Security Automation Workshop 2014 – Minutes

Disclaimers:

- *The intent of this workshop was to gather together individuals from industry, standards bodies, and government to engage in an open and honest dialogue about Security Automation activities, standards, and technologies. One of the main goals was to gain a shared and more complete understanding of the issues and challenges faced across the the Security Automation community. The individuals that attended the event represented only themselves as subject matter experts with experience in the field and a necessary perspective to share. At no time did they speak as official representatives for their organizations.*
- *Due to the intentional "conversational" environment of this workshop, the terminology used in some cases was not that which would be used in a more formal setting. As such, please know that the items captured in this document were written to reflect more of the idea or intent represented by the words used and not the exact statements made by participants. In many cases, multiple statements are consolidated or summarized into a single thread/item. Lastly, sometimes terms were used to represent a concept and not necessarily a literal idea. In those instances where there was no easily derived phrase to use in lieu of the original phrase, the phrase is presented in quotations. For example, when it was suggested that someone "stand up a server", the intent was to indicate that an authoritative information resource be established and managed.*
- *The action items identified during the minutes were derived from the conversations and were not items that any one individual committed to executing, nor were they agreed to explicitly by the group.*
- *The presenters did not review these notes prior to distribution. The presentations have been summarized by the MITRE staff to capture the important points that were directly associated with the follow on discussion. The summary is not intended to be a comprehensive representation of the slides.*

Contents

Introduction

The Security Automation Workshop 2014 was held from August 26nd to August 28th at MITRE's McLean, VA location. The meeting was jointly organized and facilitated by representatives from the Department of Homeland Security (DHS), the National Security Agency (NSA), and the National Institute for Standards and Technology (NIST). The event focused on the next generation of security automation efforts and standards, including the ongoing efforts in the IETF [1] SACM [2] Working Group.

The registration for the event totaled 69 attendees with an estimated 55 registrants attending each day.

These notes capture the most important aspects from each of the three days of the event. The reader is also encouraged to review the accompanying slides for each of the sections where possible, in order to get a better understanding of each presentation and the ensuing discussion.

Day One – The Context

Current Capabilities - Continuous Monitoring from Government Perspective

Introduction

Three U.S. Government representatives provided insight into the different programs they use to address the Continuous Monitoring challenge. The presenters supplied their top challenges to help drive security tool improvements. The DHS Continuous Diagnostics and Mitigation (CDM) program aims to automate risk posture assessment on federal networks by identifying defects in assets such as devices, software, accounts, etc. These defects get reported through a series of dashboards using SCAP [3] standards where available. The DISA Continuous Monitoring and Risk Scoring (CMRS) program has similar data/information requirements, along with a need to express operational context for all assets. CMRS can provide countermeasures or vulnerability-patch correlations to automate security decisions. Lastly, NIST's National Vulnerability Database (NVD) [4] provides a standardized view of automation reference data. This data includes checklists for configuration guidance, the official Common Platform Enumeration (CPE) dictionary, and Common Vulnerabilities and Exposures (CVE) metadata.

Details

CDM and CMRS aim to provide a single view of all systems on their respective networks. The common issues seen among CMRS and CDM programs involve too much information being collected in different formats, and too much manual interaction. Despite being collected in an automated fashion, controls-based guidance might not have the same impact for one system as it could for another, and likely requires manual tweaking to provide machine readable formats. Furthermore, the variety of sensors deployed do not all report consistent names for software installations. Formats such as the CPE dictionary in NVD are useful but inadequate to provide a standardized name.

Important talking points brought up during the presentations are listed below:

[1] Internet Engineering Task Force (https://www.ietf.org/)
[2] Security Automation and Continuous Monitoring (https://datatracker.ietf.org/wg/sacm/charter)
[3] Security Content Automation Protocol (http://scap.nist.gov/)
[4] National Vulnerability Database(http://nvd.nist.gov/)

- Multiple presenters agreed on the difficulty in correlating different software inventory solutions.
    - One possible solution would be to all use one tool, but this would be impractical given the variety of environments and platforms to evaluate.
- One presenter noted that CDM needs a better way to handle hardware inventory, particularly in identifying endpoints. Once this has been established then "people information" would be included for contextual purposes.
    - "People information" answers questions such as who might be responsible for configuring, defending, or patching a system.
- Several people lamented that there are too few automated STIGs[5] from the Defense Information Systems Agency (DISA) and USGCB[6] baselines from NIST.
- Both presenters explained issues with circumventing information overloading. Due to the volume of information being collected only failed checks are returned.
    - In some cases, these results suffer from the dashboard user not knowing whether the failed check was due to not being applicable or not being checked.
    - The main issue is insufficient metadata about cause of failure.
- Endpoint identification must be greatly improved.
    - Some attendees would like to see an industry standard emerge in this area. Presently some tools assume two IPs are the same endpoint, which may not be the case.
- Despite managing different programs, the different presenters agree that any solutions for CDM should also work for CMRS.
- There was a comment on how the correlation of CVEs to CWEs and to CPEs is time consuming.
    - One attendee wished to see this become crowd-sourced.
- Current software inventory reporting does not account for bundled software/libraries.
    - The example of OpenSSL was provided, where one would need to know all instances where that library was used or had copied code.
        - The purpose for which bundled code is utilized is also useful for determining risk. It could be possible to use something like OpenSSL for encryption but not communication.
    - People were still eager to push the creation of reliable metadata onto the software providers to better understand such bundled or statically linked libraries.
- The NVD has been experiencing a variety of issues relating to generation and distribution of content.
    - Automation guidance and CPE creation are labor intensive processes.
    - Sometimes difficult to link to a patch for a particular vulnerability.
    - Distribution is limited to manually accessing a website to select the correct guidance.
        - Repository protocols could address this.
- One presenter pointed out the lack of tools and editors to deal with these complex formats.
    - Most solutions handle the simple cases but cannot handle advanced ones.
- Some attendees perceive a lack of a professional community (in the domain of security automation standards, tools and practices) for vetting ideas and tuning business logic.

[5] Security Technical Implementation Guides (http://iase.disa.mil/stigs/Pages/index.aspx)
[6] United States Government Configuration Baseline (http://usgcb.nist.gov/)

- Software inventory tools and vulnerability scanners (and associated standards) do not scale well to millions of endpoints. Tools and standards cannot assume that it is feasible to run hundreds of thousands of tests on each endpoint, do so regularly, and transport all that data straightforwardly over the network to a central repository. Scaling issues need to be considered at the very beginning of tool/standard design activities.

Community Action Items

- Research crowd sourcing capability for CVE to CWE mapping.
- Research crowd sourcing capability for CVE to CPE mapping.
- Research better authoring tools for security automation data formats.

Current Capabilities - Continuous Monitoring from Industry Perspective

Introduction

This session focused on commercial industry's view of continuous monitoring, covering the current landscape and the outstanding challenges. The session had three sections, with three different presenters. The three continuous monitoring topics covered conventional endpoints (e.g. servers and workstations), infrastructure endpoints (e.g. switches, routers, etc.), and mobile endpoints (e.g. cell phones).

Discussion was held on various topics throughout the presentations and Q&A sections.

Details

The following high level points were noted:

- The security automation community is poised to take critical next steps in refining extant standards to be more industry- and user-friendly. Current "standards" often reflect "drafts" which became "final" before they were truly ready.
- Some members of industry found the CAESARS[7] architecture to be very helpful in moving the conversation forward.[8] CAESARS indicated that the U.S. Government needs were very similar to the needs of industry's other customers.
- The interaction points or interconnects between different functional capabilities within the current continuous monitoring framework (CAESARS) are poorly defined or missing entirely. This is one of the primary areas of concern to vendors in the space.
- It is important to balance the existing, short-term operational needs with long term standardization efforts like SACM. Also it is important to realize that multiple venues may be appropriate for these efforts.
- Several attendees asked about the line between proprietary, premium content and community shared items, specifically asking how it is determined what content is shared and what content is not. The responses suggested that there is no clear line, but rather would vary from case to case. All seemed to indicate that some sharing was desirable and encouraged. Vendors and

[7] http://csrc.nist.gov/publications/drafts/nistir-7756/Draft-NISTIR-7756_second-public-draft.pdf
[8] See http://www.dhs.gov/continuous-asset-evaluation-situational-awareness-and-risk-scoring-reference-architecture-report.

other commercial entities must strike a careful balance between being good community members and exercising and acting in the best interest of their stakeholders.

