Proceedings of the 51^{st} Hawaii International Conference on System Sciences | 2018

Host Inventory Controls and Systems Survey: Evaluating the CIS Critical Security Control One in Higher Education Networks

Philip Kobezak^{*†}, Randy Marchany^{*†}, David Raymond[†], Joseph Tront^{*} *Bradley Department of Electrical and Computer Engineering, Virginia Tech, Blacksburg, VA [†]Information Technology Security Lab, Virginia Tech, Blacksburg, VA {pdk, marchany, raymondd, jgtront}@vt.edu

Abstract

Within the field of information security, the identification of what we are trying to secure is essential to reducing risk. In private networks, this means understanding the classification of host end-points, identifying responsible users, and knowing the location of hosts. For the context of this paper, the authors are considering the challenges faced by higher education institutions in implementing the first Center for Internet Security (CIS) Critical Security Control: inventory of authorized and unauthorized devices. The authors developed and conducted a survey of chief information security officers at these institutions. The survey evaluated their confidence in meeting the goals of host inventory tracking. The results of the survey, along with analysis of the implications for information security operations, are presented in this paper. Changes in technology, such as BYOD, IoT, wireless, virtual machines, and application containers, are contributing to changes in the effectiveness of host inventory controls.

1. Introduction

At the beginning of a normal workday, an analyst is monitoring for incidents in a security operations center. The analyst is enjoying a slow start so they are catching up on emails from the previous day. Unfortunately, it does not take long before they see an alert from one of the institution's intrusion detection systems. The analyst is concerned because this alert is for a particularly nefarious type of malware associated with theft of personally identifiable information. As the analyst creates a ticket to begin the response process, another alert comes up. This time for a host identified with a ransomware download. The analyst recognizes the IP address as being in one of the administrative areas of the institution. The analyst knows that if the ransomware executes, it will begin encrypting the user's local files and any folders on a file server. Even if backups of the data are available, either incident could lead to data exfiltration. Now the analyst must work fast to notify responsible individuals quickly. If the tools available to the analyst cannot provide an answer to who they should contact, or the tools provide the incorrect person, more time will be spent finding the responsible user while the malware is in control and potentially doing harm.

In incident response, the time between initial identification and containment is critical to reducing damage particularly when sensitive or high-risk data is involved [1]. This is particularly true with modern malware moving to mobile devices and evolving to include theft of messages, position data, and banking credentials, all with real-time attacker command and control [2].

2. CIS Critical Security Control one

Organizations must prioritize the application of resources in the defense of cyber-attacks to minimize risk to their networks. Cyber security controls frameworks help with this prioritization, and often recommend specific methods, software, and systems to implement individual controls. Johnson states "all security and corporate managers now need to be concerned with compliance and governance of risks, security, and the information usage in their systems" [3]. This is especially true for higher education institutions that conduct research and must comply with mandates to defend against cyber-attacks or risk losing funding.

CIS is a not-for-profit organization "dedicated to enhancing the cyber security readiness and response among public and private sector entities" [4]. The CIS Critical Security Controls (CSC) for Effective Cyber Defense exist as a framework to help organizations improve their information security strategy. The Controls were developed by experts from many different organizations who "pooled their extensive first-hand knowledge from defending against actual cyber-attacks to evolve the consensus list of Controls, representing the best defensive techniques to prevent or track them" [5]. The twenty Controls are "a prioritized,

URI: http://hdl.handle.net/10125/50486 ISBN: 978-0-9981331-1-9 (CC BY-NC-ND 4.0) highly focused set of actions that have a community support network to make them implementable, usable, scalable, and compliant with all industry or government security requirements" [5]. The CSC framework is intended to provide an organization with key areas where they should specifically focus their efforts. Each Control gives example technologies that an organization can implement to help achieve their goal of reducing risk. As no single measure is guaranteed to prevent cyber security incidents, organizations are encouraged to implement all the Controls to have a defense in depth strategy.

