SECURITY INFORMATION & EVENT MANAGEMENT (SIEM)

KRISTOF VERSLYPE

Abstract. De IT-gerelateerde veiligheidsrisico's voor bedrijven en overheidsinstellingen kennen een stijgende trend. Er bestaan verschillende – vaak complementaire – methoden om hiermee om te gaan. Dit rapport focust op SIEM (Security Information and Event Management), waarbij veiligheidsgerelateerde logs naar een centraal systeem gestuurd worden voor analyse om security threats te detecteren. Een belangrijk aspect hierbij is de correlatie, waarbij verbanden tussen logs gezocht worden.

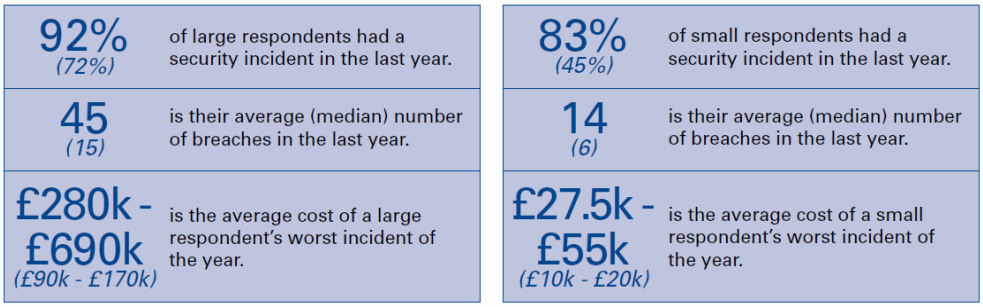
Résumé. Les entreprises et les institutions publiques sont de plus en plus souvent confrontées à des risques de sécurité informatique. Il existe plusieurs méthodes - souvent complémentaires - pour y faire face. Ce rapport approfondit le SIEM (Security Information and Event Management), une méthode qui consiste à envoyer les logs de sécurité à un système central afin de détecter des menaces à la sécurité. Un aspect important dans ce cadre est la corrélation, qui permet de chercher des liens entre les logs.

Inhoudsopgave

1. IT Security Trends

IT-infrastructuren van zowel bedrijven als overheden worden complexer, maar ook dynamischer. De tijd dat er één centrale mainframe was, waarbij een fysieke deur de enige toegang ertoe was, ligt al een tijdje achter ons. Vandaag de dag zijn er in een bedrijf al snel honderden genetwerkte servers en eindgebruikerssystemen met daarop complexe besturingssystemen en andere software. Virtuele servers kunnen on demand gecreëerd worden en bedrijven en overheidsinstellingen worden in toenemende mate geconfronteerd met de Bring Your Own Device (BYOD)[1] realiteit, waarbij de controle over de end-point systemen vervaagt. Dit kan gaan van draagbare computers, tabletcomputers en smartphones tot USB-sticks.

Figuur 1 geeft enkele onthutsende cijfers wat betreft cyberaanvallen op bedrijven in het Verenigd Koninkrijk in 2010 (tussen haakjes ter vergelijking de situatie in 2008). Deze cijfers zijn afkomstig van de tweejaarlijkse Information Security Breaches Survey[2] van PriceWaterhouseCoopers. Volgens datzelfde rapport was er bij 62 % van deze inbreuken sprake van malware[3], bij 15 % van inbraak, bij 25 % van een Denial-of-Service[4] en bij 46 % van verlies of diefstal van gevoelige data door personeel. Het lijkt aannemelijk dat we ervan uit mogen gaan dat voor België gelijkaardige cijfers van toepassing zijn.



Figuur 1: Enkele cijfers betreffende cyberaanvallen op bedrijven in het Verenigd Koninkrijk in 2010 (2008). Grote bedrijven hebben ten minste 250 werknemers.
© PriceWaterhouseCoopers

[1] BYOD: trend waarbij werknemers hun eigen toestellen, zoals draagbare computers en smartphones, voor professionele doeleinden gebruiken op de werkplek.

[2] Information Security Breaches Survey 2010. PriceWaterhouseCoopers. http://www.infosec.co.uk/files/isbs_2010_technical_report_single_pages.pdf

[3] Malware is een verzamelnaam voor kwaadaardige en/of schadelijke software.

[4] Een Denial-of-Service (DoS) is een situatie waarin een computersysteem niet meer in staat is te functioneren. Een denial-of-service-aanval is een poging om een computer, computernetwerk of dienst onbruikbaar te maken voor de bedoelde gebruiker. (Bron: Wikipedia)

Bovenstaande cijfers worden geïllustreerd door diverse aanvallen die de voorbije periode uitgebreid de pers haalden. We denken bijvoorbeeld aan de diverse DDOS[5]-aanvallen en diefstal van gegevens door het hackerscollectief Anonymous[6], aan geavanceerde aanvallen zoals operatie Aurora die allicht georchestreerd worden door buitenlandse overheden en aan malware zoals recentelijk Flame[7].

In de Verenigde Staten zijn er de afgelopen jaren een aantal regelgevingen wettelijk opgelegd om zo het aantal IT-veiligheidsincidenten te verminderen. We denken onder meer aan de Sarbanes-Oxley Act[8] (SOX, 2002), de Healthcare Insurance Portability and Accountability Act[9] (HIPAA, 1996) en de Federal Information Security Management Act[10] (FISMA, 2002). Daarnaast zijn er regelgevingen die vanuit de sectoren zelf komen. Zo wordt er vanuit de financiële sector verlangd dat organisaties die financiële transacties verwerken voldoen aan de Payment Card Industry Data Security Standard[11] (PCI DSS, 2004). In Europa zijn er nog geen door overheden opgelegde regelgevingen, maar het is niet onwaarschijnlijk dat deze in de toekomst ingang zullen vinden. Het toepassen van deze regelgevingen, alsook het nagaan of een organisatie deze naleeft, vergt een aanzienlijke inspanning.

Samengevat zijn er twee trends die de beveiliging van IT-infrastructuren bemoeilijken: 1) de IT-infrastructuren worden complexer en dynamischer en 2) de aanvallen worden gesofistikeerder en de aanvallers professioneler. De evidente vraag die zich dan stelt is hoe er met deze realiteit op een adequate manier omgegaan kan worden. Dit voedt de trend naar meer reguleringen, wat dan weer extra kosten met zich meebrengt.

2. Governance Security

De mogelijke maatregelen zijn in vier categorieën onder te verdelen: informeren, detecteren, verhinderen en reageren.

- Onder *informeren* verstaan we het proces van ontrading en bewustmaking van de betrokken partijen. Een voorbeeld van ontrading is het publiceren van de top 5 van de meest bezochte,

[5] Een distributed-denial-of-service (DDoS) aanval is een DoS-aanval waarbij meerdere computers tegelijk de aanval uitvoeren.

[6] Anonymous is een wereldwijd verspreid ideologisch geïnspireerd hackerscollectief.