- The speakers all suggested that better standards for both protocols and data exchange were needed and were critical to the security automation space as a whole.
- It was noted that many non-U.S. Government customers are happy with SCAP as it is, including international customers.
- One speaker felt that the emerging threat standards (STIX[9], TAXII[10], IODEF[11], etc.) need to be better integrated with other security-related standards to support the full threat life cycle.
- SNMP[12] and NETCONF[13] were offered as potential ways to provide posture attributes for networked devices. Both have limitations, but can provide important information. There were several attendee questions on the specific capabilities of SNMP and NETCONF as well, highlighting some of the limitations, including lack of hardware posture information, assessment capabilities, and lack of policy and remediation features.
    - Generally the attendees agreed that SACM must address infrastructure/networked devices.
- Attendees asked generally about roles and responsibilities with respect to posture assessment for both networked and conventional endpoints, with a key focus on limiting the required effort on the end user. This was a point of conversation across the event and the group generally agreed that ease of use for end users is critical. Additionally, attendees suggested that solutions also need to be clearly communicated and standardized where possible and valuable.
- MDMs[14] are centralized locations for collecting information about mobile devices, but have limitations. Specifically, the devices do not necessarily have to remain configured in the way the MDM dictate and there also needs to be a way to handle unmanaged devices in some cases.
    - Working with MDM vendors is challenging. These vendors don't make use of standards when providing configuration information and are not currently motivated to engage with security automation stakeholders. They also do not necessarily provide information to end users very easily and not all of the information required is available from all MDMs.
    - Several attendees have tried to write content against MDMs, but have found it too challenging to be successful at this point.
    - There are data freshness issues with MDMs—they cannot always tell when a given mobile device attribute was checked, or when an action was taken.
- A large customer organization has looked previously at the challenges associated with assessing the security posture of mobile devices, and found the effort to be very difficult. The effort has since been significantly scaled back due to this.

[9] Structured Threat Information eXpression (https://stix.mitre.org/)
[10] Trusted Automated eXchange of Indicator Information (http://taxii.mitre.org/)
[11] Incident Object Description Exchange Format (http://www.ietf.org/rfc/rfc5070.txt)
[12] Simple Network Management Protocol
[13] Network Configuration Protocol (http://tools.ietf.org/html/rfc4741)
[14] Mobile Device Management (http://en.wikipedia.org/wiki/Mobile_device_management)

- Some attendees asked about the need for the assessment of unmanaged devices. It was pointed out that under FISMA, all devices that touch a network must be understood, including unmanaged devices that are allowed any network access.
- Several attendees mentioned that the key for dealing with mobile devices seems to be the data that it touches and by using proper access controls, one should be able to better manage these devices.
    - One counter to this was that the music and movie industry have struggled with this and should be used as a cautionary tale.
    - The ability to delete data when lost was also mentioned as something that could help.
    - Finally, it was noted that so far all efforts to manage the data with respect to mobile devices has failed.
- Patch management is performed centrally for mobile devices by software that has been recently shown to be insecure. This poses a significant risk as well.
- One attendee added a general comment that historically with the Open Vulnerability and Assessment Language (OVAL) we have tried to get as close to the trusted root of information as possible and that with many of the more recent topics we have pushed further and further away from that trusted root (MDMs being the latest example of this.). This, along with OS fragmentation in the mobile space, makes this harder and harder.

Community Action Items

- Better standardization support for mobile workforce is required to handle changing enterprise landscape.

General SACM Introduction

Introduction

The SACM Working Group (WG) is an IETF WG chartered to develop a set of standards to enable the assessment of endpoint posture. This effort includes standards for interacting with repositories of content related to the assessment of endpoint posture. The SACM WG represents the evolution of the existing SCAP specifications as well as the development of new specifications and protocols to build on the past work by the community to develop international standards that are both scalable and sustainable. This session was intended to provide the attendees with background on the formation of the SACM WG, what the WG is trying to achieve, an update on the WG's activities (past, current, and future), as well as a call for participation by the broader security automation community.

Details

- A brief history of how the SACM WG came into existence was presented. As the security automation community matured, it became clear that SCAP needed to evolve and move into an international standards organization to increase adoption and guard against the development of multiple competing standards and ended with the chartering of the WG in 2013.
- The SACM WG has involvement from a number of organizations from both industry and government including Avaya, Cisco, DHS, Juniper, MITRE, NICT, NIST, Oracle, ThreatGuard, Tripwire, US CERT, and Goldman Sachs among others.
- The WG is expected to develop an informational architecture document as well as standards track information models and data formats for configuration and policy information, driving collection and analysis, and expressing posture information.

    - The terminology and use cases documents have been adopted by the WG. The terminology document will likely continue to get revised, but, the use cases document is nearly ready to be submitted to the Internet Engineering Steering Group (IESG).
    - The WG is also working on five individual submissions: (1) requirements document, (2) architecture document, (3) Extensible Messaging and Presence Protocol (XMPP[15]) Extensions for Use in SACM Information Transport (XMPP-Grid), (4) SACM Information Model Based on the Trusted Network Connect (TNC), and (5) the Information Model for Endpoint Assessment standards track documents that are currently undergoing further development.

- One attendee asked for clarification on the difference between an IETF informational draft and standards track document for those in the audience that may be unfamiliar with documents in the IETF. Informational drafts are meant to assist the reader in understanding why something is being done or to examine something in greater detail whereas standards track documents define specifications and protocols that the WG wants to standardize.
- An abstract view of the SACM use cases was presented which breaks the work down into five components.
    - Express: the ability to define and publish collection and evaluation guidance as well as the ability to query and retrieve posture attribute data from endpoints in an enterprise.
    - Scope: the ability to discover, characterize, and target endpoints in an enterprise.
    - Collect: the ability to acquire and collect posture attribute data from an endpoint.
    - Evaluate: the ability to acquire and query evaluation guidance, detect changes in the posture attribute data for an endpoint, and evaluate the current state of an endpoint against some evaluation guidance.
    - Common Communication Infrastructure: the ability to transmit requests and data in a standardized way facilitating transport interoperability. This area hasn't been addressed in SCAP, but needs to be addressed in the future.
- One attendee asked if there was an obligation or requirement in the charter to provide enough information to support standardized evaluation results or if the WG was bound to work within the Express, Scope, Collect, and Evaluate components. The speaker explained how it was still to be determined whether or not this work is in scope for the charter, but believes that standardized results is implicitly required.
- Another attendee asked how the SACM group planned to authenticate participants that are sending and receiving information. It was explained that XMPP-Grid includes PKI and manages authorizations. It was also noted that there will be a need to carry out certificate revocations. While those details haven't been worked out yet, it was pointed out that these things have been well thought out in other domains and should reference existing specifications where possible and only develop new ones where there are gaps.
- A speculative timeline was provided for the WG's activities moving forward.
    - 2014: finalize the requirements document and adopt the architecture and information model documents.

[15] http://xmpp.org/

    - 2015-2016: develop specific data models and architectural interfaces.
    - 2017: start the next iteration of the security automation work (e.g. remediation, etc.).
- A question was asked about whether or not proprietary protocols were supported in addition to the protocols that SACM may mandate. It was explained that the WG will have protocols that are mandatory to implement, but that would not prevent vendors from using other transport mechanisms beyond what is mandated. Another attendee explained that the IETF doesn't mandate anything, support of a particular RFC defined by IETF is up to individual organizations.
- Another attendee explained that SCAP vendors already have ways to transport data and that we want to support interoperability among different vendors. Given this, there needs to be a suitable transition for vendor products.
- One attendee noted that the security automation efforts need to evolve organically in the community (SACM is this place to do this) to keep pace with the rapid changes in the security space. Community participation is critical to ensure that implementation is feasible and that the needs of the organizations are satisfied. The security automation community cannot have the U.S. Government telling industry what standards they need to develop and support. An attendee asked why "Define" was in the Express component of the abstract view of the use cases. It was explained that organizations need to define the information that they want to express in the guidance that products will be able to leverage.

Community Action Items

- Review working group documents to determine if they fit the needs of your organization[16]
- Join and participate in the SACM WG e-mail list[17] and attend meetings (face-to-face meetings, virtual interim meetings, etc.)

Other Standards Efforts

Introduction

The primary focus of these sessions was to provide background on the Endpoint Compliance Profile (ECP), Network Endpoint Assessment (NEA), Interface to Metadata Access Points (IF-MAP) and the Extensible Messaging and Presence Protocol (XMPP) efforts to the broader security automation community. Furthermore, the sessions were intended to show how these other standards efforts complemented existing SCAP standards, and to highlight any points of contention so that SACM can learn from the development of these efforts.

During these sections, a good deal of discussion occurred. The following captures the most important topics discussed for each section:

Endpoint Compliance Profile

Details

- Under the aegis of the Trusted Computing Group (TCG), the Trusted Network Connect (TNC) WG developed ECP which is a set of schemas and protocols that help one discover what endpoints

[16] https://datatracker.ietf.org/wg/sacm/documents/
[17] https://www.ietf.org/mailman/listinfo/sacm

are on a network, what those endpoints are running, and whether or not they are compliant with the specified policy.