In this paper we focus on the first Control outlined in the CSC: inventory of authorized and unauthorized devices. As of version 6.1 of the CSC, six sub controls are defined for the first Control.

CSC 1.1, "deploy an automated asset inventory discovery tool..." [5] is common for Internet Protocol version 4 (IPv4) networks. Organizations can scan their network address space to identify hosts, and even attempt operating system identification. Nmap and other tools can provide this ability [6]. Unfortunately, scanning an Internet Protocol version 6 (IPv6) network is not so straightforward, due to the extremely large address space and time it would take to iterate through each to send probe packets to solicit a response [7]. In more recent years, passive scanning, or the listening for active hosts on the network, has become more common. This involves "the process of monitoring network traffic at the packet layer to determine topology, services, and vulnerabilities" [8].

CSC 1.2, "deploy dynamic host configuration protocol (DHCP) server logging..." [5] is something that most organizations can easily implement. By simply logging DHCP server events, we can better track hosts on the network. This is commonly used in IPv4 networks; however, depending on the IPv6 deployment, DHCP may or may not be used. IPv6 networks may use Stateless Address Autoconfiguration (SLAAC), DHCPv6, or statically assigned addresses [9], [10].

CSC 1.3, "ensure that all equipment acquisitions automatically update the inventory..." [5] is fundamentally a business process. To comply, organizations must make sure there are automatic updates to the inventory based on new acquisitions. This can be accomplished by integrating an Enterprise Resource Planning (ERP) application with the inventory system. Doing these updates manually becomes problematic for many organizations, requiring data entry in the business and financial applications that is IT-specific. In some organizations there is a fundamental decoupling of business operations from network operations.

CSC 1.4, "maintain an asset inventory of all systems connected to the network..." [5] describes an all-

encompassing inventory. This Control seems to be solved by using a database-driven application to track this information. This is common, as are spreadsheets, in many organizations. However, the accuracy of these manual processes usually erodes over time, given the significant effort required by personnel to enter and update each host's details. This technique also does not scale for networks with tens or hundreds of thousands of hosts. Something that is common to research institution networks is the ability to Bring Your Own Device (BYOD) and connect it to the network. While many corporate networks are able to resist BYOD, higher education has seen this for decades. This means that CSC 1.3 is not applicable in this situation since the owner of the device is not the same as the owner of the network.

CSC 1.5, "deploy network level authentication via 802.1x..." [5] requires every host to be authenticated to the network. This is commonly deployed for wireless and some wired networks. Sometimes it is also deployed to authenticate Voice over IP (VoIP) devices to separate Virtual Local Area Networks (VLAN). Depending on the end points, this may be less feasible to deploy across the entire network of an organization. There may also be limitations on the deployment of 802.1x with older networking equipment.

Finally, CSC 1.6, "use client certificates..." [5] requires the use of certificates to authenticate each device instead of a username and password. Client certificates are a highly secure method of authentication but do carry significant management overhead.

3. Survey design

Our higher education CISO survey was designed to answer the following high-level questions: Are new technologies changing the accuracy of inventory controls? How quickly can the location of a host and the responsible user be identified? Are current host inventory controls effective? Have there been changes in effectiveness due to increases in Internet of Things (IoT) and BYOD hosts? Do the responses vary with subsets of the population such as size of the network or number of employees dedicated to information security operations?

The survey was also intended to look for correlations between sizes or types of institutions and network architectures. Network architectures will vary with the type of institution to include the amount of research, number of residential students, and user population size. Some of the questions were based on similar surveys that identified current challenges in information security for higher education institutions [11]. The survey was reviewed by several information technology professionals for the quality of questions and answers. Many questions were modified or removed to reduce ambiguity and improve readability. The survey was tested prior to release and the respondents' test results were also used to improve the questions.

No personal information of the respondents was solicited. The only identifying attribute recorded was the respondent's IP address. This was used to identify whether multiple responses were recorded for the same institution. This also gave the ability to delete a response at the request of a respondent by asking them to verify the IP address they used. Even with this single piece of information that could be linked to the institution, it will be removed once the survey is closed. This is to encourage honest responses without fear of the respondent being identified.

