current/introduction.html.

Long, C. 2020. WORKING THROUGH SPLUNK'S BOSS OF THE SOC - PART 1. BOTSv3 läpikäynti. Viitattu 22.4.2021. https://clo.ng/blog/bots-part1/.

Luedtke, S. 2017. Dashboard Digest Series - Episode 6: Traveling on Time with Trellis. Splunk artikkeli kojelaudasta. Viitattu 22.4.2021. https://www.splunk.com/en_us/blog/tips-and-tricks/dashboard-digest-series-episode-6-traveling-on-time-with-trellis.html.

Managing access keys for IAM users. N.d. AWS accecs key dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/IAM/latest/UserGuide/id_credentials_access-keys.html.

Marcho, C. 2019. Windows Performance Monitor Overview. Microsoftin blogi artikkeli Performance monitorista. Viitattu 23.4.2021. https://techcommunity.microsoft.com/t5/ask-the-performance-team/windows-performance-monitor-overview/ba-p/375481.

Mikkola, J. 2020. Lokianalyysi ja valvonta ELK-järjestelmää hyödyntäen. Opinnäytetyö, AMK. Laurea-ammattikorkeakoulu, liiketalouden ammattikorkeakoulututkinto, tietojenkäsittely, tradenomi. Viitattu 22.4.2021. http://urn.fi/URN:NBN:fi:amk-202002122318.

Multi-factor Authentication. N.d. AWS MFA tuote esittely sivu. Viitattu 23.4.2021. https://aws.amazon.com/iam/features/mfa/.

Network ACLs. N.d. AWS ACL dokumentointi sivu. Viitattu 24.3.2021. https://docs.aws.amazon.com/vpc/latest/userguide/vpc-network-acls.html.

Network Visibilty Module. N.d. Ciscon tuote esittely PDF. Viitattu 23.4.2021. https://www.cisco.com/c/en/us/td/docs/security/vpn_client/anyconnect/anyconnect45/administration/guide/b_AnyConnect_Administrator_Guide_4-5/nvm.pdf.

Output plugins. N.d. Elasticin ulostuonti lisäosien listaus sivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/logstash/current/output-plugins.html.

Pilici, S. 2017. How to remove JS/CoinMiner Trojan (Virus Removal Guide). Malvaretipsin ohje JSCoinminerin poistoon. Viitattu 23.4.2021. https://malwaretips.com/blogs/remove-js-coinminer-trojan/.

Pivot Manual. 2019. Splunk pivot työkalun dokumentointi ohje. Viitattu 22.4.2021. https://docs.splunk.com/Documentation/SplunkCloud/latest/Pivot/IntroductiontoPivot.

Query. N.d. Techopedia query määrittely sivu. Viitattu 23.4.2021. https://www.techopedia.com/definition/5736/query.

RFC 5321. 2008. Simple Mail Transfer Protocol IETF standardi. Viitattu 28.4.2021. https://tools.ietf.org/html/rfc5321.

RFC 7231. 2014. Hypertext Transfer Protocol IETF standardi. Viitattu 28.4.2021. https://tools.ietf.org/html/rfc7231.

Reporting manual. 2020. Splunk Enterprisen raportti työkalun dokumentointi sivu. Viitattu 22.4.2021. https://docs.splunk.com/Documentation/SplunkCloud/latest/Report/Aboutreports.

SA-ctf_scoreboard. 2020. BOTS pistetaulukon GitHub repositorion readme sivu. Viitattu 23.4.2021. https://github.com/splunk/SA-ctf_scoreboard.

Salát, M. 2019. The End of Coinhive; The end of cryptojacking? Blogi kirjoitus Avastin verkkosivulla. Viitattu 23.4.2021. https://blog.avast.com/coinhive-shuts-down.

Search tutorial. 2019. Splunk Enterprisen hakutyökalun dokumentointi sivu. Viitattu 22.4.2021. https://docs.splunk.com/Documentation/Splunk/8.1.2/SearchTutorial/Createnewdashboard.

Splunk BOTS - Boss Of The SOC (v3) Walkthrough & Analysis. 2020. BOTSv3 ohjeistus video. Viitattu 22.4.2021. https://www.youtube.com/watch?v=RXDbir6B5mE.

Splunk Enterprise overview. 2021. Splunk Enterprise esittely dokumentti. Viitattu 22.4.2021. https://docs.splunk.com/Documentation/Splunk/8.1.3/Overview/AboutSplunkEnterprise.

Spulk Products. N.d. Splunk tuotteitten esittely sivu. Viitattu 22.4.2021. https://www.splunk.com/en_us/software.html.