[7] Flame is malware die gebruikt werd voor doelgerichte cyberspionage in het Midden-Oosten. Het bestaan van Flame werd in 2012 ontdekt.

[8] Zie: The Sarbanes-Oxley Act 2002, http://www.soxlaw.com/

[9] Zie: Health Insurance Portability and Accountability Act, http://www.hipaa.org/

[10] Zie: Full text of FISMA, http://csrc.nist.gov/drivers/documents/FISMA-final.pdf

[11] Zie: Security Standards Council > PCI Standards & Documents > Documents Library, https://www.pcisecuritystandards.org/security_standards/documents.php

niet-werkgerelateerde websites door de werknemers aan die werknemers. Een voorbeeld van bewustmaking is het opleiden van de werknemers, waarbij ze leren wat de bedrijfspolicy is, hoe ze verdachte e-mails kunnen herkennen, etc.

- *Detecteren* kan zowel fysiek, onder de vorm van bijvoorbeeld camera's, als virtueel, onder de vorm van logging. In beide gevallen is er sprake van monitoring.
- *Verhinderen* is het onmogelijk maken voor onbevoegden om bepaalde acties te ondernemen. In de fysieke wereld denken we aan stevige sloten om mensen uit datacenters te houden. In de virtuele wereld denken we onder meer aan paswoorden en firewalls.
- *Reageren* is het ingrijpen achteraf, dus nadat een incident of een poging daartoe plaatsgevonden heeft. Hierbij onderscheiden we technische en maatschappelijke maatregelen. Als technische maatregel kan een zwakheid (vulnerability) in het systeem gedicht worden en als maatschappelijke maatregel kan een hacker berecht worden of een werknemer ontslagen.

Idealiter wordt alle misbruik gewoon verhinderd, maar de realiteit leert ons dat dit praktisch niet haalbaar is. Vandaar dat er ook inspanningen geleverd worden op de andere niveaus. SIEM situeert zich hierbij in hoofdzaak op het niveau van de detectie. Het zal uiteraard ook zijn impact hebben op het informeren als gekend is dat een SIEM-systeem in gebruik is. Een SIEM-systeem situeert zich in mindere mate in het reactiegedeelte: het kan bewijsmateriaal leveren voor forensisch onderzoek en kan in sommige gevallen automatisch scripts uitvoeren wanneer incidenten gedetecteerd worden.

3. Uitdaging

In een moderne IT-infrastructuur van een bedrijf worden vandaag de dag al snel miljoenen veiligheidsgerelateerde log entries door honderden of meer toestellen gegenereerd, potentieel vanaf verschillende geografisch verspreide locaties. Eén enkele firewall in een bank genereert bijvoorbeeld al snel 20.000 à 30.000 log entries per seconde. Hoewel log entries potentieel heel veel vertellen over de (on)veiligheid van een IT-infrastructuur, zijn er op een dag hoogstens enkele ernstige veiligheidsgerelateerde incidenten.

Nochtans is volgens het Verizons Data Breach Investigations Report 2012[12] bij 84 % van die incidenten bewijs van het incident zichtbaar in de bewaarde logs. Dat dit niet 100 % is heeft te maken met 1) onvoldoende of slecht geconfigureerde systemen, 2) het niet lang genoeg bewaren van

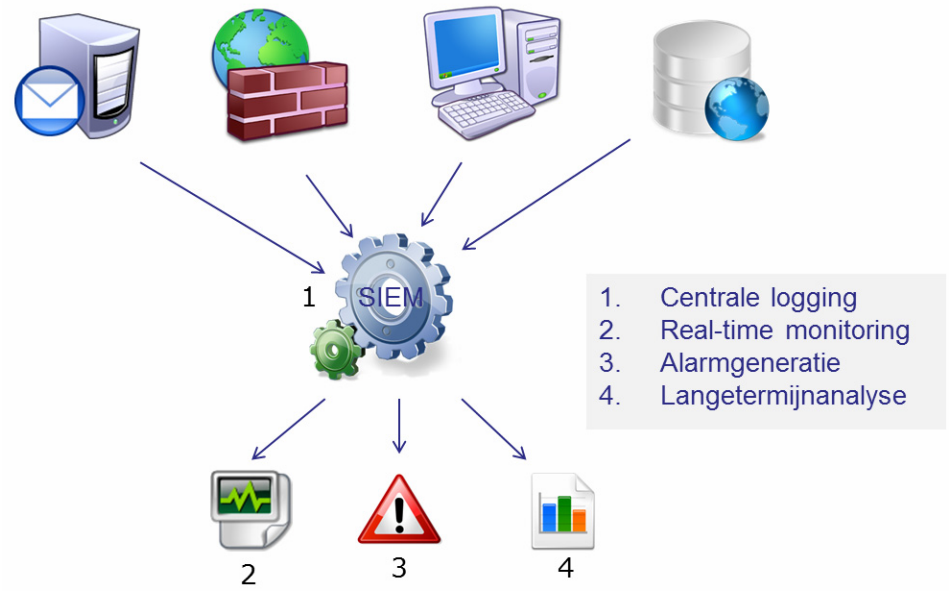
[12] Zie: 2012 Data Breach Investigations Report, Verizon, http://www.verizonbusiness.com/resources/reports/rp_data-breach-investigations-report-2012_en_xg.pdf

logs en 3) antiforensische technieken, waarbij de aanvaller zijn sporen verwijdert.

De uitdaging is om aan de hand van miljoenen log entries die paar ernstige incidenten in real-time te extraheren. Indien nodig kan 1) in real-time ingegrepen worden om de schade te minimaliseren of 2) kunnen achteraf maatregelen genomen worden om dergelijke incidenten in de toekomst zoveel mogelijk te vermijden. Het in real-time extraheren van incidenten uit logs is wat SIEM-systemen op een geautomatiseerde manier beloven te doen, alsook het aanbieden van de mogelijkheid om achteraf uitgebreide analyses op de verzamelde logs uit te voeren als hulp bij de reactieve maatregelen.

4. Basisprincipes

Figuur 2 toont op een hoog niveau de werking van SIEM. Het centrale SIEM-systeem ontvangt de veiligheidsgerelateerde logs vanuit de hele IT-infrastructuur. Dit gaat van firewalls tot Intrusion Detection Systems[13] (IDS), applicatieservers, end-user devices, etc. Zowel het besturingssysteem als de bovenliggende applicaties kunnen in principe logs naar het SIEM-systeem sturen, alsook firmware en middleware.



