EPJ Web of Conferences **214**, 03029 (2019) https://doi.org/10.1051/epjconf/201921403029
CHEP 2018

Operational security, threat intelligence & distributed computing: the WLCG Security Operations Center Working Group

David Crooks[1,*], *Liviu* Vâlsan[2,**], *Kashif* Mohammad[3,***], *Shawn* McKee[4,****], *Paul* Clark[5,†], *Adam* Boutcher[5,‡], *Adam* Padée[6,§], *Michał* Wójcik[6,¶], *Henryk* Giemza[6,‖], and *Bas* Kreukniet[7,**] on behalf of the WLCG Security Operations Center Working Group

[1]STFC Rutherford Appleton Laboratory, Harwell Campus, Didcot, Oxfordshire OX11 0QX, United Kingdom
[2]CERN, The European Organisation for Nuclear Research, 1211 Geneva 23, Switzerland
[3]University of Oxford, Department of Physics, Denys Wilkinson Building, Keble Road, Oxford, OX1 3RH, United Kingdom
[4]Physics Department, University of Michigan, Ann Arbor, MI 48109-1040 USA
[5]Institute for Particle Physics Phenomenology, Ogden Centre for Fundamental Physics, Department of Physics, University of Durham, Science Laboratories, South Rd, Durham DH1 3LE, United Kingdom
[6]Narodowe Centrum Badań Jądrowych, ul. Andrzeja Sołtana 7, 05-400 Otwock-Świerk, Poland
[7]SURFsara, SURF Science Park Building, Science Park 140, 1098 XG Amsterdam, The Netherlands

Abstract. The modern security landscape for distributed computing in High Energy Physics (HEP) includes a wide range of threats employing different attack vectors. The nature of these threats is such that the most effective method for dealing with them is to work collaboratively, both within the HEP community and with partners further afield - these can, and should, include institutional and campus security teams. In parallel with this work, an appropriate technology stack is essential, incorporating current work on Big Data analytics. The work of the Worldwide LHC Computing Grid (WLCG) Security Operations Center (SOC) Working Group (WG) [1] is to pursue these goals to form a reference design (or guidelines) for WLCG sites of different types. The strategy of the group is to identify necessary components - starting with threat intelligence (MISP [2]) and network intrusion detection (Bro [3]), building a working model over time. We present on the progress of the working group thus far, in particular on the programme of workshops now underway. These workshops give an opportunity to engage with sites to allow the development of advice and procedures for deployment, as well as facilitating wider discussions on how to best work with trust groups at different levels. These trust groups vary in scope but can include institutes, National Grid Infrastructures and the WLCG as a whole.

*e-mail: david.crooks@stfc.ac.uk
**e-mail: liviu.valsan@cern.ch
***e-mail: kashif.mohammad@physics.ox.ac.uk
****e-mail: smckee@umich.edu
†e-mail: paul.w.clark@durham.ac.uk
‡e-mail: adam.j.boutcher@durham.ac.uk
§e-mail: adam.padee@ncbj.gov.pl
¶e-mail: michal.wojcik@ncbj.gov.pl
‖e-mail: henryk.giemza@ncbj.gov.pl
**e-mail: bas.kreukniet@surfsara.nl

© The Authors, published by EDP Sciences. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (http://creativecommons.org/licenses/by/4.0/).

1 Introduction

The threat landscape faced by HEP computing sites includes attack vectors from a wide range of sources. These vectors, which can include ransomware and phishing campaigns, have the potential of affecting many sites and communities within a short time; as a result, the most effective way of dealing with these threats is via collaboration and the sharing of threat intelligence. This sharing can benefit members of the HEP computing community as well as related communities.

Effective threat intelligence forms one part of the work of the WLCG SOC WG, whose primary goal is to provide guidance to WLCG computing sites. The second area of work is a set of guidelines for an appropriate technology stack to make effective use of this intelligence. That implies monitoring the state of individual clusters using modern analytics techniques, while being mindful of recent increases in the use of virtualisation and containers for the development of HEP computing site structure. This collection of tools for monitoring the state of a cluster, sharing threat intelligence and performing analysis is known as a Security Operations Centre.