- Vulnerability alerts are published by software vendors and patches are released, but determining what software is running on the vast range of endpoints on an organization's network is still a very difficult task. Often times, system administrators don't know all the software that is running on endpoints or even all the endpoints on the network.
- Software identification[18] (SWID) is an XML-based standard for expressing information about software (name, publisher, version, patch level, etc.) and can help towards enabling system administrators to know what software is installed on an endpoint because they provide a common format for this information as well as documented locations where they can be found on an endpoint. (NB: the SWID tag standard has been developed separately within the International Standards Organization, as ISO/IEC 19770-2.)
- The TNC architecture provides a standardized way of transferring SWID data so that it can be used to do things like storing SWIDs in a Configuration Management Database (CMDB), determine compliance with the specified policy, and identify vulnerable software on an endpoint among other things.
- ECP relies heavily on a secure device identifier. The TCG is currently working to define what a secure device identifier would be.
- The presenter gave a quick overview of the TNC architecture and its different components.
- One attendee explained how many of the standards produced by the TCG have been submitted to the IETF and then updated to reflect the changes from the international community.
- The presenter gave a brief overview of the TNC interfaces as well as how the different messages are encapsulated in other messages for people who want a basic understanding of the low-level details.
- One attendee asked if SWIDs were adequate for identifying all software, vulnerabilities, and misconfigurations in software. It was explained how SWIDs do not solve all problems related to software, but can identify software installed on an endpoint, operating system, patch and version numbers, etc.
- Another attendee asked if registration was required for SWIDs. It was explained that registration isn't required and that an organization is in control of the SWID tags in their enterprise (e.g. can create your own, etc.), but tools will need to put them in the correct locations according to the specification, which may vary by platform.
- If an organization is creating their own SWID tags, there will be some maintenance that one must do, but SWIDs that are created and managed by software publishers are required to be installed and uninstalled from the endpoint along with the software.
- It was noted that the SWID community is working with industry to collect feedback, and is modifying the centralized locations for SWID tags because it doesn't work well for some sandboxed software (e.g. software that came from an app store, etc.).
- TagVault is looking to get SWIDs adopted in incremental steps with trying to get software vendors to start publishing SWIDs for their new software and then SWIDs for old software will be handled over time.

[18] http://tagvault.org/swid-tags/

- One attendee noted that one of the attractive things about SWID tags is that it is an artifact that can be built during the software build where most of the relevant knowledge is present. SWIDs also address some technical challenges present in CPE such as ambiguity between marketing revision and technical revision and allowing patches to become first class citizens which can help support some of the offline analysis processes.
- A couple of attendees discussed how the SACM community needs to address virtual environments and similar things like Linux containers as well as the network protocols to work with them.
- The presenter mentioned that it will be important to have a secure endpoint identifier (TPM, hardware, etc.) and then being able to map observed endpoint identifiers to it.
- Another attendee explained that though SACM is currently working on the architecture document, it hasn't picked specific protocols yet, despite the existence of some proposals. What protocols that get selected will be dependent on further discussion in the WG.
- Another attendee noted that different organizations have different network setups and that SACM will need to accommodate multiple approaches.

Community Action Items

- Determine what level of interoperability primary source vendors are willing to support in order to help the community develop standards that are both scalable and sustainable.
- What are the potential issues and considerations that arise with virtualized environments that SACM needs to be aware of and take into account?
- What barriers to adoption do primary source vendors face and what can be done to eliminate them? What challenges do primary source vendors face when trying to integrate SWIDs into their build processes?
- Brainstorm scalable and sustainable approaches for addressing the generation of SWIDs for software that may no longer be supported by a primary source vendor or published by a primary source vendor that chooses not to adopt SWIDs.
- Have authoritative software vendors "stand up a server" of all their SWID tags representing all the software that they provide allowing programs like NVD to download that information and use it to augment their vulnerability analysis processes and build knowledge repositories of what files belong with what software enabling the creation of vulnerable product mappings.

Network Endpoint Assessment

Details

- The IETF has previously tackled posture attributes and how to collect them in the NEA WG, which just concluded.
- The presenter briefly explained how certain specifications and interfaces have been standardized by the NEA WG and how there are two bindings for IF-T[19] (TLS and EAP). It was noted that EAP can do the posture assessment at the time of connection whereas TLS requires the endpoint to be connected first.

[19] http://www.trustedcomputinggroup.org/resources/tnc_ift_binding_to_tls

- It was also explained how much of this information collected from the endpoints is sensitive and would need to be stored and transmitted in a secure manner and that there are many existing protocols such as TLS that allow you to transmit this data securely. Clients on the endpoint are critical to securely getting this information off the endpoint.
- A couple of attendees agreed that it would be beneficial to try and re-use existing infrastructure to transport SCAP data rather than creating another infrastructure. It was also noted that this was something that SACM is trying to address and would like to get to a point where organizations only need a single infrastructure to support their management, incident response, and compliance activities. This might be achieved through proxies that utilize the standardized formats and interfaces that are developed in SACM.
- One attendee highlighted the need for the community to determine the minimum acceptable level of interoperability to encourage primary source vendor participation; understanding that primary source vendor participation is critical.
- An attendee asked whether or not there were any special considerations for endpoints operating on other networks that they don't own or endpoints accessing network resources remotely. Another attendee replied explaining how TNC is extensible and can be used as long as components such as policy server or a policy decision point are available for the endpoint to connect to. If so, one could do the TNC interfaces regardless of where the endpoints physically sits, acknowledging that the infrastructure is required.
- Another attendee noted that SACM must address unmanaged endpoints within the architecture, because in some cases those endpoints will access enterprise resources.

Community Action Items

- It would be beneficial to document the assumptions about who owns and manages components in the TNC architecture and the requirements so that SACM can determine where they agree and where they don't.

XMPP Introduction

Details

This presentation introduced attendees to the XMPP standard. This protocol has been identified as one that could address the control plane requirement within the SACM architecture with respect to endpoint communications. XMPP supports out-of-band PKI and certificate based trusted connections, flexible APIs, and is highly scalable. Being platform agnostic, this standard is applicable to the wide range of endpoints that can be expected of continuous monitoring efforts. XMPP utilizes the concept of a grid controller to arbitrate authentication and communications.

There were several discussions that took place during this presentation that clarified the capabilities of XMPP. Some high level topics are as follows:

- All connecting clients need to be registered to communicate with the grid controller.
    - This could be used for unique endpoint identification.
    - Policies can further define authorization and subscriptions for updates.
- The standard offers real-time updates through subscription notifications.
- Out-of-band communications occurs through peer-to-peer directed queries.
- The ability to create grid subtopics allow for highly targeted data queries.

    - The format supports information such as location and domain.
    - These queries can be filtered through subtopics.
- XMPP can be automatically substituted with IF-MAP standard interfaces or work off an IF-MAP enabled network to build off other solutions.
- It supports large bandwidth communications.
- SACM can help identify what subtopics should be defined.
- One attendee raised a concern about XMPP being associated with a messaging protocol. The speaker assured that person that messaging is one application of XMPP.

Endpoints and Architecture

Introduction

The IETF SACM WG has created a document that defines the overall architecture that will fulfill the requirements that have been documented and agreed upon by the group. The document defines high level components that need exist to achieve the goals of the effort. This includes Posture Assessment Information Producers and Consumers, a Management Plane, and Interfaces that connect these components.

During this section a brief, high level overview of the architecture defined by SACM currently was provided.

This overview was provided in order to give the audience a shared context for the work going on in the IETF SACM WG. Limited discussion occurred during this section, though some relevant points were made:

- The speaker noted that the term "Management Plane" was used instead of the original proposed "Control Plane" for no reason other than preference expressed by some community members.
- There was some discussion as to how well the defined architecture supports more complicated use cases. Generally the attendees agreed that the architecture accurately reflects the way simple assessment works, but there was some concern that revision might be necessary in the future to handle these more complicated scenarios.

Open Conversation

Introduction

At the end of the first day, there was a loosely structured open conversation intended to focus more on some of the specific themes that had been brought up during the course of the day.

Details

The following is a summary of the conversations:

- It was generally agreed that both a "publish/notify" model (where posture attributes are published only as needed or when attributes change) and a "collect all" model (where all required attributes are collected on a periodic basis) are required. One attendee also suggested that broadcast and/or un-authenticated data be considered for collection as well.
- The topic of Endpoint Identification was also discussed. There was general consensus that this was one of the most important topics of the event.

