# Toward the SIEM Architecture for Cloud-based Security Services

Jong-Hoon Lee
Information Security Research Division,
ETRI
Daejeon, Korea
mine@etri.re.kr

Young Soo Kim
Information Security Research Division,
ETRI
Daejeon, Korea
blitzkrieg@etri.re.kr

Jong Hyun Kim
Information Security Research Division,
ETRI
Daejeon, Korea
jhk@etri.re.kr

Ik Kyun Kim
Information Security Research Division,
ETRI
Daejeon, Korea
ikkim21@etri.re.kr

***Abstract*—Cloud Computing represents one of the most significant shifts in information technology and it enables to provide cloud-based security service such as Security-as-a-service (SECaaS). Improving of the cloud computing technologies, the traditional SIEM paradigm is able to shift to cloud-based security services. In this paper, we propose the SIEM architecture that can be deployed to the SECaaS platform which we have been developing for analyzing and recognizing intelligent cyber-threat based on virtualization technologies.**

***Keywords*— SIEM, Security Information and Event Management, SECaaS, Security-as-a-service, cloud-based security service.***

## I. INTRODUCTION

The cloud computing represents one of the most significant changes in the field of information security technology such as cloud-based security-as-a-service. Although there are many information security technologies for this purpose, the SIEM (Security Information and Event Management) has been developed as an important component of enterprise network and network infrastructures and it has been a purpose-built solution to collect, aggregate, parse, normalize, store, distill tremendous event logs and correlate data from traditional security systems such as firewalls, intrusion detection /prevention systems, anti-malware systems, and others that are deployed at both the host and network domains [1,2].

We have been developing the SOA (Security-on-Air) project which is cloud-based security platform. In cloud data center, it enables to provide various security services to the multi-tenants by applying SDN / NFV technologies and virtualizing the security sensors such as virtual firewalls, virtual IPS, virtual DLP, virtual DPI, anti-malware system and others that are deployed at both the host and network domains. The proposed SIEM can be applied to maintain a huge number of security event log which is generated from virtualized security systems for ensuring cloud-based security service.

For managing and analyzing the various logs and events which are generated by cloud-based security sensors in the SOA project, the SIEM needs to be designed not only to manage log and security events from various security systems, but also to achieve relevant correlation analytics for recognizing cyber threats. To do so, we referenced the OpenSoC [3] and complemented to our SIEM architecture for providing the various analysis model and data enrichment. In addition, because the main goal of the SIEM is to provide valuable security information provisioning and to perform a large-scaled data correlation for detecting cyber threats, we apply the Big Data platform which is composed of the distributed units based on Kafka, Spark, Elasticsearch and MongoDB [4, 5].

## II. DESIGNED SIEM ARCHITECTURE

The designed SIEM architecture mainly consists of the SIEM Engine for processing the collected data, the SIEM Storage for storing the collected data and analysis results, and the SIEM user layer for ensuring the security service to the user as shown in the Figure 1.

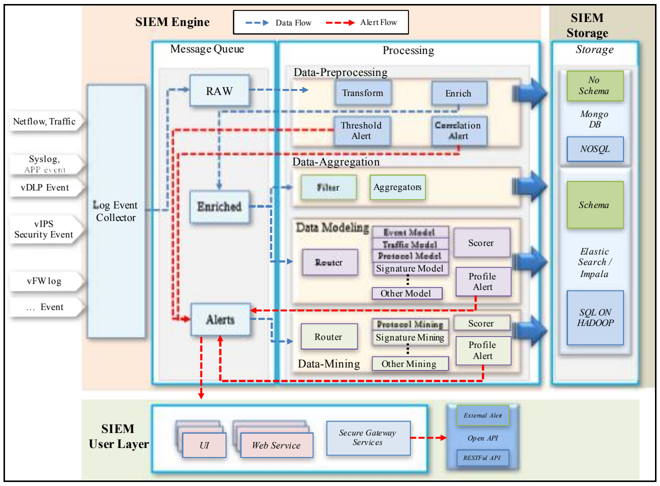

Fig. 1. The SIEM Architecture for cloud-based service

The SIEM engine aims to support provision of the intelligent threat analytics and relevant data output based on its various data processing such as data modeling and data mining. The details are explained in the next chapter. For the SECaaS service, the SIEM user layer is a specialized component which includes the application for accessing the SIEM engine and it enable to support incident response activities from a wide variety of sources.

In order to support the SIEM service in cloud platform, the SIEM user layer is executed on the virtual machine. In detail, upon the user activates the virtual machine which includes the SIEM user layer, and the SIEM Engine retrieves the related information which is retrieved from the SIEM Storage and instantly becomes running state. For this, it is necessary that the Data Identifier Manager in SIEM is able to identify each event and separate security log per each tenant of cloud-based security services.