2 Areas of work

2.1 Technology stack

The first goal of the working group, as alluded to in the previous section, is the creation of a model SOC technology stack primarily for use by WLCG sites. It is anticipated that the varying structure and configuration of sites (particularly concerning network layout and available sources of data, for example) suggests a strategy of providing a set of options for sites of different kinds. Given the potential complexity of a SOC, the strategy used thus far is to identify the key areas discussed above and use those to form a core set of components. This would then be built upon with other tools, both to extend SOC capabilities over time and to reflect individual site requirements.

The growing use of virtualisation technologies in HEP computing sites means that the visibility into individual job payloads and processes can be considerably reduced over that found in "traditional" grid sites using worker nodes deployed on bare metal. As a result, network monitoring is a critical area. Building on experience in the US HEP community and that of the CERN Computer Security Team, the Bro [3] network Intrusion Detection System (IDS) was selected by the WG as a starting component.

2.2 Collaboration and intelligence sharing

A parallel, but no less important, goal of the working group is to pursue the sharing of quality threat intelligence within our community and beyond. Unlike the technology stack, this is largely a cultural and social area of work; the sharing platform MISP is well established and in use by many organisations [4]. The WG selected MISP based on extensive operational and development experience by the CERN Security Team, who use this tool as part of daily operations. The strategy pursued thus to date is to encourage participating sites to install MISP and synchronise with the central WLCG instance hosted at CERN. In addition, the UK is pursuing the idea of a national MISP instance which can then be used as a central hub for the other UK sites.

A key element of this work is encouraging collaboration between grid and institutional security teams; we hope to pursue this by identifying sites with experience of this and developing case studies and examples.

3 Status

An effective means of development was found to be a set of workshops. The first of these, held in December 2017, was largely an introductory event lasting 1.5 days. A second workshop was held in June 2018, which explored more in depth topics. These included network topology (section 3.1), real-time indexing and visualisation (section 3.2), and advanced processing, enrichment, aggregation and correlation of notifications (section 3.3).

In this paper we predominantly cover the work done during the June 2018 workshop. The work achieved during the December 2017 workshop, which largely covered an introduction to the deployment of MISP and Bro, has been addressed in a previous publication [8].

A general outcome of this workshop was a need for validation of deployments to ensure the correct operation of each component. This, along with work on the integration of these components, was found to be a key area of future work.

3.1 Network topology and IDS deployment

During the June 2018 workshop five different WLCG sites presented five different network topologies, with an emphasis on different possible network tap points and strategies for traffic mirroring.

3.1.1 CERN

CERN presented their complex networking layout used to aggregate a mirror of relevant networking traffic and distribute it among a cluster of IDS systems running Bro. Such a complex, scalable and extensible approach is required given the high volume of networking traffic entering and leaving CERN. At the time of the workshop CERN was having a total bandwidth of 160 Gbps to the general Internet, while connections to scientific networks amounted to 660 Gbps. Some of the guiding design principles for the CERN setup were the ability to analyse traffic at boundaries between different networks domains, aggregate and load-balance the networking traffic across a set of intrusion detection servers, and leverage advanced features of networking hardware

In terms of requirements, symmetrical load-balancing of network flows represents the core functionality of the system. That ensures that for a given network flow, both directions are being forwarded to the same IDS server.

Lastly, one of the main requirements was to build a system that is easy to maintain and extend, providing flexibility to adapt to changing needs and increased network rates.

The system build by CERN is using the Extreme Networks SLX 9540 high-end data centre networking switch on the hardware side in conjunction with StackStorm [5] for automated network device configuration.

The overall networking topology for mirroring traffic to IDS systems is shown in figure 1(a). The input Link Aggregation Group (LAG) is used to aggregate traffic from different network borders: between the CERN campus network and the public Internet as well as between different network domains for scientific computing. The output LAG is used to load-balance the aggregated traffic to a cluster of IDS servers.

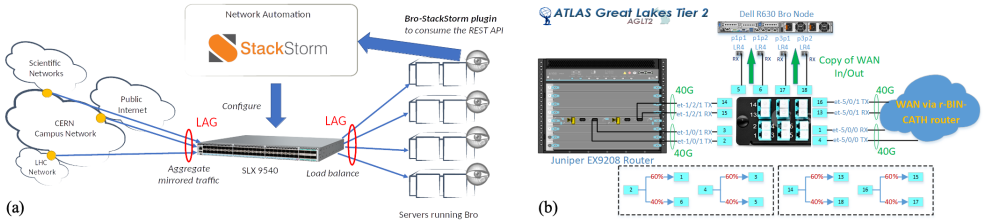

Figure 1. (a) CERN networking layout (b) The AGLT2 network configuration of the optical tap used to provide Bro with a copy of incoming and outgoing network data.