Stannard, E. 2020. Splunk Boss of the SOCs (BOTS) V3 — Part 1. BOTSv3 läpikäynti. Viitattu 22.4.2021. https://ellisstannard.medium.com/boss-of-the-socs-bots-v3-part-1-3a0d92b851b4.

Symantec Endpoint Protection Manager. N.d. Symentecin SEP apu sivu. Viitattu 23.4.2021. https://help.symantec.com/cs/ATP_3.2/ATP/v117615376_v127300344/Symantec-Endpoint-Protection-Manager?locale=EN_US.

Tittel, E. 2017. Symantec Endpoint Protection and the details for buyers to know. Ostajan ohje TehcTargetin verkkosivulla. Viitattu 23.4.2021. https://searchsecurity.techtarget.com/feature/Antimalware-protection-products-Symantec-Endpoint-Protection.

What are Elasticsearch Beats? N.d. Objectrocket sivun artikkeli. Viitattu 22.4.2021. https://www.objectrocket.com/resource/what-are-elasticsearch-beats/

What is Amazon API gateway? N.d. Amazon API gatewayn dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/apigateway/latest/developerguide/welcome.html.

What is Amazon S3? N.d. Amazon S3 dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/AmazonS3/latest/userguide/Welcome.html.

What is AWS. N.d. AWS esittely Amazonin verkkosivuilla. Viitattu 23.4.2021. https://aws.amazon.com/what-is-aws/.

What is AWS CloudTrail? N.d. AWS CloudTrail dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/awscloudtrail/latest/userguide/cloudtrail-user-guide.html.

What is DNS? N.d. Cloudflaren DNS protokollan selitys sivu. Viitattu 23.4.2021. https://www.cloudflare.com/en-gb/learning/dns/what-is-dns/.

What is Elasticsearch? N.d. Elasticin elasticsearch ohjeistus sivu. Viitattu 22.4.2021. https://www.elastic.co/guide/en/elasticsearch/reference/current/elasticsearch-intro.html.

What is IAM? N.d. Amazon IAM dokumentointi sivu. Viitattu 23.4.2021. https://docs.amazonaws.cn/en_us/IAM/latest/UserGuide/introduction.html.

What is Kibana? N.d. Elasticin tuote kuvaus sivu Kibanasta. Viitattu 22.4.2021. https://www.elastic.co/what-is/kibana.

What is Monero (XMR)? N.d. Moneron esittely sivu. Viitattu 23.4.2021. https://www.getmonero.org/get-started/what-is-monero/.

What Is Security Information and Event Management (SIEM)? N.d. McAfeen McAfee SIEM kuvaus sivu. Viitattu 22.4.2021. https://www.mcafee.com/enterprise/en-us/security-awareness/operations/what-is-siem.html.

WHAT IS A SECURITY OPERATIONS CENTER? WHY IS IT IMPORTANT? 2020. EC-Councilin blogi artikkeli. Viitattu 4.5.2021. https://blog.eccouncil.org/what-is-a-security-operations-center-why-is-it-important/.

What is SIEM? N.d. IBM aihe kirjoitus sivu SIEMistä. Viitattu 22.4.2021. https://www.ibm.com/topics/siem.

What is SOC Training? 2020. Welp Magazine artikkeli. Viitattu 4.5.2021. https://welpmagazine.com/what-is-soc-training/.

Why Splunk N.d. Myyntipuhe Spunkin verkkosivulla. Viitattu 22.4.2021. https://www.splunk.com/en_us/about-us/why-splunk.html.

Your AWS account ID and its alias. N.d. AWS account ID ja alias dokumentointi sivu. Viitattu 23.4.2021. https://docs.aws.amazon.com/IAM/latest/UserGuide/console_account-alias.html.

Liitteet

Liite 1. Boss Of The SOC versio 3 200-sarjan kysymykset

Kysymys 200

List out the IAM users that accessed an AWS service (successfully or unsuccessfully) in Frothly's AWS environment? Answer guidance: Comma separated without spaces, in alphabetical order. (Example: ajackson,mjones,tmiller)

Kysymys 201

What field would you use to alert that AWS API activity have occurred without MFA (multi-factor authentication)? Answer guidance: Provide the full JSON path. (Example: iceCream.flavors.traditional)

Kysymys 202

What is the processor number used on the web servers? Answer guidance: Include any special characters/punctuation. (Example: The processor number for Intel Core i7-8650U is i7-8650U.)

Kysymys 203

Bud accidentally makes an S3 bucket publicly accessible. What is the event ID of the API call that enabled public access? Answer guidance: Include any special characters/punctuation.

Kysymys 204

What is the name of the S3 bucket that was made publicly accessible?