    - The central question is whether or not it is possible to come up with a truly unique identifier for endpoints. There is no obvious answer. This is an area for more research.
    - The Internet of Things[20] is an area that needs to be considered when discussing endpoint identification.
    - The ability to (in a standard way) generate an identifier for an endpoint dynamically was generally determined to be important.
    - There is existing effort around discovery that should be leveraged where possible.
    - How do we assign names to new things we discover that aren't already named?
- Crowd-sourced content and related Quality Assurance (QA) was also discussed. Generally the attendees agreed that content quality currently is a significant challenge, especially with free and open content.
    - Most believe that a validation type function must exist for high quality content to be feasible.
    - Some potential issues with crowd-sourced issues were mentioned:
        - Trust issues – How to convey provenance information about content
        - Lack of tools – Current methods for crowd-sourcing content are email-based, which breaks down
        - Better ability to share content is needed
    - Some attendees feel that crowd-sourcing content will not work and that it must be more directly paid for.
    - One general theme that emerged during the conversation is that many non-technical issues exist with respect to sharing content across boundaries.
    - One attendee suggested that a central repository for SWID tags would be very helpful.
- Another significant topic during the open discussion was posture attributes, including what types of attributes to collect, how to collect them, and how to associate attributes with endpoints.
    - The general consensus was not to worry about whether attributes were 'security-related' or not until such time as it is important to make such a decision.
    - Several categories of attributes were identified, including characterization attributes, targeting attributes, and configuration attributes. While some overlap exists, generally the attribute types are:
        - Configuration – those attributes that drive settings that can change the operation of the asset in question
        - Characterization – those attributes that provide context that can be relevant to identifying the appropriate posture guidance of the asset in question
        - Targeting – those attributes that can be used to determine things like applicability and used to target an asset for assessment
    - The topic of Endpoint Identification needs to be solved to allow for the proper collection of endpoint posture attributes.
        - A related and important topic is the ability to reconcile multiple, duplicate endpoint identifiers. This is a topic that requires additional research.

[20] http://en.wikipedia.org/wiki/Internet_of_Things

    - All seemed to agree that while it is difficult and possibly impossible to identify endpoints in all cases, it is still very important to collect a great deal of endpoint posture attributes.
- Some attendees brought up how existing SCAP will fit into IETF SACM. While some of the SCAP components may be of value to the SACM work, it is too early to know precisely how and if things will fit into both the Architecture and the Information and Data Models.

Community Action Items

- Consider creation of prototype "repository" for SWID tags to include identification of mechanisms to discover, "feed", and "consume" this information.

Day Two – Software Management

A CIOs Perspective on Software Management

Introduction

During this session one of the organizers of the event gave some insight on the perspective of CIOs, garnered from publicly available talks. This was provided as an aggregation of views from various interviews and presentations rather than one single speaker's opinions. Additionally, several shortcomings were identified with some existing solutions.

Details

The chief points of concern for the typical CIO are due diligence, regulatory compliance, and fine avoidance. The details about how software licenses and inventories are managed support those concerns. Better understanding of both software inventory and software licensing is important. Software inventory can encourage healthy networks by utilizing whitelists to block specific application activity. For the purposes of this talk, malware is considered unwanted software.

The following talking points highlighted by the presenter were discussed:

- Proper software license management has the capacity to save money.
    - It also ensures no unlicensed software is being executed on endpoints.
- Accurate software inventories have many security-related benefits.
    - The correlation of software inventory to vulnerability information such as CVE can indicate which endpoints are unpatched.
        - One attendee wished to clarify that CVEs are a subset of vulnerability information that can correlate to what is vulnerable software.
    - Collecting the correct information is difficult for open source software instances, where a patch may be applied but the tool may still be reporting the previous version. This would provide a false positive for vulnerable software when evaluated using current methods.
    - The mapping of software inventory to specific processes on endpoints could clarify which prohibited versions are actually running rather than just what is installed.
- Existing software inventory tools are mostly developed by 3rd party vendors. By not being provided by the primary source vendor, inaccurate information can be collected.

    - One attendee discussed the level of effort required to re-purpose existing sources in an automated fashion.
        - IAVM[21] automated mappings and Microsoft security bulletin spreadsheets do not contain all the required information for creating content to check for vulnerable software in a comprehensive or automated manner. .
        - Severity ratings from Microsoft bulletins are for the full bulletin rather than individual software pieces.
    - For Windows platforms, registry scraping can only get some required information; not all information may be available.
- Endpoint identification is a hard problem.
    - One is required to report on all endpoints that touch the network, not just ones owned by an organization.
    - This problem is further complicated with new implementations of BYOD[22].
        - Some licenses for software on these endpoints could span commercial and non-commercial usage.

Community Action Items

- Push for primary source vendors to provide detailed patch and version information that enables accurate identification of known vulnerabilities using software inventory information.
- Look to improve mapping vulnerability and/or mitigation detection information with existing sources used by NVD.

Software Identification and Inventory – SWID Tags

Introduction

The current landscape of software identification was discussed during this presentation. Additionally, information was provided on the way SWIDs could alleviate some concerns with current solutions. The presenter talked about both challenges with usage of SWIDs and asked several questions of the audience.

Details

No commercially available tools currently perform software inventory in a satisfactory manner. It is inherently a hard problem to discover what software is installed on an endpoint. The SWID standard aims to provide common data and data structures for usage, with non-profit TagVault.org actively encouraging new organizations to adopt it. Correct software identification becomes difficult as different tools collect in different formats, with no proven method of reconciliation. One hope for the expanded adoption of SWIDs is to appeal to the CIO's business case of saving money. Many organizational processes require some form of software inventory such as patch or policy management, licensing, and virus scanners. Some observed hurdles with adoption are from cultural perceptions about the standards body offering the solution. ISO has typically provided access to their standards for a fee, however one may implement the SWID standard without needing to purchase it. Furthermore progress is being made to address SWID interoperability requirements in a form that will be made publically available.

[21] Information Assurance Vulnerability Management
[22] Bring your own devices

Some discussions and opinions on SWIDs that occurred during the presentation are outlined below:

- Most attendees agree that the best case scenario would be for software producers to provide their own SWIDs.
    - This is required for scalability and also allows the most authoritative source to provide this information.
    - When a software publisher provides SWID tags with their software, it decreases the value of other vendor's proprietary software catalogs and libraries.
- The presenter noted that having a standard created means nothing. It is up to the producers to adopt and make that standard viable.
- There has been difficulty in providing SWIDs for platform or browser specific app stores. These may not have the correct permissions to write to a common directory where a SWID may be queried. Newer drafts of the SWID standard have changed this expected behavior to allow an application to provide its own SWID directly to support app stores and other virtualized software.
- Several questions of trust were raised, specifically voicing concern that SWID tags could be subject to tampering.
    - SWID tags provide integrity against tampering through a combination of a digital signature and a hash of the footprint embedded within a "media" tag.
- The presenter raised the question of what could be done to cause software vendors to provide their own SWID tags for identification.
    - U.S. Government sponsors could mandate that software purchased must conform to this practice, but would be difficult in the current environment.

Software Identification and Inventory – Lifecycle of SWID - MS

Introduction

The presenter's perspective on the lifecycle of SWID tags in Microsoft products was presented. The presentation began with an overview of the problem as seen from the presenter's perspective. Information about how SWIDs are currently used in Microsoft products and how SWIDs could be used in the future was also provided.

During and following this presentation a series of questions and related conversations occurred.

Details

A variety of customers both commercial and government (not limited to the U.S. Government) are asking for better ways to identify and trust the software on their endpoints. They are also concerned about malware.

Currently Microsoft supplies SWID tags to begin helping with the software inventory problem, including some very simple tags in Windows Server 2012, Windows 8, Office 15, and VisualStudio. There is a proposal to make SWIDs a formal part of the centralized software release process and to add additional details to the SWID including installation details, more information regarding dependencies and patch information, installation source, and package footprint.