The target population for the survey is all higher education institutions in the United States. According to the Carnegie Classification, there are approximately 4,600 institutions [12]. The first question of the survey was to identify the respondent's institution's Basic Carnegie Classification. Additionally, questions were asked to determine the size and attributes of each institution to include numbers of students, employees, employees in information security roles, and estimated research expenditures. This first section of the survey was used to provide a framework for comparison of like institutions only. The second, third, and fourth sections of the survey asked questions pertinent to network size, host identification, and evaluation of controls. Samples of those survey questions are mapped to the appropriate CSC control in Table 1.

Table 1: Select	questions	mapped to	the CSC
-----------------	-----------	-----------	---------

Survey Question	CSC
2.1. What is your best estimate for the peak	1.2,
number of hosts on your network at one time?	
	1.5
2.2. What is the average number of BYOD	1.3,
hosts each type of end-user connects to your	
network?	
2.4. What is your best estimate for the number	1.1,
of sub-networks (Local Area Network	
segments or broadcast domains)?	
2.10. Where does your institution allow	1.3,
embedded hosts or Internet of Things (IoT) on	
your network?	
3.1. What is the estimated percentage of each	1.3,
type of host on your network?	1.4
3.3. What IP addressing methods do you use?	1.1,
	1.4

3.7. How confident are you in your organization's ability to identify hosts with multiple, changing addresses, to include application containers (Docker) and IPv6 privacy extensions (RFC 4941)?	1.1, 1.4
3.8. What percentage of hosts on your network	1.5,
utilize some form of network authentication to connect (IEEE 802.1x, NAC, etc.)?	1.6
4.1. For the purposes of your host inventory	1.3,
controls, what types of hosts do you track?	
4.2. During a potential security incident or	
event, how long does it usually take to track	
down the responsible user or owner of these	
host types?	
4.6. How accurate have you found the	1.4,
following tools and technologies to be in	
keeping track of hosts in your network?	
4.7. Do you consider embedded devices or IoT hosts more difficult to track than other hosts?	1.4

The survey had 42 questions but some asked the respondent to answer for different cases which results in up to 96 total data points. Only one question had a required response due to validation needed to constrain the sum of the response to one hundred percent.

4. Results

The survey was opened for distribution to participants on May 24, 2017. Since the survey was targeting Chief Information Security Officers of higher education, several email lists were used to distribute the anonymous link. Most respondents completed the survey in less than 20 minutes.

4.1. Institution classification

These survey results cover 51 responses. More than half of the respondents reported their institutions to be R1, R2, or R3 doctoral granting universities with research activity as shown in Figure 1.

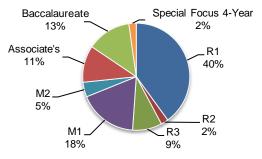


Figure 1: Percentage of respondent institutions by Basic Carnegie Classification

Using the high-end of the selected ranges for total employees, employees in IT, and employees in security operations, ratios were calculated. The results showed the ratios of employees to be 5.3 percent for IT to total employees and 5.4 percent for information security to IT employees.

There was a wide variety of reported enrolled students with most reporting between 2,000 and 50,000. There was some variance with the number of reported remote students. However, more than half of the respondents reported 10 percent or fewer remote students.

4.2. Network characteristics

For question 2.1, most of the respondents selected the peak hosts on their network to be 10,000 to 50,000. Eleven respondents said their networks were greater than 50,000. One stated they had more than 500,000 peak hosts on their network.

One important question asked, "How is the respondent's network managed?" The results are shown in Table 2.

Table 2: CISC	D's ability to	deny or allow hosts
---------------	----------------	---------------------

	Response Option	%
1	The CIO or CISO has the ability to deny or	64
	allow all hosts on the network	04
2	The CIO or CISO has the ability to deny or	26
	allow most hosts but not all	20
3	The network is mostly federated. Most	
	organizational units control their networking,	2
	to include network equipment and hosts	
4	The network is completely federated. The	
	CIO or CISO has no ability to allow or	0
	disallow hosts.	
5	Other	7

This question may have been interpreted differently than anticipated. The intent was to determine how federated or completely centralized the institution's IT functions were. If the respondents understood the question, it is possible that their institutions are mostly centralized in terms of managing the network.