StackStorm is being used for automation, to manage the configuration of the network traffic aggregator and also to provide shunting functionality. Different policies can be implemented as Bro scripts for detecting trusted bulk data transfers, triggering a StackStorm action to deploy an Access Control List (ACL) for preventing further data packets for being forwarded, while allowing the TCP control flags that are used to signal the end of the connection. Like that the IDS systems are offloaded by preventing the deep packet inspection of trusted bulk transfers, while still retaining basic connection details, including duration.

3.1.2 Oxford WLCG Tier-2

The WLCG Tier-2 site at Oxford has a small setup where data is captured through port mirroring on a 10 Gbps switch and fed to Bro running on a server with 16 cores and 96 GB of RAM. Oxford is using the PF_RING [6] enabled Bro packages provided by CERN. The use of PF_RING allows to distribute the network traffic among a set of Bro worker processes, with proper CPU pinning.

3.1.3 Michigan WLCG Tier-2

The ATLAS Great Lakes Tier-2 (AGLT2) at the University of Michigan has a single border router (a Juniper EX9208) with two 40 Gbps connections to the wide-area network. Shown below in figure 1(b) is a diagram of how a dedicated server running Bro is analysing a copy of all the incoming and outgoing traffic on these two 40 Gbps connections.

3.1.4 Narodowe Centrum Badań Jądrowych[1]

The initial version of the Narodowe Centrum Badań Jądrowych (NCBJ, Otwock-Świerk, Poland) SOC architecture was designed as a test framework to evaluate and experiment with different SOC components. The solution presented during the workshop was a basic, cost free implementation which utilised existing hardware and software. The design principles included analysis of the external network traffic in select networks only, evaluation of the capabilities of hardware and software traffic mirroring solutions, and integration with the existing software solutions

During the analysis it was decided that the most important networks to monitor are the DMZ and the campus network. Because of a physical network separation implemented at NCBJ, DMZ could only be monitored at the firewall or the first switch stack behind it. Since initially all 10 Gbps Ethernet ports in the stack were occupied an alternative solution, i.e.

[1]National Centre for Nuclear Research

software port mirroring (daemonlogger [7]), was considered. However later it turned out that there there might be a possibility to free up one port so the port mirroring capability provided by Juniper EX-4500 was utilised in the final solution.

In case of the campus network a similar traffic mirroring strategy was considered, however, the tap was planed to be done at the switch stack (Juniper EX-4300) connected to the edge router (Juniper MX-80) as this was the closest common point between the campus network and the HPC centre where the IDS system was located.

As the last stage the traffic was redirected to a dedicated IDS node (32 CPU cores, 128 GB RAM, 2x 10 Gbps links) via a chassis switch which was capable to join the traffic from both networks. This however was not a perfect solution as only one node link could be used to monitor the traffic since the second one was required for internal purposes.

Nevertheless such solution was acceptable since only partial traffic of the NCBJ 6.5 Gbps Internet connection was analysed and the available link capacity was never exceeded.

Future work involves extending the monitoring to cover the entire external traffic and some specific traffic between critical internal networks as well. This plans however require an upgrade of the network equipment installed at the edge and an increase of the monitoring ports in the IDS system.

3.1.5 SURFsara

SURFsara is the National Compute Centre in Holland, providing networking connectivity for computing infrastructures dedicated to different science fields, including HEP. They currently employ multiple lines of defence: stateful firewall, extensive rules sets, IDS / Intrusion Prevention System (IPS), Deep Packet Inspection (DPI), vulnerability scans, and more.

SURFsara is currently investigating the addition of Bro as yet an additional line of defence at the border of its network.

3.2 Real time indexing and visualisation

3.2.1 CERN

Real time indexing and visualisation of security logs collected by the CERN Security Operations Centre is done using Elasticsearch. Two different Elasticsearch are used, one dedicated for storing Bro logs and the second used for storing all other types of security logs. Both clusters are being centrally managed by the team in charge of the CERN central IT Elasticsearch service.