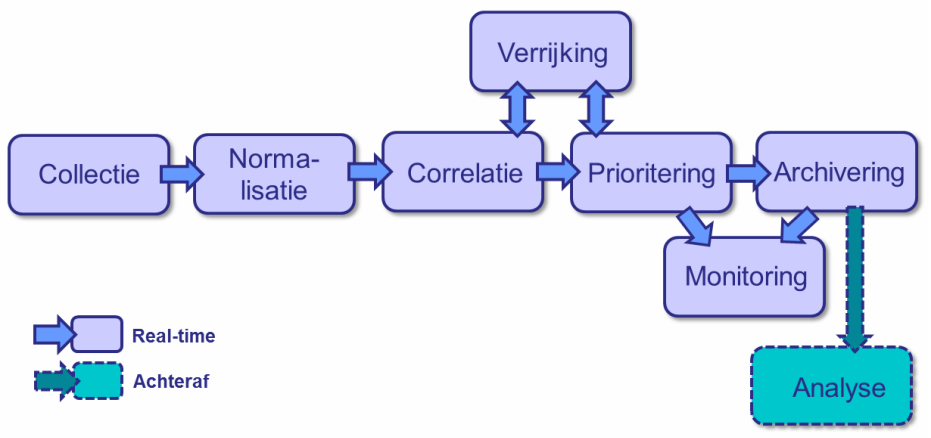
Figuur 2: Werking SIEM

De SIEM-functionaliteit kunnen we opdelen in: 1) centrale logging, 2) het aanbieden van real-time monitoring, 3) alarmgeneratie bij een incident en 4) het aanleveren van gegevens voor langetermijnanalyse.

[13] Een Intrusion Detection System of IDS is een geautomatiseerd systeem dat hackpogingen en voorkomens detecteert van ongeautoriseerde toegang tot een informatiesysteem of netwerk . (Bron: Wikipedia)

5. Verwerking logs

Figuur 3 onderscheidt acht stappen bij de verwerking van log entries. Deze stappen worden nu toegelicht.



Figuur 3: Verwerking logs door SIEM-systeem

5.1. Collectie

Bij de collectie worden de log entries afkomstig van verschillende bronnen door het SIEM-systeem verzameld. Hierbij onderscheiden we het push- en het pull-principe[14].

In de praktijk zien we vaak push over UDP[15], wat efficiënt is maar geen aflevergarantie biedt en bovendien ontbreekt authenticatie van de bron (het systeem en de service die de log entries versturen) meestal. Het gebrek aan aflevergarantie kan er theoretisch voor zorgen dat cruciale logs verloren gaan (bijvoorbeeld op een zwaar belast netwerk) al zou dit in de praktijk maar uiterst zelden voorvallen. Het gebrek aan bronauthenticatie betekent dat om het even welk vijandig of slecht geconfigureerd toestel grote aantallen log entries kan sturen naar het SIEM-systeem dat bijgevolg overbelast kan raken.

Bij pull worden de nieuwe log entries van de bron op regelmatige tijdstippen opgehaald, waardoor het real-time gehalte daalt. Pull blijkt bovendien in de praktijk moeilijker te configureren, aangezien het SIEM-systeem credentials nodig heeft om toegang tot de logs op het brontoestel te krijgen en aangezien autodetectie van nieuwe toestellen en services op

[14] Push is het principe waarbij het sturen van informatie over een netwerk door een informatieproducent naar een informatieconsument geïnitieerd wordt door de producent. Pull is het tegenovergestelde.

[15] UDP (User Datagram Protocol) is één van de basisprotocollen van het internet en opereert op hetzelfde niveau als TCP (Transmission Control Protocol). In tegenstelling tot TCP biedt UDP geen aflevergarantie van de verstuurde pakketen.

het netwerk door het SIEM-systeem moeilijker wordt. In de praktijk wordt push dan ook het vaakst gebruikt.

Niet alle bronnen bieden standaard de ondersteuning aan om log entries naar een externe locatie te sturen. In deze gevallen zal er nood zijn aan een agent op het brontoestel. Deze agent kan bijvoorbeeld lokale logbestanden monitoren op nieuwe log entries en deze vervolgens doorsturen. Het SIEM-systeem moet uiteraard ook ondersteuning bieden voor alle bronnen waar het log entries van ontvangt. Als bijvoorbeeld een Database Activity Monitoring (DAM) systeem log entries stuurt naar een SIEM-systeem, moet deze laatste uiteraard weten hoe het de log entries van het eerste kan interpreteren.

Bemerk ten slotte dat het in real-time doorsturen van de gegenereerde log entries het gebruik van antiforensische technieken wel heel erg moeilijk maakt voor aanvallers. De aanvaller hoeft niet enkel in het gehackte systeem sporen uit te wissen, maar ook in het SIEM-systeem.

5.2. Normalisatie

Na de collectie volgt de normalisatie waarbij de binnengekomen log entries, die verschillende formaten kunnen hebben (Windows Event logs, syslog, Cisco logs, etc.), omgezet worden naar een uniform formaat. Dit vergemakkelijkt de volgende stap, de correlatie.

Deze normalisatie situeert zich potentieel op vier niveaus.

1. De encodering kan verschillen. Er is onder meer keuze tussen ASCII en UTF-8.
2. Er zijn diverse logstructuren. In sommige logtypes worden eenvoudige scheidingstekens zoals puntkomma's gebruikt om de verschillende velden te begrenzen, terwijl andere logtypes XML of nog een andere structuur zullen hanteren.
3. Veldwaarden kunnen, hoewel ze conceptueel dezelfde informatie bevatten, toch verschillen. Zo kunnen "Mon, 19 Mar 2012 13:38:11 GMT" en "1332164291" exact dezelfde tijd aanduiden.
4. Ten slotte kunnen velden ook namen hebben die opnieuw kunnen verschillen. Denken we maar aan "time" en "timestamp".

5.3. Correlatie

Correlatie komt erop neer dat er verbanden tussen log entries gezocht worden, ook al zijn deze op verschillende momenten door verschillende devices op verschillende geografische locaties gegenereerd.

We illustreren dit aan de hand van twee voorbeelden.

- Te midden van talloze andere log entries detecteert een firewall een port scan. Kort daarna detecteert een applicatieserver een hele reeks gefaalde SSH-authenticaties, een succesvolle SSH-authenticatie en ten slotte een installatie van een stukje software

op die applicatieserver zelf. Daaropvolgend detecteert de IDS een reconnaissance scan[16]. Een SIEM-systeem zou als resultaat van een eerste correlatie kunnen detecteren dat er een geslaagde brute force SSH-authenticatie gebeurd is. Bij een tweede correlatie zou deze brute force-aanval geplaatst kunnen worden in een bredere aanval.

- Het correleren kan helpen bij het detecteren van anomalieën. Wanneer een gebruiker ingelogd is op zijn bedrijfscomputer op het bedrijfsnetwerk, en tegelijkertijd een systeem met een extern IP-adres probeert in te loggen onder dezelfde gebruiker, kan dit via correlaties gedetecteerd worden.

Correlatie is een uitdagend aspect bij SIEM-systemen. Het SIEM-systeem staat of valt hier mee. Er zijn immers diverse SIEM-implementaties die niet efficiënt zijn door slecht geconfigureerde correlatieregels. Vandaag de dag is het correleren nog regelgebaseerd, maar we zien een voorzichtige evolutie richting analytics.