Attendees had a number of questions and other thoughts on the subject, which are summarized below:

- Software installed on top of Microsoft OSs was a popular topic. Many attendees had questions about how the installation of this software could be integrated more closely with the OS and SWID tags.
    - All seemed to like the idea of influencing installation software vendors to include a SWID tag for installed software in the case that the vendor themselves did not see fit to include one.
        - Some attendees stated concerns that 3rd party SWID tags would not be able to generate truly unique identifiers for the software. While the SWID ID is intended to be globally unique, any time 3rd party SWID tags are used, the possibility for clashes exists.
    - The strongly preferred method of getting SWIDs in place is for the primary source vendors to provide them themselves, as it will prove more authoritative, trusted, and accurate than 3rd party efforts.
    - One attendee pointed out that the top software vendors account for a large amount of the most commonly used software products. He gave the example that were the top 20 vendors (according to CVE data) to provide SWID tags with their software, it would account for around 45% of all relevant software.
- One idea that came up several times was how to handle legacy OSs and other software. OSs like Windows XP and others are still in use in some enterprises and need to be accounted for.
    - Some attendees suggested that one could take already collected information about legacy software (presumably without SWID tags) and generating SWID tags for those pieces of software. The app store could be used to collect information for this purpose, with respect to mobile devices.
- Some asked about the cloud and whether specific care should be given to that use case. In general, attendees agreed that the cloud case needed to be considered, but to focus first on traditional installs.
- Another topic discussed throughout the section was the use of APIs to collect information regarding software inventory. It was suggested that one option would be to allow the OS to keep native solutions (like rpm for Linux or MSI database for Windows) and provide APIs that could reply using SWID tags or other standardized data formats to provide inventory data.
- One attendee pointed out that another issue that needs to be solved is how to handle features (either for an OS or other software) that can be "turned on" after install. The general consensus here was to solve install-time inventory first and then consider how to address this more advanced case.
- There was an overall concern about 1st vs. 3rd party SWID tags and the possibility of double counting software due to unique identifier issues or other problems that arise with multiple sources of data. The group recognizes this as an area of research.

Community Action Items

- Research and prototype the ability to create SWID tags for software published by non-adopters of the SWID tag standard. Several potential data sources were mentioned.
- Research the feasibility of unique software ids, considering both the creation of such identifiers a priori, as well as the dynamic creation of identifiers by automated tools. Another related topic is the ability to reconcile duplicate ids for endpoints.

Software Identification and Inventory – Lifecycle of SWID - Linux

Introduction

This section began with an overview of how Red Hat handles software inventory, including how rpm works and what the contents of the rpm database. Some shortcomings of rpm and the other Linux tools were also highlighted.

Following and during the presentation, several attendees asked questions or made comments. The conversation is summarized below.

Details

The rpm database and related tools can provide a great deal of the information required to do Software Inventory. The presenter reviewed how rpm works, including providing details on what information is found within the rpm database. It was also highlighted that the rpm database binary file is already around 300KB in size as is. Additionally, he spoke about some of the shortcomings of rpm and its related tools, including the possibility for duplicate information, issues with Virt[23], etc.

Finally, the presenter shared some concerns with SWID adoption/investment, noting the current indicators of community involvement like public mailing lists and forum interest are lacking. For example, a SWID-focused e-mail list established by the NIST NCCoE[24] (National Cyber Security Center of Excellence) has seen minimal traffic. TagVault posts only occasional updates, and the recent revision process for the ISO/IEC 19770-2 standard has not been transparent. SCAP and SACM e-mail lists are essentially silent on the role and status of SWID tags within the security automation community. The fear that SWID may not succeed in the marketplace results in a hesitation to commit to fully supporting the standard. Additionally, the extra data storage required for SWID tags could result in a significant increase in the size of the inventory information (as much as 1GB could be required).

The following questions and comments both following and during the presentation were made:

- How to handle the support of SWID was discussed at length. Several attendees asked about how Red Hat could/would support the usage of SWID. The overall answer seemed to be that they would continue to store inventory information natively in rpm, and provide an API that could return that information in SWID format.
    - In the span of an hour during lunch, one attendee was able to write some rough code that accomplished this task.
    - The rpm database seems to already encapsulate the necessary data fields for supporting SWID.
- Another concern is that not all software published for Linux uses rpm. The feeling was the enough software uses rpm such that the capability to get SWID-formatted inventory information would be a great start, despite the issue with software that does not use rpm.
- It was noted that rpm is non-interoperable with the native package management systems deployed on other Linux distributions.
- One attendee pointed out that Strongswan[25] has a tool for generating SWID tags already.

[23] http://virt-tools.org/
[24] National Cybersecurity Center Of Excellence (http://nccoe.nist.gov/)
[25] https://www.strongswan.org/

- Another attendee shared that his company was considering halting its use of rpm due to complexity issues. It was far easier for them to simply use a JAR file to deploy their application across platforms. A JAR to SWID generator would be helpful in cases like this, but one does not exist as far as the attendees knew.
- Whether SWID is 'real' or not generated a good deal of conversation. The overall theme was that corporations like Red Hat can be hesitant to support efforts like SWID adoption without more visible evidence that the standard is 'alive and vital'.
    - The NCCoE was mentioned a few times during this conversation. It was suggested that the stakeholders within Security Automation use the NCCoE as a forum to help drive adoption of things like SWID.
    - Several ways were suggested to better demonstrate growth of SWID adoption and interest:
        - A non-profit group could be funded to help drive this effort.
        - The U.S. Government could invest in forming and organizing a community, and could play an evangelizing role complementary to industry. U.S. Government direction-setting is a powerful form of marketing. U.S. Government requirements and plans/intentions need to be documented and widely shared.
        - Identify compelling features of SWID and market these features in a way that naturally creates interest and adoption without a further driving force, assuming such features exist.
- One attendee noted that the recent OpenSSL vulnerabilities[26] were very challenging to assess in an enterprise because patched vs. unpatched versions of the software were not immediately evident simply by knowing an application's version and patch level. While OVAL content was created so that it could provide that information by the OS vendor, the use of OVAL is seen as a very heavyweight solution for a large enterprise—it doesn't scale to millions of endpoints.

Community Action Items

- Create a JAR to SWID tag generator to help aid in dealing with software that is published as a JAR file only.

Software Identification and Inventory – Data Repository and its Interface

Introduction

The Data Repository and its Interfaces session focused on the various considerations and challenges that need to be thought through when considering the establishment of a repository to store a wide variety of posture attribute data. The session first highlighted the types of information that might be stored in such a repository and how the repository might be used. It also discussed the current options and challenges associated with endpoint assessment, what applicability statements are, how they are used, and the current applicability language situation along with the new vision. Lastly, the challenges with standardizing a data repository were discussed.

[26] https://www.openssl.org/news/vulnerabilities.html

Details

- The presenter explained how the goal of security automation is to reduce the attack surface on endpoints, and to do this, it requires collecting lots of information about an endpoint which can then be evaluated against guidance to drive mitigation and remediation decisions. The presenter also provided an overview of how endpoint assessment is performed today and the shortcomings of not having a standardized solution.
- The presenter then explained the need for an applicability language and provided multiple examples of where applicability statements are currently used.
    - Common Vulnerability Reporting Format (CVRF): source for identifying vulnerable software from an endpoint's software inventory.
    - National Vulnerability Database (NVD): source for identifying vulnerable software from an endpoint's software inventory. NVD uses CPE for its applicability statements.
    - Extensible Configuration Checklist Description Format (XCCDF): source of automated identification of targets based on software inventory.
- It should be noted that all of these examples need to figure out how to incorporate SWID tags.
- The presenter then discussed the state of our current applicability language CPE 2.3.
    - It is used in XCCDF and NVD.
    - It would not be easy to update CPE 2.3 so we are looking to develop a new applicability language.
- One attendee asked why it was decided not to use CPE in favor of SWIDs.
    - There are problems with CPE such as the centralization problem and is constrained to expressing only 11 attributes.
    - It is up to vendors to document what data goes into what fields which can result in inconsistencies in the content.
    - SWID provides a much richer set of data, is extensible and better for cataloging software, and is generated by vendors.
    - Primary source vendors are already involved with SWIDs so it makes sense to use it rather than fixing CPE and then trying to gain vendor support.
    - CPE mixes software IDs, matching, and metadata all together whereas SWID does not.
- Multiple attendees agreed that applicability statements are critical to targeting and could be used to determine which evaluation guidance applies to an endpoint, what data needs to be collected form an endpoint, and what information needs to be reported back.
- Next, the challenges associated with standardizing on a data repository were discussed.
- One attendee commented on whether or not SACM was ready for this work and encouraged that it be brought to the list, a draft be written, and then the community can start discussing it and provide feedback. The attendee also provided feedback that they don't agree with the statement that an applicability statement is a query.
- Another attendee explained how there are three types of data they may want to store.
    - Identifiers of an endpoint.
    - Data that characterizes an endpoint from business perspective.
    - Posture attribute data that describes the state of an endpoint.
- One attendee mentioned that they were not aware that all of the SCAP work was being done on an endpoint. Another attendee explained that this is not the case and vendors could make things like proprietary data stores. It is just that the problem is not standardized or extensible.