Figure 2 shows the responses to question 2.2, BYOD host percentages by user type. The differences are particularly pronounced for residential students in the 4 to 6 range.

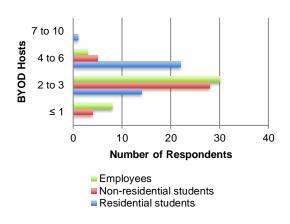
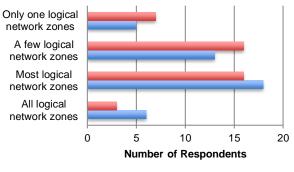


Figure 2: BYOD host quantity by user type

For the average number of BYOD hosts connected by non-residential students, 67 percent of respondents said 2 to 3. This differs from the number of BYOD hosts connected by residential students, which was split between 2 to 3 and 4 to 6. This is not surprising, as you can assume that residential students will connect devices in their dorm rooms that they would otherwise keep in an off-campus residence. What we found surprising is that 73 percent of respondents said employees connected 2 to 3 BYOD hosts. This means that most institutions expect employees to connect 2 to 3 personal, BYOD devices that are not institutionally owned.

Questions 2.9 and 2.10 in the survey asked where the institution allowed BYOD and embedded or IoT hosts. The results in Figure 3 show that most respondents chose "most logical network zones" for BYOD hosts.



Where do you allow embedded or IoT hosts
 Where do you allow BYOD hosts

Figure 3: Where BYOD, embedded, and IoT hosts are allowed

For embedded or IoT hosts, the response was evenly distributed. The exception to both is that a few institutions allow BYOD and embedded or IoT hosts on all network zones.

4.3. Defining a host

The authors were interested in understanding what the average distribution of host types are on an institution's network. Question 3.1 asked respondents to provide their estimated percentage of each of four host categories: embedded devices, servers, institutionally owned end-user devices, and BYOD end-user devices. The results in Table 3 show that, on average, embedded devices (IoT, printers, cameras) and BYOD end-user devices make up half of an institutions network.

Table 3: Host type percentage

Host Type	Min %	Max %	Mean	Std Dev
Embedded devices (IoT, printers, cameras, etc.)	1	25	9.11	6.03
Servers with full operating systems (either physical or virtual)	1	80	15.89	14.3 9
Institution owned end- user devices (desktops, laptops, mobile devices)	4	75	33.56	16.2 8
BYOD end-user devices (desktops, laptops, mobile devices)	0	92	39.44	20.6 4
Other	0	11	0.86	2.57

When asked about the percentage of hosts that use statically assigned IP addresses, all but one respondent said 10 or 20 percent. In addition, the respondents were asked what addressing methods they used. The results are shown in Figure 4.

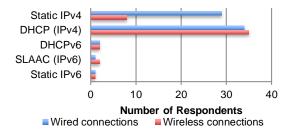


Figure 4: Types of addressing used on wired and wireless connections

Interestingly, eight respondents stated that they used static IPv4 addressing on wireless connections. These could be embedded devices such as printers or copies using wireless however, the authors would expect DHCP Reservations to be used for wireless devices

Questions 3.6 and 3.7 asked how confident the respondent was in identifying unique individual hosts

for virtual machines and application containers. The results are shown in Figure 5.

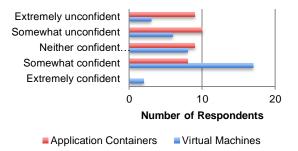
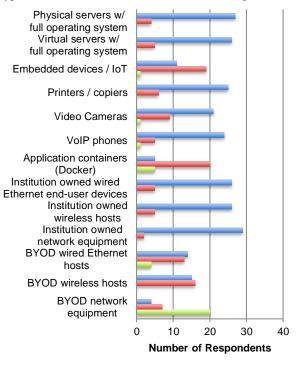


Figure 5: Confidence in identifying virtual machines and application containers

The last question of this section asked how respondents identified a unique host. Most all stated, in their own words, that a MAC address was the unique identifier.