5.4. Verrijking

Bij verrijking wordt externe informatie aan de logs gekoppeld. Enkele voorbeelden zijn 1) gekende kwetsbaarheden van een betrokken server, 2) de functie van deze server, 3) de volledige naam en functie van de persoon horend bij de betrokken account en 4) de geografische locatie van het externe IP-adres. Deze informatie haalt het SIEM-systeem uit bronnen zoals een CMDB[17] en Active Directory[18].

We illustreren dit a.d.h.v. een voorbeeld. Onderstaande log entry van een SSH-authenticatie ziet er op het eerste gezicht niet bijzonder uit.

> "Mar 20 **02:44:35 192.168.56.41** sshd[263] Accepted pass for **krive** from **172.16.254.1** port 56946 ssh2"

Bij verrijking zou er extra informatie over de delen in het vet opgehaald kunnen worden. Daarbij zien we het volgende:

1. De gebruiker krive is iemand die sinds vorige maand op de sectie Onderzoek werkt.
2. Het IP-adres 192.168.56.41 komt overeen met een cruciale databaseserver in productie.
3. Het systeem met IP-adres 172.16.254.1 bevindt zich in China.

---

[16] Een reconnaissance scan is een verkenning van het netwerk; welke zijn de actieve systemen, welke besturingssystemen gebruiken ze, welke services draaien ze, etc.

[17] Een configuration management database (CMDB) is een opslagplaats voor gegevens over de componenten in een informatiesysteem, waaronder ook configuratiegegevens.

[18] Active Directory staat beheerders toe om het beleid (rechten en instellingen) in het netwerk van een volledig bedrijf te beheren. Ook het automatisch installeren van software en patches behoort tot de mogelijkheden. (Bron: Wikipedia)

4. 02:44:35 is buiten de werkuren.

We zien dus dat een ogenschijnlijk normale log bij nader inzien uitermate verdacht is: met behulp van het account van iemand die nieuw is in het bedrijf probeert iets of iemand toegang te krijgen tot een productiesysteem, terwijl krive, als onderzoeker, niets met productie te maken heeft. Bovendien is ook het tijdstip wel erg verdacht.

Verrijking is nuttig bij het inschatten van het risico (zie sectie “prioritering”) van een mogelijke aanval. Een poging om misbruik van een Windows-kwetsbaarheid te maken is bijvoorbeeld ongevaarlijk als de aanval gericht is tegen een Linuxmachine, maar potentieel zeer gevaarlijk indien gericht tegen een Windowsmachine die deze kwetsbaarheid heeft.

5.5. Prioritering

Bij prioritering wordt aan de hand van een aantal parameters berekend hoe risicovol een event is, waarbij een event uit één of meerdere log entries bestaat. Hierbij wordt gekeken naar diverse parameters zoals o.a.

1. De waarde (vb. kritieke server of client machine),
2. De impact (vb. root access of slechts heel beperkte privileges),
3. De waarschijnlijkheid dat het event tot een effectief incident leidt (vb. 5 gefaalde SSH-authenticaties of 10.000 gefaalde SSH-authenticaties).

5.6. Archivering

Bij de archivering zijn de voornaamste vereisten: confidentialiteit, integriteit, beschikbaarheid, efficiëntie en juridische bewijswaarde.

1. *Confidentialiteit* vertaalt zich in het geven van toegang tot enkel deze verzamelde log entries die voor de specifieke gebruikers van het SIEM-systeem noodzakelijk zijn. Log entries kunnen immers gevoelige gegevens bevatten. Bijvoorbeeld, een log entry als gevolg van een mislukte authenticatiepoging waarbij de gebruiker per ongeluk zijn paswoord ingaf in het veld waar de gebruikersnaam vereist was, zal het paswoord bevatten.
2. *Integriteit* en *beschikbaarheid* van de bewaarde logs zijn noodzakelijk om correcte analyses te kunnen doen in real-time en achteraf. De beschikbaarheid vertaalt zich in de noodzaak om de verzamelde logs ook voldoende lang te bewaren. Integriteit vertaalt zich in de onmogelijkheid om log entries te wijzigen of te verwijderen of om vervalste log entries toe te voegen. Dit geldt eveneens voor de beheerders van systemen waarvan het SIEM-systeem gebruik maakt (vb. de database waar de log entries bewaard worden) en het SIEM-systeem zelf.
3. Daarnaast is ook de *efficiëntie* cruciaal. Enerzijds worden er dagelijks potentieel miljoenen log entries toegevoegd, en anderzijds moet er efficiënt in deze verzameling log entries gezocht

kunnen worden. Dit vertaalt zich op het niveau van de databanken in gegevensstructuren die geoptimaliseerd zijn voor het wegschrijven van log entries zonder toe te geven op de efficiëntie van het lezen.

4. Genormaliseerde logs hebben typisch geen *juridische bewijswaarde*, wat ook de opslag van de originele logs noodzakelijk maakt. SIEM-systemen zijn dus doorgaans ook geschikt om de originele logs te bewaren waarbij confidentialiteit, integriteit en beschikbaarheid gegarandeerd worden.

SIEM-leveranciers gebruiken vaak geoptimaliseerde appliances om aan de bovenstaande vereisten maximaal te kunnen voldoen.

5.7. Monitoring

Monitoring houdt in dat via dashboards de kwetsbaarheden, de verdachte activiteit en de effectieve incidenten op een visuele en overzichtelijke manier aan gebruikers van het SIEM-systeem gepresenteerd worden. Andere zaken zoals status, uptime en belasting van de verschillende toestellen en services alsook informatie over het netwerkverkeer kunnen eveneens via dashboards gevisualiseerd worden. Veelal wordt gebruik gemaakt van web interfaces.

Het SIEM-systeem kan alarmen genereren wanneer events met een voldoende hoog risico gedetecteerd worden. Alarmen kunnen via o.a. dashboards, e-mail en sms verstuurd worden.

Typisch wordt er gewerkt met tijdsvensters, waarbij de activiteit van bijvoorbeeld de laatste 24 uur getoond wordt op de dashboards. In SIEM-systemen worden via de dashboards events op een eerder hoog niveau gepresenteerd, maar is het wel steeds mogelijk om door te klikken tot op het niveau van de individuele log entries.

Gezien een SIEM-systeem door verschillende mensen met verschillende taken gebruikt kan worden, is een user management systeem geen overbodige luxe. Met behulp hiervan kunnen verschillende gebruikers toegang krijgen tot verschillende dashboards en kunnen ze verschillende (types) queries uitvoeren en verschillende types rapporten genereren (zie volgende sectie).

5.8. Langetermijnanalyse

De langetermijnanalyse op reeds verzamelde logs resulteert in een dieper inzicht in de veiligheid van de eigen IT-infrastructuur, is waardevol bij forensisch onderzoek alsook bij het nagaan en aantonen van het naleven van regelgevingen zoals HIPAA en PCI DSS. Bij forensisch onderzoek kan het SIEM-systeem een overzicht bieden van wat een verdachte persoon in de IT-infrastructuur gedaan heeft in een gegeven periode en kan een selectie van de verzamelde log entries gebruikt worden als bewijsmateriaal.