- An attendee noted that to achieve true interoperability, you will need to define the data formats, interfaces that define operations, and the transport protocols for exchanging the information.
- One attendee felt that the focus shouldn't be on standardizing a repository, but rather, how to get the data out and exchange it back and forth. They also noted that they believe the ability to query is more important than having data structured in standardized fashion.
- Another attendee commented that it shouldn't matter whether or not they are querying an endpoint or some central repository as long as it is the fastest way to retrieve the data. Vendors should be able to query endpoints, and information that they need, in any way that they want.
- An attendee stressed the idea that they want to be able to collect the data once and allow everyone to use it. Another attendee stated that this could already be done using best practices and proprietary solutions.
- Another attendee explained how it is important to make sure that standards are only being developed to solve the correct technical problems and not being developed to solve political and organizational issues.
- One attendee noted that developing assertions in a machine readable fashion has been solved a bunch of times and that we should pick one of those. They also noted that, with respect to the repository, it is very important that the minimal set of attributes are selected for querying and evaluation.
- An attendee suggested that agencies should be surveyed to determine how well they understand the existing specifications before new ones are created. Otherwise, the agencies may not be able to understand the new specifications.
- A couple of attendees believed that this is less about an actual database and more about exposing a minimum set of attributes using standardized queries and interfaces.
- Another attendee expressed the need for SACM to define the interfaces between the collection, orchestration, and aggregation subsystems. This emphasizes the need for a tasking language.
- One attendee offered an XMPP-Grid perspective where a capability provider would state its capabilities and provide a schema of attributes it is exposing along with specific ways to invoke those attributes through queries. It is important to note that how the data is stored is implementation specific and it is more about interfaces and the attributes that you want to expose.
- An attendee asked whether an interface was equal to a schema, a set of protocols, and data models which led to the discussion of what is an interface.
    - One attendee responded that an interface is a set of protocols, data models, and any command instructions necessary to retrieve or connect.
    - Another attendee defined interfaces as being the operations you would want to take on a given set of data.
    - Another attendee discussed the order in which they care about the different parts of an interface.
        1. Common data format (payload)
        2. Operations
        3. Protocol
- One attendee expressed that how information is expressed and transported should be decoupled. The vocabulary should also be separated from the expression. The attendee also encouraged that the community should start with the smallest vocabulary before getting to edge cases because simplicity drives adoption and lets you understand the political and technical issues that may not be

seen in advance rather than trying to make the perfect solution that covers everything and then find out it doesn't work. Another attendee suggested that data types also need to be defined.

- An attendee explained how encodings, date and time, and knowing if the data has changed are all key problems when working with data.
- One attendee mentioned how information about the perspective of the data needs to be collected as different sensors will see things differently and it is not always clear why things have happened. For example, was the data not reported because the tool did not see it, the tool did not look for it, or because the data was not there. This further emphasizes the need for the community to select a minimum set of attributes to support.
- One attendee suggested the following priorities for different efforts discussed at the workshop.
    - Software inventory.
    - Guidance (complexity, cost, quantity).
    - Getting out of defining settings at the government level.
    - Creating accurate content to get more accurate results.
- One attendee explained how NIST's Cyber Security Framework helped them understand risk and discover inconsistencies and unknowns that would have otherwise gone undiscovered.
- One attendee explained how control substantiation is a different problem than whether a control manages your risk. While they are related, there are different questions that you need to ask to see if the controls are effective.

Community Action Items

- Propose ideas to SACM (or to another appropriate community) regarding a data repository standard. This may come in the form of informational drafts. A few ideas of things to discuss include:
    - Set of things needed to identify an endpoint.
    - Set of things to contextualize the endpoint.
    - Set of minimal posture attributes that we care about.
- Determine requirements for an applicability language and enumerate existing languages that could satisfy those requirements.
- Survey government agencies to determine how well they understand the current security automation specifications to see if we can make any improvements and develop more accessible specifications moving forward.
- Define what the security automation community means by interface.

Working Session

Introduction

The final session of the day addressed several significant, open questions that came up over the course of the day. One of the event organizers led a loosely structured conversation on the topic of software inventory, highlighting specific questions.

Following the brief introduction, the group discussed specific aspects of the topics.

Details

The organizer opened by noting that the group agreed that software inventory is the most significant broad challenge facing the group. With respect to this important challenge, the following high priority questions remain open and need careful consideration:

1. What data is important to include within a SWID tag, minimally? Related, where within the SWID format would this minimal data reside?
2. How should platform-specific vs. platform-agnostic issues related to SWID publication, installation, and collection be handled?
3. What is the best approach to balance primary source/authoritative SWID tags vs. 3rd party generated tags?

After introducing the questions, an open conversation was held. The following topics were discussed:

- The organizer reiterated that the intent for this event is to identify challenges and issues within the relevant security automation topics, and not to attempt to solve them during the event.
- It was suggested that SWID tags could be created at several points in the software lifecycle, including development, compilation, installation, and interpretation.
    - At any of these stages, some different information could be available. One example given was the possibility that at installation time, an installer could add information to a generated SWID tag that would not be available at development time, such as installation directory.
- Another challenge discussed by a few of attendees was the how to handle the case where original software binaries are altered by the end user. In some cases this could be done in an acceptable (desired) manner, which in others it might be undesirable, either intentionally or unintentionally.
- Several attendees pointed out that there are great challenges in data normalization both with and without SWID tags:
    - Registry scraping data is unreliable and cannot be successfully used in comparison to SWID tag data, without significant investment in mapping exercises.
    - One could attempt to generate SWID tags for software that does not currently supply such a tag. NIST's National Software Reference Library (NSRL)[27] was mentioned as a possible source for such a task.
    - One attendee suggested that the U.S. Government should seed the effort to provide mappings for normalization purposes as a way to help SWID be successful. A counter to this suggestion was that type of maintenance was costly and not feasible over time.
    - Finally, it was also pointed out that many software products have multiple installable components, adding to the challenge of normalizing naming data.
- Another attendee pointed out that SWID tags have two value propositions:
    - The first, short term value is in providing a better way to normalize the naming of software, similar to CPE.
    - The second, longer term value will be found when there is a more mature, supported SWID ecosystem in place. At this point, SWID can provide things that CPE could not.

[27] http://www.nsrl.nist.gov/

- The importance of having high quality data was also discussed.
    - In looking at success stories, automatable solutions seem to provide the best approach for achieving high quality data.
    - Better data is also achieved when the primary source vendors provide the information, as opposed to 3rd parties.
- Finally, one attendee suggested that the community could continue to use CPE until the SWID solution was viable.
    - The general consensus was that SWID would be a better solution immediately, while providing better value long term as well.
    - It was also pointed out that one could easily generate CPE names from SWID data.

Community Action Items

- Consider how to get primary source vendors incentivized to fully support SWID tags.
- Research how and if CPE names generated from SWID tags could provide value to the community.

Day Three – Configuration Items and Assessment

Current Challenges with Configuration Guidance and Standards

Introduction

Over the years, the security automation community has faced a variety of challenges and issues with several of the core SCAP standards (Common Configuration Enumeration (CCE), XCCDF, and OVAL) for assessing the posture attributes associated with an endpoint. This session primarily focused on highlighting those challenges and issues from different perspectives in the community including a Program Manager's perspective, a Developer's perspective, and an Implementer's perspective.

Details

A Program Manager's Perspective

- The creation of guidance is important, but the community needs to get away from needing the low-level technical details and experience necessary to create and maintain the guidance because it is time consuming and expensive. In most cases, the high-level security mechanism is the only relevant item and not the low-level, platform-specific details that implement that mechanism.
- There is a need be able to roll up and drill down the results and be able to have meaningful and actionable results that allow an administrator to know exactly what needs to be changed to mitigate risk or remediate an issue.
- One attendee explained that they don't like how current tools roll up vulnerabilities in multiple instances of software into a single CVE for reporting purposes because it then requires system administrators to go and figure out what software on an endpoint has that vulnerability.
- Many of the existing efforts such as CPE and CCE are extremely flexible in how content is created leading to lots of ambiguity and inconsistencies in content. Furthermore, efforts such as CPE and CCE rely on too much centralization and suffer from a lack of incentive for primary source vendors to create and maintain content.

- There is a need to be able to extend efforts, like OVAL, more quickly to address the current needs of the security automation community and allow organizations to write the content they need. Also, there are efforts like XCCDF which is currently in ISO and is not easily revised.
- A couple attendees felt that the security automation community has been too "enumeration happy" and maybe there are other approaches where one can identify configuration concerns without having to know about it first.
- Another attendee suggested that it may be useful to set up time budgets for how long it should take for a CCE to be developed or a checklist to be written and then creating a workflow that supports that time budget. Another attendee suggested that the time budget may vary depending on the scenario (e.g. the time budget for compliance would be much longer than incident response).
- There is a need to expand the content breadth. Right now, SCAP has good support for core operating systems (Windows, Linux, etc.), but, it needs to support emerging platforms such as platforms that provide internet services (DNS, SNMP, etc.), databases (SQL, Hadoop, etc.) applications, cloud stacks, industrial control systems, etc.
- We need better ways to related endpoint posture to risk.
- There is a strong need for content authoring tools to take away barriers for people to develop content.
- Content creation is challenging. Understanding the business context, what information needs to be checked, and how to check it are the challenges, while the XML itself is relatively straightforward to compose.
- Developing international standards takes time. There is still room to improve upon SCAP, especially as products are still being validated against SCAP 1.2. There are no plans to abandon SCAP.