The Computer Security Elasticsearch clusters comprise two types of data nodes: SSD based nodes with a capacity of 1.6 TB / node and HDD based nodes with a capacity of 120 TB of storage / node. New indices get created and data gets written to SSD based nodes, with Curator moving old data from SSD based nodes to HDD based nodes.

Custom made templates are used for each different source of data, ensuring that the most efficient field type is being used, both in terms of the storage of data but also from the point of view of indexing, querying and aggregating the data.

Indices of different size are using a varying number of shards. As a general rule we try to have around 10 GB of data per shard, providing a good balance between the amount of data per shard and the total number of shards.

For convenience, users can access one single Kibana instance that's providing transparent access to data stored in both clusters.

3.2.2 Oxford WLCG Tier-2

Bro captures the network traffic and classifies it on the basis of application layer protocol, generating separate log files for protocols such as SSH, HTTP, etc. We are using Filebeat [11] to transfer these log files from the Bro server to a server running the Elasticsearch, Logstash, Kibana (ELK) stack. Filebeat is configured to add logtype tags based on protocol for conditional processing.

Logstash [12] receives data from Filebeat and passes it to Elasticsearch through the use of filters. Input is configured to capture everything on a particular port and output is configured to pass all filtered data to Elasticsearch but most of the interesting processing happen at the filter stage.

Filter matches every line of Bro logs with a predefined pattern and assigns values to various fields. It also applies string and mathematical functions like mutate, join, add, etc on the data to create a meaningful data structure. Logstash can also use external plugins to enrich data; one of the plugins which is used in our setup is geo_ip lookup to translate IP addresses to geographical locations.

Data from Logstash is consumed by ElasticSearch where we use Kibana to visualise and analyse interesting network patterns.

3.2.3 Michigan WLCG Tier-2

The configuration of the SOC at Michigan relies upon a local Elasticsearch cluster, originally dedicated to capturing system and device logging. While the integration between this Elasticsearch cluster and Bro is not yet completed, having Elasticsearch made it very simple to incorporate add ons. In addition to Bro monitoring, we wanted to have better visibility into our network traffic. To provide this visibility we installed ElastiFlow [10]. Elastiflow enables Netflow/IPFIX records to be incorporated with Elasticsearch and provides a very nice set of Kibana dashboards to visualise and track the flow data. We had some challenges getting the sflow-codec and the Kibana elastiflow index imported and ended up just "learning" the ElastiFlow index from the incoming data. One example of how to install and configure ElastiFlow is shown in [14]. Once Elastiflow was setup we just needed to configure our Juniper router to send sflow records to it. Shown in figure 2 is an example of a typical ElastiFlow dashboard.

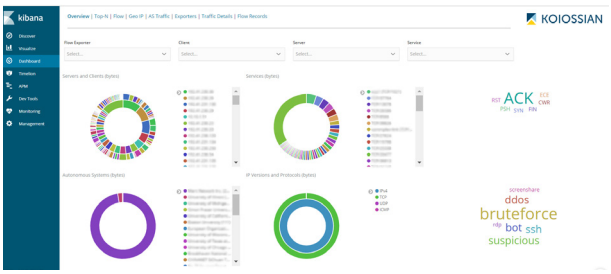

Figure 2. Example of the main Elastiflow Kibana page at AGLT2.

3.3 Advanced processing, enrichment, aggregation and correlation of notifications

The work presented in this section has been contributed by CERN and is based on their system for advanced processing, enrichment, aggregation and correlation of alerts produced

by the CERN SOC. The primary aim of this system is to reduce the number of false positives and to generate alerts that can be quickly and effectively handled by the security analyst.

Alerts are processed in near real time, in 5 minute intervals. Related alerts are aggregated and correlated. Different alerts are considered to be related if the different Indicators of Compromise (IoC) that triggered the alert are part of the same MISP event or of related (correlated) MISP events.

Generated alerts are enriched with additional sources of information to allow easy triage of alerts. These additional sources of information include, but are not limited to: MISP, CERN's networking database, reverse DNS, WHOIS information, GeoIP data and external threat intelligence services.

The alerts also contain additional information providing context around the time of the alert and the involved computing resources. For example, besides the basic network connection metadata, the generated alerts also include application specific logs (e.g. HTTP connections, files transferred, DNS queries made, etc).