4.4. Evaluation of inventory controls

In this section of the survey, the first question asked respondents to identify whether or not a particular host type was tracked. The results are shown in Figure 6.

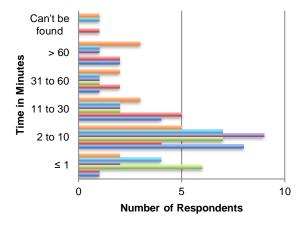


Tracked Not tracked but allowed Not allowed

Figure 6: Host tracking by type

It is worth noting that fewer respondents said BYOD, embedded or IoT, and application containers were tracked. Most respondents tracked physical and virtual servers, VoIP phones, video cameras, printers, and institutionally owned network equipment.

Figure 7 shows the time it takes to track down the physical location of a host for non-research institutions. It is worth noting that a greater number of research institutions (R1, R2, and R3) selected the more than 60 minutes option for multiple host types as shown in Figure 8. This could be due to the larger number of hosts on research institution networks or the distribution of IT responsibility.

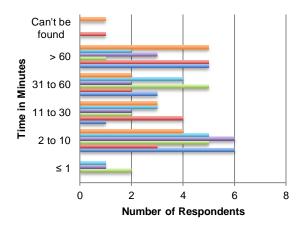


- A wireless end-user host
- A wired end-user host
- Virtual machines and application containers
- A wired Ethernet server with full operating
- system A wireless embedded device, including printers, cameras, and IoT
- A wired embedded device, including printers, cameras, and IoT

Figure 7: Time to find physical location of different host types for non-research institutions

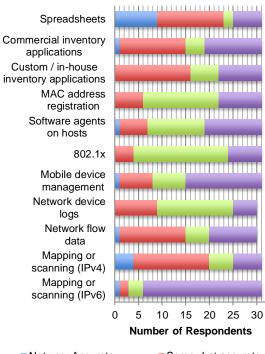
For question 4.4, respondents were asked how often their inventory controls and tools lead to someone who is not the current responsible user and most respondents selected a few times a month. Some wrote in that it varies widely and that it is worse for lab environments.

Question 4.6 asked respondents how accurate they thought various inventory tools were. The results are show in Figure 9.



- A wireless end-user host
- A wired end-user host
- Virtual machines and application containers
- A wired Ethernet server with full operating system
- A wireless embedded device, including printers, cameras, and IoT
- A wired embedded device, including printers, cameras, and IoT

Figure 8: Time to find physical location of different host types for research institutions



Not very Accurate
 Very Accurate

Somewhat accurate
 Not used

Figure 9: Accuracy of inventory tools

It is worth noting that five respondents said that mapping or scanning of IPv6 was somewhat or very accurate. It would be interesting to know their methods given the large address space.

Questions 4.7, 4.8, and 4.9 asked if certain host types were more difficult to track than others as shown in Figure 10.

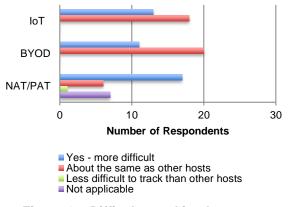
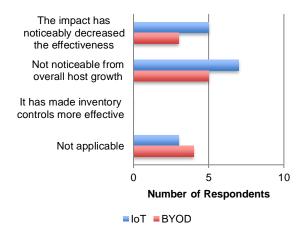
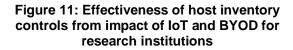


Figure 10: Difficulty tracking host types

The third host type, NAT/PAT, was used to determine whether address translation has an impact on inventory controls. Interestingly, seven respondents selected not applicable for NAT/PAT. The authors surmise that these institutions may have enough IPv4 addresses for all hosts, and therefore have no need for address translation.