De analyse gebeurt d.m.v. rapporten en queries. Rapporten geven eerder algemene informatie, terwijl queries toelaten om specifieke informatie te zoeken. Typisch geven SIEM-systemen een hele resem rapporttemplates en querytemplates mee, die naar de specifieke noden van het bedrijf aangepast kunnen worden.

6. Topologie

Er zijn drie topologische modellen te onderscheiden.

1. *Standaard.* De gemonitorde devices sturen hun log entries naar een centrale SIEM-server. Dit model is weinig schaalbaar.
2. *Hiërarchisch.* De gemonitorde devices sturen hun log entries naar intermediaire servers die een deel van de verwerking, zoals de normalisatie, op zich nemen en vervolgens de verwerkte log entries of een selectie ervan doorsturen naar de eigenlijke SIEM-server. Dit model is onder meer praktisch wanneer het bedrijf geografisch verspreid is.
3. *Mesh.* In dit model sturen de verschillende gemonitorde devices hun logs naar een cluster van SIEM-servers, waarbij er onderling een taakverdeling kan zijn. Dit model is complexer dan het hiërarchische model.

In de praktijk zien we in grotere productieomgevingen vaak het hiërarchisch model. Ook hybride situaties zijn mogelijk.

In plaats van de communicatie tussen de SIEM-componenten over het productienetwerk te laten gebeuren, kan dit of een deel hiervan over een apart netwerk om de betrouwbaarheid en veiligheid te verhogen.

7. SIEM-producten

State-of-the-art producten zijn ArcSight (HP), QRadar (IBM) en NitroView (McAfee)[19]. Een gedeeltelijk gratis open source SIEM-product is AlienVault.

De SIEM-markt is in volle ontwikkeling. In november 2010 werd ArcSight door HP overgenomen, in november 2011 werd eerst NitroSecurity door McAfee overgenomen en Q1Labs, die Q1Radar aanbiedt, door IBM.

AlienVault en ArcSight worden nu kort besproken.

7.1. AlienVault

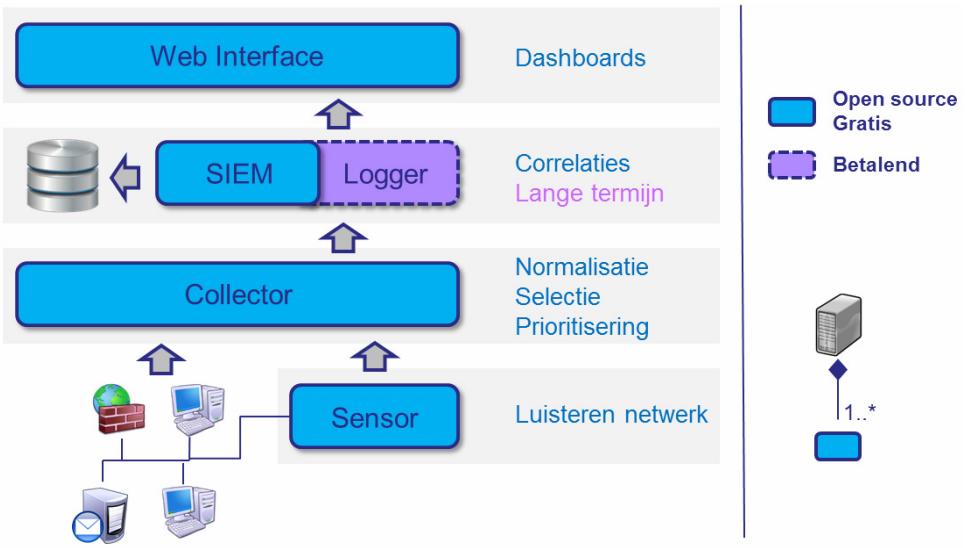
Het Debian-gebaseerde AlienVault is getest door Smals onderzoek in een beperkte virtuele omgeving. AlienVault is erin geslaagd heel wat open

[19] Bron : Magic Quadrant for Security Information and Event Management, Gartner, May 24 2012.

source producten zoals onder meer openVAS, Snort, ntop en Nagios in één SIEM-tool te integreren.

7.1.1. Componenten

Figuur 4 toont de verschillende componenten van een AlienVault-systeem. Een Sensor component monitort een netwerksegment en stuurt, net zoals de relevante devices, zijn log entries door naar de Collector. Meerdere Sensor componenten kunnen logs sturen naar een Collector component. De Collector component bevindt zich een niveau hoger en is verantwoordelijk voor het verzamelen van de logs, het normaliseren ervan, voor prioritering en voor de selectie van de logs die doorgestuurd zullen worden naar de centrale SIEM-component. Meerdere Collector componenten kunnen logs sturen naar een SIEM-component. De SIEM-component is verantwoordelijk voor de real-time analyse. Voor de langetermijnanalyse (rapporten genereren en queries uitvoeren) is de betalende Logger component vereist. Ten slotte is er de Web Interface die de dashboards aan de gebruiker aanbiedt en de databank. De verschillende componenten kunnen elk op een apart systeem geïnstalleerd worden, maar het is eveneens mogelijk om meerdere componenten op één machine te installeren.

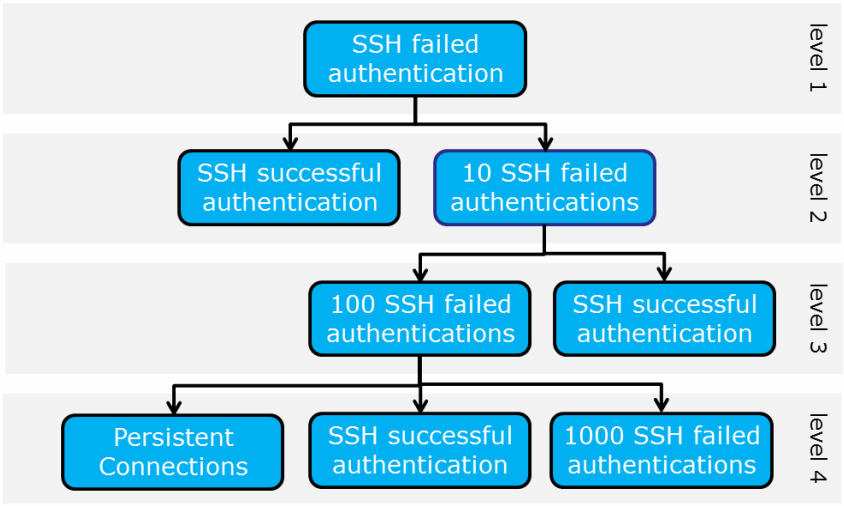


Figuur 4: De verschillende componenten in AlienVault

7.1.2. Correlatieregels

Voor de correlatieregels heeft AlienVault een 90-tal directieven, die elk een set gerelateerde, hiërarchisch geordende regels groeperen. In de SSH brute force attack directive die in figuur 5 getoond wordt, vinden we bijvoorbeeld op het laagste niveau “SSH failed authentication”, dus één enkele poging om via SSH-toegang te krijgen tot een systeem. Nadat

deze regel vervuld wordt door een binnenkomende log entry, zal daarna getracht worden om met de daaropvolgende binnenkomende log entries één van de regels op het hogerliggende niveau te vervullen ("SSH successful authentication" en "10 SSH failed SSH authentications"). Voor elke vervulde regel zal een risico berekend worden, zodat eventueel een alarm gegenereerd kan worden.



