Enriching network and security events for better incident detection

Nick Buraglio[a,*], Ralph Koning[b,a], Cees de Laat[b,a], Paola Grosso[b]

[a] *Energy Sciences Network, Lawrence Berkeley Lab. Berkeley, CA, USA*

[b] *Universiteit van Amsterdam, Science Park 904, Amsterdam, The Netherlands*

Keywords: Security, IDS, netflow, detection, DDoS, Networks

Biography

Nick Buraglio has been involved in the networking industry in varying roles since 1997. Prior to joining the Network Engineering group at ESnet, Nick was employed by the University of Illinois as the Lead Network Engineer working on research, HPC, and wide area connectivity. In this role, Nick also functioned as the Lead Network Engineer and IP architect for the National Association of Telecommunications Officers and Advisors (NATOA) broadband project of the year, UC2B. Nick has also held Network Engineering positions at early regional broadband internet providers as well as at the National Center for Supercomputing Applications. Nick has participated in the SCinet working group on many occasions and has been involved in R&E, high performance networking and security since 2002. In addition to Network Engineering positions, Nick has been involved in cybersecurity from the campus, enterprise and service provider perspective and acted as a resource and trainer for the Federal Bureau of Investigation. Nick has been active in the SDN community since 2009 and is currently actively involved in several SDN and security related initiatives and projects. https://www.es.net/about/esnet-staff/network-planning/nick-buraglio/

Abstract

Given the current uptake of cyber attacks, record breaking DDoS events, and increasing frequency of attacks targeting infrastructure, carefully monitoring Internet systems and components for suspicious activities becomes not only desirable, but imperative. There are many developments in monitoring and intrusion detection systems (IDS) that enable them to trigger alerts when such activities are present [1, 2]. When such an event occurs it is the responsibility of the security and incident response teams that monitor this information to further investigate these events; this often requires them to locate and combine information from multiple sources to make a more informed judgment. This becomes increasingly difficult in a wide area network context due to the distributed nature, geographical diversity, and typically passive transit nature of the ecosystem. Automation and correlation of diverse data sets allowing for only the critical data has become necessary for near real-time responses to be realistic. Given the scale and frequency of large scale incidents and the diverse sources of attack, a higher level approach should be adopted by service providers, leveraging all data sources available in order to quickly judge and ascertain similarity of events based on vastly different perspectives and circumstances.

[*] Presenter

Email addresses: buraglio@es.net (Nick Buraglio), r.koning@uva.nl (Ralph Koning), delaat@uva.nl (Cees de Laat), pgrosso@uva.nl (Paola Grosso)

To aid in automating this process we introduce CoreFlow, a prototype framework to enrich network and security happenings; this enhancement provides more context to network and security events and this in turn creates more targeted alerts and the potential for faster, more advanced responses. This is in particularly important for transit networks that due to their characteristics require correlation of information sourcing from distant elements within the network.

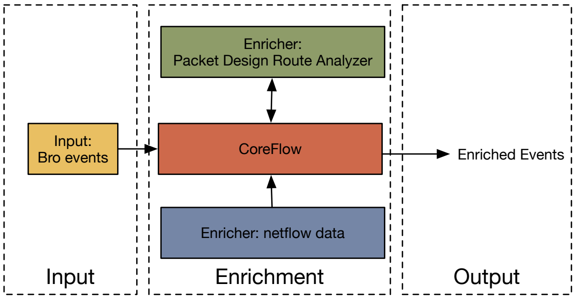

Figure 1: CoreFlow correlates input data from Bro to NetFlow and uses the enriched data to query the route analyzer. Finally, it outputs the security event with additional data from both enrichers.

CoreFlow [3] ingests data from many sources, focusing primarily on the Bro[1] IDS and augments this with other available network event data sources such as flow data, syslog messages, and topological data from the devices within a given network. The architecture of CoreFlow is composed of three distinct phases: input, enrichment, and output. This is shown in Figure 1.

The input phase ingests events from an IDS, in this case Bro. The enrichment phase then correlates the information with other data sources, detecting similarities in patterns and traffic behavior. The output module then displays the enriched data to the console, a log file or into elastic search. Events that were not correlated will pass through CoreFlow unaltered so no information will be lost during this process. Table 1 shows a list of the available data input modules.

Input	**Enrichment**	**Output**
• raw bro notice log • raw bro notice log via stdin • bro notice log via splunk	• raw bro connection log • nfdump formated netflow data via NFS • bro connection logs via splunk • Palo Alto logs via splunk (under development)	• stdout • log file • elasticsearch

Table 1: The input formats currently supported by CoreFlow for correlation

Initial CoreFlow prototype testing was done within the ESnet network[2], correlating inputs from 3 Bro systems that were collected by Splunk[3] and netflow information of more than 50 routers, shared a from central storage. By confirming the presence of the malicious flow on routers in the network we are able to deduce the path that the flow took utilizing the internal routing table, provided by a packet design route explorer appliance. This allows for placement of countermeasures at borders or within the interior of

[1] https://www.bro.org/
[2] http://www.es.net
[3] https://www.splunk.com/

the network, and also provides confirmation of valid or invalid routing sources, aiding in the detection of potential spoofed addresses. An updated build of CoreFlow was tested at the Super Computing 2017 network infrastructure where we leveraged Bro connection logs instead of NetFlow data for correlation, correlating with log data as well as other sources of network event information.

Enriching IDS data with NetFlow information gives a better view of an attack for several reasons, but in very large service provider networks 1:1 flow data records are difficult to generate, collect, store, and consume so in many cases the flows are sampled in order to reference them for capacity planning and traffic baseline creation. Because of this, forensic data on core backbone flow data can prove to be difficult. With that in mind, CoreFlow provides a correlation framework that can combine data sources based on the flow tuples. The successfully enriched data provides more context that can be used for more advanced attack mitigation. Additionally, this context can also improve alerting and action, research is underway on how to automate actions based on this new information. Examples of this continuous cycle include feeding data back into the IDS system to reduce false positives; it may even be beneficial to lower the threshold for IDS events sent into CoreFlow to discover malicious events that previously went undetected.

Acknowledgments

This work is funded by the Dutch Science Foundation project SARNET (grant no: CYBSEC.14.003 / 618.001.016) and by the Dutch project COMMIT (WP20.11). Special thanks go to our research partners CIENA, TNO and KLM.

Initial research and prototype work was performed by Nick Buraglio and Ralph Koning at Lawrence Berkeley National Lab during the summer of 2016. This research used resources of the ESnet, which is supported by the Office of Science of the U.S. Department of Energy under contract DE-AC02-05CH11231.

[1] H. Debar, M. Dacier, A. Wespi, Towards a taxonomy of intrusion-detection systems, Computer Networks 31 (1999) 805–822.

[2] M. V. Mahoney, A machine learning approach to detecting attacks by identifying anomalies in network traffic, Ph.D. thesis, Florida Institute of Technology, 2003.

[3] R. Koning, N. Buraglio, C. de Laat, P. Grosso, Coreflow: Enriching bro security events using network traffic monitoring data, Innovating the Network for Data Intensive Science (INDIS) workshop at Super Computing 2016, Salt Lake City (UT) (2016).