Questions 4.11 and 4.12 asked if respondents believe the effectiveness of their host inventory controls changed in the past five years for either BYOD and IoT hosts. As shown in Figure 11, nearly half of the respondents from research institutions stated that both host types impacted the effectiveness of their inventory controls.





Question 4.13 asked respondents how much time they spend updating host inventory control tools. The results, charted in Figure 12, show that more than half of institutions spend a moderate to significant amount of time updating records.

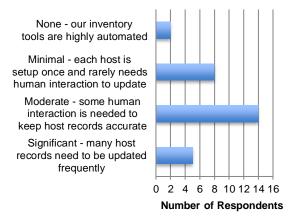


Figure 12: Time spent updating inventory tools

The final question of the survey asked if the respondent had any specific challenges with host inventory controls. A couple of respondents stated that it is difficult to have a unified inventory with a distributed IT responsibility. One respondent also stated that NAT/PAT can be an issue for their DMCA complaints. Another stated that they would like to raise awareness of keeping inventories current and correct.

5. Discussion and insights

It is worth noting that even though the CSC 1 provides methods for inventory, these are corporate enterprise-centric. Even though a higher education institution network may be a special case, the way it works may actually become more common. With BYOD, wireless, and virtualization, the methods of traditional inventory are becoming more difficult to deploy and scale. Specifically with BYOD, corporate networks are allowing more personal devices in their environments [13]. Some will segment their wireless networks; however, there is ever growing pressure for these corporate networks to allow personal devices on their more restrictive network segments.

5.1. Network access

We must consider the user-base as we discuss access and authenticating to a network. In a higher education institution's network, there is an expectation that access to the Internet should be unhindered. This is because faculty and students need to complete their work by collaborating with other higher education institutions and industry partners. To them, the network is only a tool to accomplish this. Additionally, many research institutions have multiple campuses along with faculty and students who are frequently traveling around the world. With this culture, there is usually greater emphasis on controlling access at the applications that are globally available.

This leads to the discussion of private versus public networks. Many higher education institutions operate networks, which could be considered hybrids. For legal reasons, most institutions consider themselves private networks, but discussions have been ongoing ever since The Communications Assistance for Law Enforcement Act (CALEA) and the USA PATRIOT Act have come into existence [14]. Even with this designation, most faculty and students expect open access to the Internet. This is a culture that has been around since the early years of the Internet. Larger higher education networks are traditionally operated like Internet Service Providers (ISPs) where their primary focus is to make sure packets are getting from one host to another. In recent years, some higher education institutions have become more limiting on the free flow of traffic in and out of their networks. Nonetheless, these networks remain much more open than other private networks such as those in corporate environments. This cultural tendency makes requiring high assurance authentication to the network, and ultimately the Internet, a challenge.

5.2. Host attributes

In previous decades, a host might have used the same IPv4 address for long periods of time, sometimes months or even years. The pace at which IT systems are changing is increasing. The life cycle of an individual host has shortened while the expectations of service availability has increased. This leads to redundancy inside a host's subsystems and to redundancy in entire hosts. With redundancy at the host level, the service may change which hosts are responding to requests. This leads to hosts that are dynamic and taken out of service for maintenance or failure. Virtualization furthers this trend of more difficulty in tracking hosts. Virtualization enables the decoupling hosts from hardware, thereby allowing movement. The Media Access Control (MAC) addresses, previously considered relatively static, are now created when new virtual machines (VM) are defined [15]. The ease of creating and moving VMs can be a challenge for traditional host inventory tools.

Some environments are moving to services being deployed in containers by which the operating system or host is considered separate. This leads to even more churn in the traditionally static hosts providing services. For example, Docker is a containerization platform that provides separation of applications from their operating system. Using Linux kernel technology, the containers even have their own network interfaces [16]. These interfaces, like the virtual machines, have their own MAC addresses. Again, this can complicate the issue of how we define a host and what attributes we inventory.