Great emphasis has been placed on dealing with false positives. The additional context from MISP is used for filtering out false positives. For example, alerts raised following a connection to a potentially malicious IP are being filtered out, if the MISP event contains a domain name linked to that IP (i.e. via a composite attribute or inside a MISP object) and the detected connection was made to a different domain that the one from the MISP event.

3.4 Associated work

The Durham site presented on their work on a distributed Fail2Ban[13] configuration, describing a set of jails which would increase the length of time a particular IP would be banned based on how persistent the activity was.

The Institute for Particle Physics Phenomenology (IPPP) based out of the University of Durham is currently investigating pre-emptive threat blocking across its public facing GridPP High Performance Computing (HPC) cluster using a customised Fail2Ban framework. The framework is used to counter unauthorised access, brute force and denial of service attacks across the site by making use of a shared resource pool of known IP addresses that are believed to be used for such attacks.

To accomplish this the system makes use of a series of nodes running the Fail2Ban client with a number of custom jails and actions to push and pull recent attack data from a shared database. When Fail2ban encounters a threat that IP address is blocked and the information pushed to the database where it is disseminated throughout the site. This halts the attackers progress in moving from system to system to attempt the same attack with the hope of gaining a successful login. This continuous delay slows the attacker down enough that they will hopefully give up and allow enough time for site admins to be alerted to the unauthorised access attempts so further actions can be taken.

To reduce the chances of permanently blocking out entire addresses ranges from known sources each block hit enters a low level jail for a certain period of time. The more hits against this IP, the more it is elevated to a more severe jail until it is finally put into a permanent ban list, requiring an administrator action for removal. In the near future it is hoped this system will be integrated to work with the Malware Information Sharing Platform (MISP) to pull down known Indicators of Compromise and to push any confirmed malicious IP addresses that may be useful to other sites. To increase the data gathering ability of the system, it will also be adapted to make use of target data gathered by remote honeypot systems that will be implemented at a later date. It is hoped that once full implementation is accomplished the system will be capable of performing highly automated defensive actions, making use of an open source, easy to manage framework that can be integrated into other environments

with little to no issues occurring. Figure 3 shows a block diagram of the deployed Fail2Ban framework.

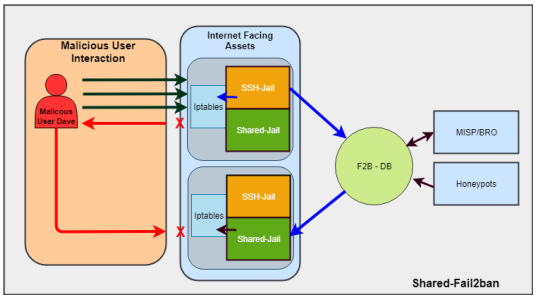

Figure 3. Shared Fail2Ban framework

4 Conclusions and future work

With the second workshop of the working group, a wide range of SOC components have been explored, including network monitoring, threat intelligence, data storage and visualisation, and associated tools. Key areas of future work include both the validation of deployed components along with the integration of these components into a working model. A future publication will report on a proposed scalable SOC reference design applicable to a wide range of WLCG sites.

At time of writing the next workshop is planned to take place at RAL, UK, in early 2019. The planned topics for this workshop include the exploration of the areas of validation and integration in particular, with a view to forming a complete, integrated model. This work will form the basis of a future publication.

References

[1] https://wlcg-soc-wg.web.cern.ch/
[2] https://www.misp-project.org
[3] http://www.bro.org
[4] https://misp-project.org/communities/
[5] https://stackstorm.com
[6] https://www.ntop.org/products/packet-capture/pf_ring/
[7] https://www.talosintelligence.com/daemon
[8] D. Crooks and L. Vâlsan (2018) Harnessing the Power of Threat Intelligence in Grids and Clouds: WLCG SOC Working Group. International Symposium on Grids & Clouds 2018 (ISGC 2018), BHSS, Academia Sinica, Taipei, 16-23 March 2018, *in press*
[9] https://www.elastic.co
[10] https://github.com/robcowart/elastiflow
[11] https://www.elastic.co/products/beats/filebeat
[12] https://www.elastic.co/products/logstash
[13] https://www.fail2ban.org
[14] https://pandaways.com/elastiflow-with-mikrotik-and-centos-7