Figuur 5. SSH brute force attack directive

Verder zijn er directieven voor het detecteren van onder meer DoS-aanvallen, spoofing[20], reconnaissance scans[21], malware-activiteit, torrentverkeer, etc. De directieven blijven dus eerder laag niveau.

7.1.3. Risicoberekening

AlienVault genereert een alarm wanneer het risico voor een event meer dan 1 is, waarbij het risico als volgt berekend wordt:

$$risico = \frac{assetWaarde * impactBijSucces * WaarschijnlijkheidTotSucces}{25}$$

[20] Spoofing is het vervalsen van kenmerken, met doorgaans als doel het aannemen van een valse identiteit. Voorbeelden zijn e-mail spoofing en website spoofing.

[21] Een reconnaissance scan is een verkenning van het netwerk waarbij getracht wordt informatie te verzamelen over het netwerk, de verschillende systemen en services.

Dit zijn dezelfde parameters die in sectie 5.5 bespoken zijn. *AssetWaarde* en *ImpactBijSucces* liggen beide in het interval [0,5] en *WaarschijnlijkheidTotSucces* in het interval [0,10]. Enkel gehele waarden zijn mogelijk. Voor elke asset bepaalt de beheerder van het SIEM-systeem de *assetWaarde*. De betrokken directief bepaalt de waarde van *impactBijSucces* en de eigenlijke correlatieregel die getriggerd werd, bepaalt de waarde van *WaarschijnlijkheidTotSucces*.

De SSH brute force attack directive heeft een *impactBijSucces*-waarde van 4, wat de op één na hoogste waarde is. Na 10 gefaalde SSH-authenticaties wordt de regel "10 SSH failed authentications" getriggered, die een *WaarschijnlijkheidTotSucces*-waarde heeft van 2. Op een cruciale DNS-server met een assetWaarde van 4 krijgen we dus een risico van 4*2*4/25 = 1,28, wat wil zeggen dat een alarm gegenereerd zal worden nog voor er sprake is van een effectieve penetratie. Zien we hetzelfde gebeuren op een wat minder belangrijke server met een assetWaarde van 3, krijgen we een risico van 3*2*4/25 = 0,96. Er zal dus geen alarm gegenereerd worden. In dit voorbeeld werden de standaardwaarden voor *impacBijSucces* en *waarschijnlijkheidTotSucces* gebruikt.

7.1.4. Ervaring

AlienVault is een waardevolle tool voor het centraal in de gaten houden van de veiligheid binnen de IT-infrastructuur, maar is niet in staat een volwaardig alternatief te bieden voor de SIEM-systemen in het topsegment, die weliswaar vrij duur zijn.

AlienVault ondersteunt momenteel zo'n 2400 bronnen, maar helaas vinden we daar nog geen geavanceerdere technologieën zoals DAM of PAM[22] terug. De ondersteuning uitbreiden naar andere bronnen is mogelijk mits het schrijven van regels om de binnenkomende logs te interpreteren.

AlienVault biedt geen uitgebreide toegangscontrole aan. De gebruiker geeft zijn paswoord en krijgt de standaard dashboards te zien. Dit kan een gebrek zijn in omgevingen waarin meerdere rollen gebruik maken van het SIEM-systeem.

De hoog niveau correlatieregels ontbreken. Ze beperken zich tot zaken als detectie van DoS-aanvallen, spoofing, reconnaissance, SSH-authenticatiepogingen, malware-infectie, malware-activiteit, flooding, het uitvoeren van bepaalde commando's en het detecteren van verdacht verkeer op het netwerk.

AlienVault kan een uitstekende kennismaking met SIEM zijn, maar zal wellicht wat tekortschieten in een complexe bedrijfscontext. Misschien loont het wel de moeite om een bijdrage te leveren aan AlienVault in plaats van het aanschaffen van een duur commercieel systeem. Een open vraag blijft evenwel de efficiëntie in een productienetwerk, wat we helaas niet hebben kunnen testen.

[22] Privileged Account Management

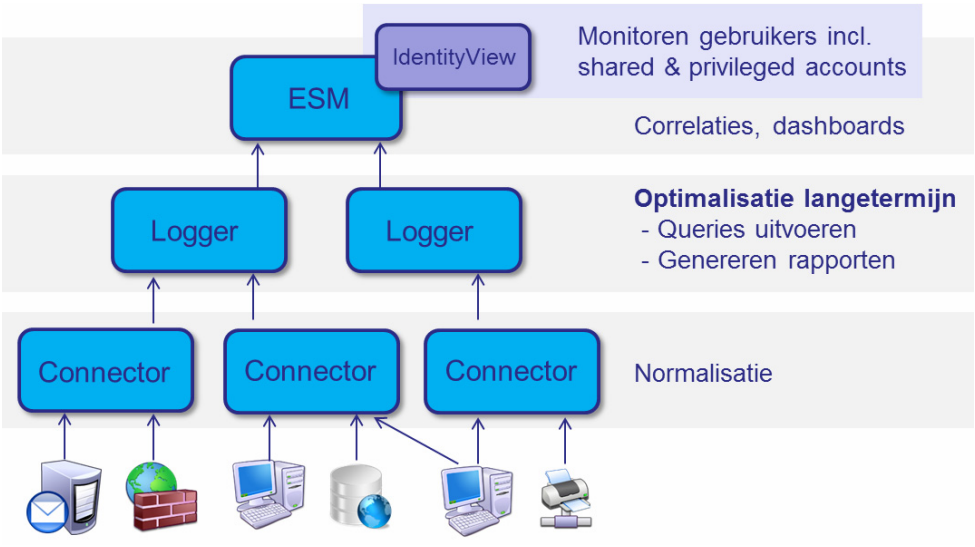
Bemerk ten slotte dat AlienVault meer aanbiedt dan de pure SIEM-functionaliteit. In de AlienVault Sensor component vinden we immers allerlei software die passief op het netwerk luistert. Zo vinden we er Snort, de de facto standaard voor open source IDS.

7.2. ArcSight

ArcSight is de SIEM-oplossing van HP. Het is een krachtige oplossing, maar tevens ook een dure.