5.3. Host responsibility and organizational inefficiencies

Answering the question of who is responsible is core to host inventory. This can be a difficult problem in a federated research institution network. There can be hosts in which the user is the responsible party, as is the case for BYOD. There are also groups of hosts in which an IT professional is responsible. In some instances the research institution can have both a central IT organization and distributed IT professionals reporting through different leadership. This federated network management model requires more effort to define and track who is responsible for any host. One common method involves the assignment of blocks of addresses to organizational units. The institution assumes that organizational units will track hosts within their assigned block. It is an honor system and can be problematic if the organizational unit has no knowledge of a host using one of its addresses.

Two of the sub controls from CSC 1 are focused on authenticating to the network. If we accept the scenario in which all devices on a network are authenticated, we still have to map the user to a group or responsible IT professional. Again, BYOD comes into play whereby the organization may not have a record of who the device belongs to or who should be contacted if there is an incident involving it.

One last consideration is the time involved in maintaining most host inventories. It is simple to keep the inventory of a twenty-host network up to date. The time it takes to maintain the inventory increases steadily with the number of hosts unless efficient tools are used. Even then, there is significant time spent on updating each host entry. This can be a burden on already busy IT personnel and takes them away from solving more high profile issues. IT professionals can also miscategorize or mistype information. This fundamentally human element makes a tedious tracking process more inaccurate as time goes on.

6. Future work

The number of security incidents occurring within many networks is increasing. The time to detection is not keeping up with the time to compromise as described in the Verizon 2016 Data Breach Investigations Report [17]. This means that we must get better at reducing time to detection and ultimately remediation. By improving a network's host inventory, we can reduce the time to remediation. This is accomplished by quickly determining where a host is and who is responsible.

An accurate host inventory is also a place of record for answering other questions. These include which hosts may need operating system updates and which hosts may be vulnerable to newly announced exploits. This valuable tool goes beyond information security to include understanding how hosts change over time.

If we can leverage automated data flows to populate a host inventory, we can also extend it to become more about crowdsourcing IT security. Presenting users with options and information pertinent to their hosts, we can enable them to make decisions rather than those personnel at the organization level. In time sensitive incidents, this can reduce risk of data exposure by getting the people who know the host the best looking at the problem. This also enables organizational IT security personnel to focus on wide trends and hunt for vulnerabilities. This encourages the philosophy that those closest to the hosts know most about them and security is local.

Much of the information needed to create a dynamic, host inventory with minimal human intervention is already available. The information is in the form of log events which are often left on servers or sent to closed systems for human review. This information should be consolidated and used for more than just ad hoc queries. Correlation of user authentication with host activity has been implemented in higher education institution networks in the form of the Grand Unified Logging Program (GULP) [18]. This system, developed at Columbia University, demonstrates that it is possible to maintain open access to a network and identify responsible users without preregistration or network authentication.

The authors have begun designing a solution that uses network device generated data, such as MAC address to IP address mappings and user authentications to applications, as shown in Figure 13. This design builds on existing solutions and utilizes near real-time data flows.

Now that data analytics has become more the norm, and compute cycles and memory are inexpensive, we can use these resources to mine relevant log events for the right information [19]. Given the right logic, we can piece together what a host is and how it is interacting with the network. This enables us to remove most of the human data entry from the host inventory. It also allows for more timely updates to the inventory and is therefore more accurate at any point in time. This will help solve the problem, identified in question 4.13 of the survey, that most institutions spend at least a moderate amount of time updating host records. It will also reduce the time necessary to identify physical locations of hosts

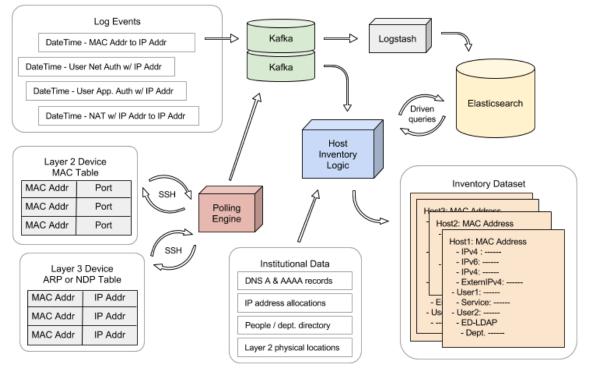


Figure 13: Diagram of a data-driven host inventory system

that is shown in Figure 7 and 8. Lastly, this approach should improve confidence in the ability to track more host types.