Kysymys 205

What is the name of the text file that was successfully uploaded into the S3 bucket while it was publicly accessible? Answer guidance: Provide just the file name and extension, not the full path. (Example: filename.docx instead of /mylogs/web/filename.docx)

Kysymys 206

What is the size (in megabytes) of the .tar.gz file that was successfully uploaded into the S3 bucket while it was publicly accessible? Answer guidance: Round to two decimal places without the unit of measure. Use 1024 for the byte conversion. Use a period (not a comma) as the radix characte

Kysymys 207

Ei olemassa

Kysymys 208

A Frothly endpoint exhibits signs of coin mining activity. What is the name of the first process to reach 100 percent CPU processor utilization time from this activity on this endpoint? Answer guidance: Include any special characters/punctuation.

Kysymys 209

When a Frothly web server EC2 instance is launched via auto scaling, it performs automated configuration tasks after the instance starts. How many packages and dependent packages are installed by the cloud initialization script? Answer guidance: Provide the number of installed packages then number of dependent packages, comma separated without spaces.

Kysymys 210

What is the short hostname of the only Frothly endpoint to actually mine Monero cryptocurrency? (Example: ahamilton instead of ahamilton.mycompany.com)

Kysymys 211

How many cryptocurrency mining destinations are visited by Frothly endpoints?

Kysymys 212

Using Splunk's event order functions, what is the first seen signature ID of the coin miner threat according to Frothly's Symantec Endpoint Protection (SEP) data?

Kysymys 213

According to Symantec's website, what is the severity of this specific coin miner threat?

Kysymys 214

What is the short hostname of the only Frothly endpoint to show evidence of defeating the cryptocurrency threat? (Example: ahamilton instead of ahamilton.mycompany.com)

Kysymys 215

What is the FQDN of the endpoint that is running a different Windows operating system edition than the others?

Kysymys 216

According to the Cisco NVM flow logs, for how many seconds does the endpoint generate Monero cryptocurrency? Answer guidance: Round to the nearest second without the unit of measure.

Kysymys 217

What kind of Splunk visualization was in the first file attachment that Bud emails to Frothly employees to illustrate the coin miner issue? Answer guidance: Two words. (Example: choropleth map)

Kysymys 218

What IAM user access key generates the most distinct errors when attempting to access IAM resources?

Kysymys 219

Bud accidentally commits AWS access keys to an external code repository. Shortly after, he receives a notification from AWS that the account had been compromised. What is the support case ID that Amazon opens on his behalf?

Kysymys 220

AWS access keys consist of two parts: an access key ID (e.g., AKIAIOSFODNN7EXAMPLE) and a secret access key (e.g., wJalrXUtnFEMI/K7MDENG/bPxRfiCYEXAMPLEKEY). What is the secret access key of the key that was leaked to the external code repository?

Kysymys 221

Using the leaked key, the adversary makes an unauthorized attempt to create a key for a specific resource. What is the name of that resource? Answer guidance: One word.

Kysymys 222

Using the leaked key, the adversary makes an unauthorized attempt to describe an account. What is the full user agent string of the application that originated the request?

Kysymys 223

The adversary attempts to launch an Ubuntu cloud image as the compromised IAM user. What is the codename for that operating system version in the first attempt? Answer guidance: Two words.

Kysymys 224

Frothly uses Amazon Route 53 for their DNS web service. What is the average length of the distinct third-level subdomains in the queries to brewertalk.com? Answer guidance: Round to two decimal places. (Example: The third-level subdomain for my.example.company.com is example.)

Kysymys 225

Using the payload data found in the memcached attack, what is the name of the .jpeg file that is used by Taedonggang to deface other brewery websites? Answer guidance: Include the file extension.

Liite 2. Korjattu sisääntuonti

Opinnäytetyön teon jälkeen, kun jäi ylimääräistä aikaa, päätettiin korjata sisääntuonti. Ongelman ratkaistiin pääasiassa käyttämällä eri järjestelmää. Tämä tehtiin omalla koneella olevaan Oracle VM VirtualBoxiin. Tähän asennetiin normaali Elastic Stack asennus, jossa kaikki ohjelmat olivat samalla koneella ja eikä lainkaan SSL salausta.