7.2.1. Componenten

De verschilende standaardcomponenten van ArcSight worden getoond in figuur 6. De toestellen sturen hun log entries naar de Connector componenten, die instaan voor de normalisatie. De genormaliseerde logs worden naar de Logger componenten gestuurd. Deze optionele Logger componenten zorgen voor optimalisatie van de langetermijnanalyse, het uitvoeren van de queries en het genereren van rapporten. Op het hoogste niveau vinden we de ESM (Enterprise Security Manager), het hart van het ArcSight-systeem, waar de correlaties gebeuren en de dashboards zich bevinden. Naast de standaardfunctionaliteit zijn er uitbreidingen voor o.a. de detectie van low-and-slow-aanvallen, optimalisatie voor analyse, monitoring van gebruikers waaronder ook gedeelde en geprivilegieerde accounts. Als goedkoper alternatief is er de lightweight ArcSight Express, die een gelimiteerde SIEM is, waarbij er slechts van één component sprake is.



Figuur 6: De verschillende componenten in ArcSight

7.2.2. Veelzijdigheid & snelheid

ArcSight kan sterk naar de specifieke noden van het bedrijf aangepast worden. Rapporten kunnen gepersonaliseerd worden. Dashboards

kunnen samengesteld worden waarbij verschillende monitoring tijdvensters van toepassing kunnen zijn (2u, 12u, etc.). Queries kunnen flexibel samengesteld worden en er is een veelzijdig, fijnmazig gebruikersbeheer. Het nemen van backups kan m.b.v. diverse opslagfaciliteiten. Het is eventueel mogelijk om in ArcSight zelfgeschreven scripts aan te roepen, al moet daarmee opgepast worden, gezien false positives ongewenste effecten met alle gevolgen van dien kunnen hebben.

In een productieomgeving duurde het, afhankelijk van de belasting, 2 tot 5 minuten eer een event op het dashboard getoond werd, wat soms wel een cruciale tijdspanne kan zijn. Het uitvoeren van een query kan een verder negatief effect op het real-time gehalte hebben. Zo duurt het uitvoeren van een query in de geobserveerde productieomgeving op de log entries van de laatste twee weken, waarbij er een stringvergelijking nodig is, 5 tot 10 minuten wanneer er geen gebruik gemaakt wordt van de logger componenten. En ondertussen zal het monitoren aanzienlijk langzamer verlopen.

7.2.3. Nadelen

Hoewel ArcSight een zeer krachtig systeem is met beperkte overhead op het netwerk, zijn er toch een aantal nadelen. Deze elementen zijn hoofdzakelijk gebaseerd op feedback verkregen van mensen die dagelijks ArcSight gebruiken.

1. ArcSight is vrij duur. We spreken hierbij al vrij snel over honderdduizenden euro's.
2. Voor niet-klanten wordt er door HP vrij weinig informatie beschikbaar gesteld.
3. Typisch heeft ArcSight bij de ondersteuning van bronnen wat marktachterstand. Zo werd de courante Cisco IDS v6 pas ondersteund op het moment dat de Cisco IDS v7 op de markt kwam.
4. Hoewel ArcSight in theorie zeer goed naar de noden van het bedrijf verfijnd kan worden, blijkt het beheer in de praktijk soms wat omslachtig en voor verbetering vatbaar.
5. Het updaten van ArcSight wanneer niet van een appliance gebruik gemaakt wordt is doorgaans niet evident. De onderliggende databank moet potentieel eveneens geüpdatet worden en een deel van de configuratie gaat soms verloren.
6. Wanneer een backup gemaakt wordt, kunnen de gebackupte logs enkel in diezelfde ESM terug geïmporteerd worden.

7.2.4. Kosten

Bij de aankoop moet met verschillende aspecten rekening gehouden worden.

1. De prijzen voor de componenten variëren naargelang er gebruik gemaakt wordt van eigen hardware, een appliance of een virtuele appliance.
2. Bij appliances zijn veelal extra licenties nodig om de aanwezige maar ongebruikte hardware, zoals typisch processor cores, te activeren.
3. Extra functionaliteit zoals ondersteuning voor de detectie van low-and-slow-aanvallen, ondersteuning voor de langetermijnaspecten en ondersteuning voor het monitoren van gebruikers, kan aangekocht worden.
4. Voor elk gemonitord toestel is een licentie nodig.
5. De ondersteuning vanuit HP is betalend.

8. Managed SIEM

De klassieke SIEM-aanpak, waarbij elk bedrijf via een eigen SIEM-systeem zijn eigen IT-infrastructuur monitort en zelf de escalatie doet bij een incident, resulteert potentieel in een aantal nadelen zoals investeringskosten, onderhoud, een gebrek aan expertise en te beperkte monitoring.

Managed SIEM is een alternatief waarbij bedrijven hun log entries naar een externe partij, de MSSP (Managed Security Services Provider), sturen, die ook het monitoren en de escalatie op zich neemt. Een MSSP heeft meer SIEM-expertise en een grotere security intelligence. Afhankelijk van de bedrijfsfilosofie kan het outsourcen van de verwerking van potentieel gevoelige log entries wenselijk of minder wenselijk zijn. Investeringskosten worden vervangen door beheerskosten. De prijs hangt af van het overeengekomen contract, maar typisch ligt dit tussen € 10.000 en € 20.000 per jaar per gemonitord device.

MSSP's kunnen ook op andere manieren ondersteuning bieden in het SIEM-verhaal, door bijvoorbeeld een SIEM in het bedrijf zelf te monitoren of door mankracht en opleiding te leveren bij bijvoorbeeld de uitrol van het SIEM-systeem.

IBM, Symantec, BT, HP, Verizon Business en Fujitsu zijn de grotere MSSP's. Een Belgische speler is Belgacom.

9. Aandachtspunten

De voornaamste aandachtspunten bij de uitrol van een SIEM-systeem verdelen we onder in "business", "infrastructuur" en "uitrol en gebruik". Bemerk dat dit geen exhaustieve lijst is.

Wat betreft het **business aspect** zijn de volgende aandachtspunten relevant.

- Een degelijk security incidence response plan op voorhand is noodzakelijk; welke procedures worden er gevolgd bij een bepaald type incident?
- De skills en mankracht moeten aanwezig zijn. Hierbij is opleiding en ondersteuning door de SIEM-leverancier wenselijk. Niet alleen in de testfase, maar ook eens het systeem in productie is. Hou er ook rekening mee dat de juiste mensen hebben één ding is, maar dat ze ook voldoende uitdaging moeten hebben opdat het bedrijf ze kan houden.
- De business requirements moeten op voorhand duidelijk geformuleerd zijn. Leg op deze prioritaire requirements de focus. Het wordt sterk afgeraden om “alles” trachten te doen. De focus zou bijvoorbeeld gelegd kunnen worden op insider attacks, het nagaan van het naleven van bepaalde regelgevingen of het monitoren van de systeembeheerders.
- Draagvlak en betrokkenheid vanuit de verschillende bedrijfsafdelingen die met het SIEM-systeem zullen te maken krijgen is essentieel. SIEM is immers veelal een tool die niet door één enkele IT-specialist gebruikt wordt, maar ook door bijvoorbeeld interne audit, de databaseploeg, de netwerkploeg, de securityploeg tot zelfs HR. Omwille van dit breedingrijpende karakter is het absoluut noodzakelijk dat ook de directie volledig achter het project staat.