The authors are also taking into consideration that any current inventory system needs to accommodate a hundred-thousand hosts or more in a given day with many moving around the network. This can be accomplished using modern, scalable technologies such as clustered message queues, flexible parsing engines, and distributed data stores.

7. Conclusion

More work needs to be done to address the needs of host inventory in higher education, and specifically, research institution networks. The Critical Security Control One provides a high-level goal that every network should strive to achieve. However, the recommended technologies for implementing the control can be difficult for some institutions.

Therefore, a data-driven host inventory system is needed to address the dynamic nature and growth of connected end-user devices. In addition, new classes of hosts, such as IoT, virtual machines, and application containers, are contributing to decreased effectiveness in higher education institutions' abilities to track locations and responsible users. Using real-time log analytics, a data-driven host inventory system can help reverse this trend.

8. References

- P. Cichonski, T. Millar, T. Grance , and K. Scarfone, "Computer security incident handling guide," *NIST Special Publication 800-61* Revision 2, 2012.
- [2] W. Hsieh, C.Wu, and Y. Kao, "A study of android malware detection technology evolution," Security Technology (ICCST), 2015 International Carnahan Conference on, IEEE, 2015.
- [3] L. Johnson, Security Controls Evaluation, Testing, and Assessment Handbook, Elsevier Science, 2015.
- [4] "About Us Center for Internet Security." [Online]. Available: https://www.cisecurity.org/about-us/. [Accessed: 10-May-2016].
- [5] "The CIS Critical Security Controls for Effective Cyber Defense Version 6.1." [Online]. Available: https://www.cisecurity.org/controls/. [Accessed: 20-May-2017].

- [6] G. Lyon, "The Art of Port Scanning." [Online]. Available: https://nmap.org/nmap_doc.html. 1997.
- [7] T. Chown, "RFC 5157 IPv6 Network Scanning", IETF Network Working Group, 2008.
- [8] R. Deraison and R. Gula, T. Hayton, "Passive vulnerability scanning: Introduction to NeVO", Revision 9.1-13, 2003.
- [9] J. Bound, et al. "RFC 3315 Dynamic host configuration protocol for IPv6 (DHCPv6)", IETF Networking Group, 2003.
- [10] S. Thomson, T. Narten, and T. Jinmei, "RFC 4862 Stateless Address Autoconfiguration", IETF Networking Group, 2007.
- [11] R. Marchany, "Higher Education: Open or Secure?," SANS Reading Room, 2014.
- [12] "The Carnegie Classification of Institutions of Higher Education," Indiana University Center for Postsecondary Research (n.d.), 2015.
- [13] K. Miller, J. Voas, and G. Hurlburt, "BYOD: Security and Privacy Considerations," *IT Professional*, vol. 14, no. 5, pp. 53-55, Sept.-Oct., 2012.
- [14] M. Fuller and D. Walker, "Acts Influencing How Higher Education Deals With Information," *Journal of Higher Education Management* 28(1), AAUA, 2013.
- [15] C. Clark, et al. "Live migration of virtual machines", Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation-Volume 2, USENIX Association, 2005.
- [16] D. Merkel, "Docker: lightweight linux containers for consistent development and deployment", *Linux Journal*, 2014.
- [17] "Verizon 2016 Data Breach Investigations Report", Verizon, 2016.
- [18] M. Selsky and D. Medina, "GULP: A Unified Logging Architecture for Authentication Data," Large Installation System Administration Conference, 2005.
- [19] G. Lee, et al. "The unified logging infrastructure for data analytics at Twitter", *Proceedings of the VLDB Endowment*, 2012.