A Developer's Perspective

- Content is complex and hard to produce. Authoring tools haven't advanced and the flexibility makes things difficult to check. To create content, it requires knowledge of both policy and the endpoint being assessed and the current standards are currently forcing people to use XML even though they don't typically work at that level.
    - There is a need for the community to create an open source tool to help people auto-generate content in a way that is familiar to them (i.e. not XML).
    - Attendees had mixed feelings regarding content creation, some thought it was relatively easy and the primary challenge was figuring out what to check whereas other attendees have seen their customers constantly struggle with content creation.
- One attendee pointed out that the U.S. Government still does pay for activities such as OVAL XML development even if indirectly through contractor support. Additionally, the indirect cost could be higher, as the contractor may not have the proper level of expertise to develop and maintain content effectively.
- One attendee explained that they do not have customers paying for content. They are creating content because the content is helping their organization. Another attendee explained that they are not seeing demand from their customers for SCAP content as the system administrators are fine with hardening bash scripts that are available.

- One attendee explained how the Script Check Engine[28] was developed to support scripting in XCCDF and is currently supported in OpenSCAP[29] and jOVAL[30]. They also mentioned that they have scripts for generating XML. Allowing scripting in content and having scripts that can generate content could help system administrators create content in a way that is familiar to them.
    - Scripting was previously excluded from OVAL due to security concerns such as executing arbitrary code in content. Organizations should be allowed to assess this risk for themselves rather than disallowing it outright.
    - In more recent discussions[31], the OVAL community seemed to open up to the idea of scripting as something that should be further investigated as it could go a long way towards satisfying their short-term needs.
- One attendee stressed that standardizing the evaluation of posture attribute data is important, but posture attribute collection should be done in a way that makes the most sense for the platform.
- One attendee thinks there is big demand for SCAP. The fault lies in the business model and the community needs to establish a business model for what is a commodity now and what is expensive now. It is a commodity to write code now. It is incredibly expensive to manually interact with people to get something reviewed and certified. As a result, there is often times no incentive to write SCAP content because it does not impact processes that people are using.
- It was noted that the SCAP Discussion List[32] is a great place to bounce ideas of the SCAP community. (Though some audience members indicated they were unaware that such a list existed.)

An Implementer's Perspective

- The presenter explained how the two main challenges for an implementer is knowing what benchmarks need to be assessed against what endpoints (targeting) and being able to get meaningful and actionable results without overloading endpoints and network bandwidth.
- One attendee discussed how part of the problem is that batch collection occurs and get lots of results rather than just collecting and getting the results that one cares about. It would be very beneficial if the standards were more event-based to complement this batching paradigm.
- Another attendee explained how the SCAP Validation Program requires that one keep the results as XML which bloats the data even though it is not really needed beyond that requirement.
- An attendee noted that the security automation community needs to revisit SCAP and the assumptions around it because it was originally designed to be run on a single endpoint with all the data included in a set of multiple XML files that are sent back and forth. It does not consider a lot of things that are relevant today.
- One attendee discussed their experiences regarding targeting and how they are getting back too much data. Under the current SCAP paradigm, they send all of their benchmarks to all of their systems, evaluate the targeting on the local endpoint, and get back all of the results. They noted that they would really just like to get the results of the things that they need rather than everything.

[28] http://www.open-scap.org/page/SCE
[29] http://www.open-scap.org/page/Main_Page
[30] http://joval.org/features/schema-platform-support/
[31] https://github.com/OVALProject/Sandbox/issues/21
[32] http://scap.nist.gov/community.html

- One attendee explained that targeting is not currently supported in the current standards and understands why large documents must be created for interoperability purposes, but did not see anything in the specifications that say one couldn't reduce data sizes or do more event-driven assessments and as result seems more like implementation problems.
- There was agreement among some attendees that targeting means applying information labels that are organizationally defined. One attendee explained that they could send information on their requirements and another attendee explained how they already do tagging for this purpose and may be able to contribute their work to the security automation community.

Community Action Items

- Prioritize what new platforms need to be covered so that the necessary extensions and content can be developed.
- Create an open source tool to help people auto-generate content.
- Investigate how to speed up the process and eliminate the bureaucracy associated with extending efforts like OVAL.
- Develop time budgets for the various efforts.
- Many attendees stated the need for scripting in security automation. Let's look at the Script Check Engine and determine if it fits the community's needs. If not, what can the group learn from it? What could be improved?
- Over the years, SCAP has been focused on scanning local endpoints on a periodic basis, investigate what it would take to make SCAP more event driven.
- Determine what would incentivize content creation for primary source vendors.
- Write a document that describes how to create a checking language because multiple checking languages might be required for different platforms.
- Determine if existing information label tagging work can be contributed to the community.

Guidance Challenges

Introduction

The focus of this section was a (relative) outsider's view of the challenges with guidance. The section began with discussion of those challenges including difficult of learning the relevant standards and tracking assets over time. The presenter also gave some suggestions to the community including making things less report-based and more web resource-based.

Again, following the presentation a great deal of conversation occurred.

Details

The challenges laid out by the presenter focused largely on the difficulty in picking up the standards and other related context on the security automation space. He lamented the lack of a good "Hello World" example. He also asked if there was a good way to track assets over time and whether there exists a way to manage waivered or exceptions to the prescribed guidance. The community generally agreed to this list of issues.

The presenter also highlighted the fact that there is relatively little SCAP content available from trusted sources. He then discussed the Tier IV[33] content from NVD to support this notion, pointing out that NVD had very few SCAP automatable (that is with OVAL content) benchmarks. Not only are there very few of them, but they are also very old and/or address old platforms like Windows XP.

Finally, the presenter also argued that instead of using a report-based solution to share assessment results, a more web-based approach would be more powerful and intuitive.

During and following the presentation, the attendees had a number of questions and comments; the following is a summary of those points:

- Some attendees believe that decisions made in the past cause some of the issues raised by the presenter. The perceived limitations on checking languages in SCAP and the lack of an open scripting ability in OVAL contribute to making OVAL harder to learn and less flexible. It was pointed out that SCAP does not necessarily limit the checking languages that can be used. While it does mandate support of OVAL (and the Open Checklist Interactive Language (OCIL)[34]), it does not limit support for other checking languages.
- Education on security automation and the related standards was a common discussion point.
    - One attendee stated that the process to submit a patch to NVD or how to get content fixed is not clear.
    - It was pointed out that NIST does not maintain all of the benchmarks found in the NVD and that the responsibility for updating those falls on external parties.
    - The Red Hat and OpenSCAP[35] documentation were held up as shining examples in the education area.
    - It was also noted that a Coursera[36] course is currently being developed to help address the education shortcomings.
        - This effort is looking for contributors.
- The difficulty in working within the U.S. Government space to quickly and effectively achieve specific goals was raised. Generally it was acknowledged to be an issue without a clear solution. One suggestion was that increased engagement from other stakeholders could alleviate some of this issue.

Community Action Items

- Develop better education and training materials to help lower the entry barrier for security automation and standards.
- Contribute to the creation of a Coursera course on security automation.

[33] http://web.nvd.nist.gov/view/ncp/repository/glossary
[34] http://scap.nist.gov/specifications/ocil/
[35] http://www.open-scap.org/page/Main_Page
[36] https://www.coursera.org/

From Lessons Learned to Possible Alternatives

Introduction

Several vendor representatives formed an informal panel to discuss lessons learned from SCAP and other previous security automation efforts so that the most effective steps can be taken. The panel brought up some lessons learned and also discussed some suggestions for moving forward.