## III. Analytics In SIEM Engine

The SIEM engine mainly includes the time-series analytics and the correlation analytics in order to provide cloud-based SIEM service. Each method is explained as follows.

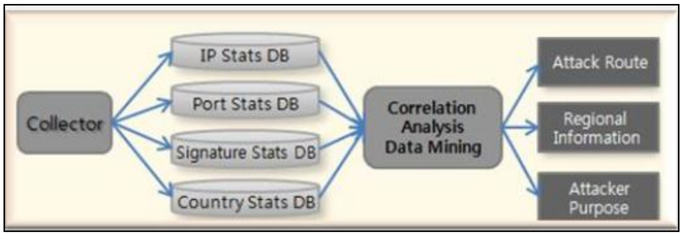

Fig. 2. Flow for Correlation Analysis using Data mining

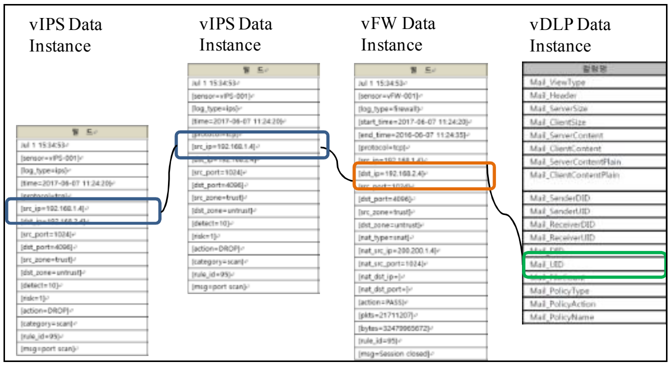

Fig. 3. Example of Correlated Information among vIPS, vFW and vDLP data instance

- Time-Series Data Analytics Function: It is to extract and calculate statistical data by security source and object in collected security event log data. The SIEM calculates predicted baseline values based on factors such as time, a day of the week, protocols and log data frequency for each service port. By doing so, it can carry out anomaly detection based on the calculated event baseline, attack name, traffic information which are stored in big data storage.
- Correlation Analytics Function: It provides the data information to analysis the correlated features among collected security events with the aggregated threat case dataset. For this, the SIEM generates statistical database using the big data analytics and data mining methodologies by IP address, port, signature of event and metadata for data enrichment from the collected data, and then it is able to provide the correlated data information to analyze and recognize the status of cyber threat by the correlated information. The Figure 2 represents the work flows of the correlation analysis module and the Figure 3 shows the example of correlated data information among the security sensor logs.

Additionally, the long-term's correlation analytics function enables to recognize and detect real-time intrusion threats by analyzing the event occurrence patterns related to past intrusion threats. It enhances the ability of the correlation analytics between attack cases which were occurred.

When the above SIEM engine performs the data analysis per each user, it must identify the collected data for each user. Therefore, this function is carried out by the Data Identifier Manager (DIM) that should recognize the data source by collected data from the various security sensor. This provides the functionality to run the customized SIEM engine according to whether the virtual machine of the cloud user is activated.

## IV. Future Works

In this paper, when the traditional security systems are virtualized in cloud platform, we designed the SIEM architecture for cloud-based security service that can help to recognize cyber threats using collected data and to provide correlation-based cyber threat analytics. Furthermore, by the reason that the correlation analytics is the most important one of the various analytics method, we will apply the Neural Network in order to detect the threat based on learning the security data model. In detail, by the neural network model which outputs the threat categories or normality by learning collected long-term data, the proposed SIEM can improve the ability to determine the threats that whether the status of the current collected data is threat or not. And such method significantly will enhance to improve intelligent cyber threat analysis in the SIEM.

## Acknowledgment

This work was supported by Institute for Information & communications Technology Promotion(IITP) grant funded by the Korea government(MSIP) (No.2017-0-00078, Cloud based Security Intelligence Technology Development for the Customized Security Service Provisioning)

## References

[1] YEN, T., OPREA, A., ONARLIOGLU, K., LEETHAM, T., ROBERTSON, W., JUELS, A. AND KIRDA, E. 2013. Beehive: Large-scale log analysis for detecting suspicious activity in enterprise networks. In Proceedings of the 29th Annual Computer Security Applications Conference, ACM, 199-208.

[2] Cloud Security Alliance, https://cloudsecurityalliance. org/group/ security-as-a-service/

[3] OpenSoC Project, http://metron.apache.org

[4] Apache Kafka Project, https://kafka.apache.org

[5] Apache Spark™ https://spark.apach.org