Sisääntuonti tehtiin komennolla, johon lisättiin asetus laittamaan oikeus tarkistukset pois päältä.

```
sudo /usr/share/filebeat/bin/filebeat -e --strict.perms=false
```

Jos **--strict.perms=false** käytetään tulee tapahtumiin lisää kenttiä, täten botsv3part2b menee yli 1000 uniikki kentän. Täten indeksi kohtaan tulee laittaa arvoksi **"botsv3part2b-%{+SS}"** tai jokin muu indeksiä jakava arvo. Normaali tapauksen Filebeats konfiguraatio alla.

```
filebeat.inputs:
- type: log
```

```
  paths:
    - /home/elk/Desktop/botsv3/full/botsv3part1.json
  json.keys_under_root: true
  json.message.key: "message"
  json.override_keys: true
  index: "botsv3part1"

output.elasticsearch:
  hosts: ["localhost:9200"]

processors:
  - drop_fields:
      fields: ["result.timestamp","result.endtime","result.backupUsage{}.lastCompletedBackupDate","result.res-
ponseElements","result.requestParameters","result.lastLoginDate","result.instance_profile","result.location"]
ignore_missing: false
```

Jos yrittää tehdä sisääntuonnin yhdellä ajolla esimerkiksi asetuksilla

```
paths:
  - /home/elk/Desktop/botsv3/full/botsv3part*
  index: "botsv3-%{+SS}"
```

Menee sisääntuonti Elastic Stackin tuhat Shardia rajoituksen yli, ja lopettaa silloin sisääntuonnin.

Liite 3. Testihenkilön täyttämä lomake

Tekijän kuvaus

JAMK tieto- ja viestintätekniikka 4. vuosi.

Tekijän aikaisemmat tähän liittyvät kokemukset

SIEM:Ei

Elastic Stack:Ei

BOTS:Ei

Kyberharjoitukset: JYVSECTECin harjoittelussa testaillut useampia kyberharjoittelu sivustoja

Muut mahdolliset aikaisemmat kokemukset:x

Kysymys 200

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:ei

Kysymys 201

Vastaus saatu (kyllä/ei): kyllä

Vastaus oikein (kyllä/ei): kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei): kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista): kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono): kohtalainen

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei): jotenkin

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei): jotenkin

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista): kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono): hyvä

Muita huomioita:

Kysymys 202

Vastaus saatu (kyllä/ei): kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei): kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista): vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono): kohtalainen

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei): kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei): kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista): vähän

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono): kohtalainen

Muita huomioita:

Kysymys 203

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):ei

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):ei

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono):huono

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):jotenkin

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):jotenkin

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):kohtalainen

Muita huomioita:

Kysymys 204

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:Vastaus löytyi JSONista

Kysymys 205

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 206

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):ei mahdollista

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):hyvä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 207

Ei olemassa

Kysymys 208

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 209

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):ei mahdollista

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 210

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä
Muita huomioita:

Kysymys 211

Vastaus saatu (kyllä/ei):kyllä
Vastaus oikein (kyllä/ei):kyllä
Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä
Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän
Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä
Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä
Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä
Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä
Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä
Muita huomioita:

Kysymys 212

Vastaus saatu (kyllä/ei):kyllä
Vastaus oikein (kyllä/ei):kyllä
Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):ei
Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä
Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä
Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä
Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä
Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä
Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):kohtalainen
Muita huomioita:Syntaksi ongelma haussa

Kysymys 213

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:nice prank bro

Kysymys 214

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 215

Vastaus saatu (kyllä/ei):ei

Vastaus oikein (kyllä/ei):ei

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):ei

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita: Ei voitu tehdä loppuun, import ongelma

Kysymys 216

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 217

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 218

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):vähän

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 219

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 220

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:raw datassa löyty githubin osoite, mutta contentissa ei ollut sähköpostin sisältöä

Kysymys 221

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 222

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 223

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Kysymys 224

Ei voi tehdä

Kysymys 225

Vastaus saatu (kyllä/ei):kyllä

Vastaus oikein (kyllä/ei):kyllä

Vastaus oikein ensimmäisellä yrityksellä (kyllä/ei):kyllä

Ohjeistus antoi mahdollisuuden selvittää asioita itse (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus selkeys (hyvä/kohtalainen/huono):hyvä

Ohjeistuksen ratkaisu logiikka oli selvä (kyllä/jotenkin/ei):kyllä

Ohjeistus oli johdonmukainen (kyllä/jotenkin/ei):kyllä

Ohjeistus opetti asiasta (kyllä/vähän/ei/ei mahdollista):kyllä

Ohjeistus laatu yleisesti (hyvä/kohtalainen/huono):hyvä

Muita huomioita:

Loppu kommentit

Mielipide kokonaisuudessa ohjeistuksesta: Hyvä ohjeistus, varmasti oppii käyttämään Elastic Stackin analysointia.

Muuta:Pieniä teknisiä ongelmia muutamassa kohdassa. Harjoituksessa pidetään käestä kiinni hyvä määrä mutta pääsee tekijä myös itse miettimään.

Liite 4. BOTSv3 Elastic Stack (salassa pidettävä)