Details

During the panel, both panelists and attendees had a number of comments and/or questions. Those comments and questions are outlined here:

- One panelist strongly recommended that any relevant work be done within the IETF SACM WG, as opposed to working out solutions in isolation and then moving things over to the Working Group.
    - The XMPP model could serve as a model here; that WG does some work directly in the official forum, but also has a non-profit stood up to help drive things as well.
    - One of the attendees suggested that it will difficult to do everything in the WG, holding out things like education and training as things best done outside the group.
- Another panelist provided the group details on how the IETF process worked, highlighting a few key points:
    - All decisions are made over email lists.
    - No direct support can be given in IETF by corporations. While organizations can pay for individuals' time contributions to IETF, the individuals can only officially represent themselves.
    - While official IETF meetings are held 3 times per year, a WG can hold many interim meetings as well to achieve goals.
- One attendee highlighted the importance of running code as a requirement for work in the IETF. By requiring running code against anything adopted by the standards bodies, higher quality standards become more likely.
- There is a concern that the U.S. Government is mandating NIST standards, but then not effectively supporting the development of these standards. Attendees would like to see the government host more events to discuss these standards and training.
- A general comment made by several attendees was that while the conversation on these topics is great, it is critical that follow up work be performed to move these efforts forward. The notion that the group needs to actually do something and not just talk about problems and solutions came up a number of times during the week.
- One attendee was concerned that the requirements for this work have not been adequately defined. Several attendees pointed out that the IETF SACM WG has created a requirements document and that if during review anyone believes that one or more requirements are missing, they are strongly encouraged to engage the group to fix the omission.

Working Session

Introduction

This session was provided to wrap up the event's three days of discussions and attempt to identify solutions for moving forward. One workshop organizer provided priorities to the attendees for their

main concerns and identified areas of potential improvement to existing efforts. This was an unstructured conversation with the intent of fostering extensive feedback.

Details

There were four main issues identified for this discussion. Priority number one was for improved standards related to software inventory collection and transport. The presenter identified a need for a vendor-neutral and platform-independent way to report software that was present, missing, or in need of patching. There is a reliance on vendor bulletins to correlate vulnerability information with versioning information as well as remediation procedures. There was discussion about the potential to have the NVD and NSRL automatically generate SWIDs, and to push for higher adoption rates of SWIDs. The latter could be coupled with the inclusion of SWIDs and other software identifiers in applicability checking languages.

The second priority was improved standards related to configuration assessment. The goal for this series of checks was to report on truly platform-agnostic controls. This would allow primary source vendors to determine the best way to check such controls. This implies that the configuration assessment solution should handle a multitude of tool outputs to de-conflict the results based on the most accurate collector.

The third priority covered the ability to discover, share, and consume standards-conformant content. This was discussed previously with the need for repository protocols to automatically acquire desired guidance or checks, however was not discussed in depth during this presentation.

The final priority was the capability to perform continuous monitoring tasks based on event notifications. These event notifications would provide specific updates to inventory changes. It would greatly reduce the time to re-scan systems if the system offered new results on significant changes rather than a central location querying on a set schedule. This priority was also not discussed in depth during this presentation.

There was much open discussion during this session. The following captures the main clarifications, comments, and questions:

- Vendor bulletin information provided to NIST such as the executables affected and vulnerable libraries could be used to further identify other vulnerable software that rely on those same packages.
- One participant had asked whether the expected bulletins were required to allow a user to take automated action against vulnerable software.
    - If the automated remediation aspect delayed the informational aspect then it could be ignored for now. The important piece was knowing whether one had to take action.
- For unknown software found during inventory, the ability to generate a unique identifier that can correlate multiple instances of this particular unknown software is highly desirable.
- There was a question on the frequency of such inventory scans. Some enterprises can collect data every 3 days while others are currently operating with less frequent collection. The more real-time this can become the better.
    - There was a strong emphasis added that accurate information far outweighs quicker access to information. This was reflected in the fourth priority of event-based updates being seen as a stretch goal once inventories were collected to satisfaction.

- Some asked about the necessary level of detail in the reporting phase of assessments. Generally attendees felt that applications should report at least to the version & patch level to correlate with vulnerability information.
- Another attendee questioned whether the goal was for a tool to provide the view of the results or for the ability to synthesize the view from the results.
    - Due to the nature of multiple sensors and tools and the potential to add additional data later on, it would have to be the option where a view could be synthesized from a fusion of the data.
- SWID adoption is not mandatory for these priorities moving forward, but it is the best current solution. Since SWIDs are more data-rich, they can be used to back-create CPEs for current needs, but CPEs lack sufficient data to be the source for generating SWID tags. Usage of auto-generated SWIDs from the NVD/NSRL would suffice until primary source vendors begin to provide their own.
    - The talking point about government contracts requiring the usage of SWIDs was brought up again. Mandating it in such a way would be a quick method to getting SWIDs adopted faster.
        - Several others agreed that a roadmap for SWID adoption would greatly increase their planning capabilities to meet this need quicker once implemented.
    - Regarding the further adoption of SWIDs, one person questioned whether there was a specific version of SWID targeted for these priorities.
        - This is currently being addressed by the SWID interoperability work.
        - There would be no plans to require signed SWIDs before increased adoption is achieved.
    - An attendee pointed out that the lack of SWID implementations slows the progress of expanding adoption. If SWIDs are too slow to be proven technically feasible to be accounted for in the SACM model, then that is a huge missed opportunity.
        - Another attendee committed to generating a SWID dashboard as proof of concept over the next few months to help with adoption.
- Regarding the second priority of configuration assessment, one participant noted the previous attempt to use CCEs were provided by platform. This approach caused great difficult with respect to scaling.
    - One other attendee added that it was fantastic to see vendor participation. The only downside was the manual hours associated with each check from that large quantity to ensure the check was of the quality to include in the official repository.
        - One organizer attributed this to the lack of validation tools to provide to vendors to check their own CCE submissions.
    - One of the organizers asked whether CCEs were ever requested of a vendor to be provided.
        - One attendee responded that he was asked to include them for completeness because they existed.
        - One other participant noted that they were told they had to map the CCEs with OVAL checks for compliance.
    - CCEs being tied to a pass or fail percentage is often misleading due to overlapping or compensating controls.

    - If there was to be a reuse of something similar to CCEs then it should be platform agnostic.
        - One person noted that it becomes another mapping problem again to keep track of all the levels of inheritance.
    - One attendee suggested that embedding the check information within the CCE itself would be a way to improve the utility in CCEs.
        - This becomes complicated depending on whether evaluation is done at a central location and not on the endpoint itself.
        - Additionally, the possibility of a CCE on a platform to be collected different ways might mean a different identifier would be used.
        - One other participant disagreed with the proposal and pushed to continue investigating solutions using existing methods rather than start something new.
    - There was a small discussion on the original purpose of CCEs.
        - One participant pointed out that CCEs were never meant to be machine readable and purely used for human data correlation.
        - One other participant suggested that they were originally for correlation between old configuration management databases, with CCEs as the common mapping.
- The level of controls targeted for this priority were described as one step below high level. There is no need here for knowing the lowest level being collected.
    - The example of execution prevention was provided. Microsoft uses ASLR while Red Hat uses DEP.

Community Action Items

- Investigate the feasibility of generating SWID tags from the NSRL data.
- Create a proof-of-concept dashboard that uses SWID input for software inventory, vulnerability management, and (possibly) targeting activities.

Conclusions

The workshop provided a great opportunity for vendors and other security automation stakeholders to have a frank and valuable series of conversations with several U.S. Government personnel. During these conversations, wider context was provided across the community and specific conversations on Software Inventory and Configuration Assessment helped work towards a shared understanding of both the current issues as well as the way forward.

During the workshop, several overarching themes emerged:

- The security automation community is 'balkanized' at present—there are disparate project-specific community e-mail lists and sites (e.g., SACM list for SACM-specific activities; SCAP lists for discussions related to particular SCAP standards; TagVault member-only lists for member-only discussions related to SWID tags). There is no central location on the web (web presence and associated discussion list) today for rallying the security automation community, or even the sub-community focusing on security automation standards. This needs to be addressed, and

there seems to be a consensus among participants that there is a role for U.S. Government leadership.

- There have been at least two prior meetings where there were extensive discussions about SWID tags and the need to promote wider adoption of tagging standards and practices. To some degree, adoption has been hindered by (1) limitations of the 2009-era standard, (2) opacity of development of the 2014 revision to the standard, (3) inability of the U.S. Government to publicly state or clearly explain their stance towards SWID tags, (4) lack of public forums for open discussion of tagging standards and practices, (5) lack of freely available tools to make tagging easy for publishers, (6) lack of "killer apps" demonstrating the utility of tags to consumers.
- Somehow the emerging pattern of handwringing without effective goal-setting and action needs to be broken.

The following is a list of broadly applicable next steps that need to be understood and worked as a community.

Community Action Items

- All members of the community have been strongly encouraged to get involved and engaged with the IETF SACM group.
- A roadmap from the U.S. Government should be created to aid vendors and other stakeholders in supporting and advancing necessary standards.
- The community should create one or more proof of concepts to show both their level of commitment and the overall value of the effort.