Samengevat kunnen we stellen dat mensen en processen even belangrijk zijn als techologie.

Op **infrastructureel** vlak vinden we volgende aandachtspunten.

- Er is een duidelijk beeld van de IT-infrastructuur nodig. Waar staan welke devices? Welke applicaties draaien er op? Hoe cruciaal zijn de devices en de services? Waar staan de gevoelige gegevens of worden de gevoelige gegevens verwerkt? Welke is de netwerktopologie en de beschikbare bandbreedte? Hoeveel log entries worden er door de verschillende bronnen gegenereerd? Etc.
- Er mogen geen zwarte gaten zijn bij het (nu nog gedecentraliseerd) monitoren. Zijn de benodigde security devices op de verschillende subnetten aanwezig? Wordt overal de noodzakelijke logging (lokaal) gedaan? Voordat er een SIEM-systeem geïnstalleerd wordt, is een degelijke veiligheid nodig opdat het SIEM-systeem enige toegevoegde waarde kan hebben.
- De klokken van de verschillende devices moeten gesynchroniseerd zijn en blijven. Zoniet kan het correleren een onmogelijke opdracht worden.
- Er is een goed beeld nodig van de opslagvereisten van de logs. 6 maanden opslag wordt als een goed begin gezien. Vaak zien we een onderschatting van de hoeveelheid gegenereerde log entries. Er is dus een degelijke inschatting per toestel nodig.

Ten slotte zijn er de aandachtspunten bij de **uitrol en het gebruik** van het SIEM-systeem.

- SIEM is geen plug-and-play-tool en fine-tuning zal nodig zijn. De standaard meegeleverde correlatieregels zullen aangepast moeten worden naar de noden van het bedijf. Dit blijkt vaak een groot struikelblok te zijn wegens een onderschatting van dit werk.
- Documentatie van de SIEM-uitrol kan tijdsverlies later voorkomen. Van welke veronderstellingen werd er vertrokken bij de uitrol van het SIEM-systeem? Waarvoor wordt het SIEM-systeem exact gebruikt? Welke topologie werd er gehanteerd: wat zijn de verschillende SIEM-devices, welke taak hebben ze en hoe interageren ze? Hoe zijn de ruimtes waar de SIEM-systemen staan beveiligd? Etc.
- Zorg voor een gefaseerde uitrol. Pas wanneer een (eventueel vereenvoudigde) use case een aanvaardbaar aantal hoge prioriteitsevents genereert, kan een stap verder gegaan worden. Een volledige uitrol vergt tijd. Als vuistregel wordt veelal 12 tot 18 maanden naar voor geschoven en soms zelfs drie jaar.
- Ten slotte veranderen de bedreigingen alsook de IT-infrastructuur constant waardoor onderhoud van het SIEM-systeem noodzakelijk is. Nieuwe patches en ondersteuning voor nieuwe logtypes gegenereerd door nieuwe bronnen zal nodig zijn, correlatieregels zullen moeten meeëvolueren, etc.

10. Te onthouden

In sectie 2 werd een onderscheid gemaakt tussen preventie, detectie, verhindering en reactie. In een ideale wereld speelt de veiligheid zich enkel af op het niveau van de verhindering en worden de andere niveaus bijgevolg overbodig. Daar staan we nog ver van af, maar dit moet wel het ultieme objectief blijven in veiligheid van IT-infrastructuren. De werkelijkheid noodzaakt ons om ook op de andere niveaus zoals detectie te investeren, en daar zien we dan SIEM in het vizier komen om dit op een efficiënte manier beheersbaar te houden. Dit gaat gepaard met een aanzienlijke investering en inspanning.

Zonder goede veiligheid op voorhand heeft het gebruik van een SIEM-systeem weinig zin.

SIEM is een project. Zonder de nodige technische voorbereiding en de nodige mensen heeft het project geen kans op slagen. Een SIEM-uitrol vergt tijd. 12 tot 18 maanden is een vuistregel, maar dit varieert uiteraard naargelang de specifieke condities. Ten slotte moet een SIEM-systeem in productie onderhouden worden.

Het heikele punt bij veel SIEM-systemen vandaag de dag zijn de correlatieregels. De afstemming ervan zal het verschil maken tussen enkele hoge prioriteitsevents of enkele honderdduizenden en bepaalt dus de mogelijke toegevoegde waarde van het SIEM-systeem.

SIEM blijft zowel in aankoop als in onderhoud een aanzienlijke kost. De aanschaf beloopt al snel enkele honderdduizenden euro's en de nodige mankracht en expertise in huis halen en houden is vaak geen evidentie. Managed SIEM kan een alternatief bieden, maar ook hier spreken we over € 10.000 à € 20.000 per managed device.

AlienVault is geschikt als eerste kennismaking met SIEM. Verschillende leveranciers zijn bereid om een appliance uit te lenen voor testdoeleinden en vaak hebben ze in hun gebouwen een testomgeving waar demo's gegeven kunnen worden.

Er zijn omwille van de bovengenoemde redenen vele SIEM-uitrollen mislukt, wat niet wegneemt dat een succesvolle uitrol, die een effectieve meerwaarde biedt, mogelijk is. We zien dit bevestigd in onder meer de financiële sector, de farmasector, de energiesector en defensie.

In de VS is 80 % van de SIEM-uitrollen gedreven vanuit de vereiste om in overeenstemming te zijn met geldende regelgevingen. We kunnen verwachten dat dergelijke reguleringen ook Europa zullen bereiken. Dit zal de interesse voor SIEM doen toenemen.

Verder kunnen we verwachten dat het gebruik van analytics in het SIEM-verhaal zal toenemen. Vandaag de dag is dit meestal uitsluitend regelgebaseerd.

Aanbevolen literatuur

- Verizon, 2012 Data Breach Investigations Report, 2012. http://www.verizonbusiness.com/resources/reports/rp_data-breach-investigations-report-2012_en_xg.pdf
- Miller, Harris, Harper, VanDyke, Blask. Security Information and Event Management (SIEM) Implementation. 10/2010
- ISACA. Security Information and Event Management: Business Benefits and Security, Governance and Assurance Perspectives. 12/2010. http://www.isaca.org/Knowledge-Center/Research/ResearchDeliverables/Pages/Security-Information-and-Event-Management-Business-Benefits-and-Security-Governance-and-Assurance-Perspective.aspx

Sectie Onderzoek van Smals brengt met regelmaat verschillende publicaties uit over een hele waaier aan topics in de huidige IT-markt. U kan deze publicaties opvragen via het extranet :

http://documentatie.smals.be

Of u kan rechtstreeks contact opnemen met het secretariaat van de afdeling 'Klanten & Diensten', op het nummer 02/787 58 24.