SCAP v2 Workshop Minutes

April 30, 2019 – May 2, 2019

The MITRE Corporation – McLean, VA

This document provides a summary of the Security Content Automation Protocol (SCAP) v2 Workshop held at the MITRE Corporation in McLean, VA from April 30, 2019 to May 2, 2019. This document describes the key points made during the discussions rather than a full transcription. As a result, slides and meeting recordings are available at the following links.

Slides: https://csrc.nist.gov/CSRC/media/Projects/Security-Content-Automation-Protocol-v2/documents/Presentations/April%202019%20F2F/AprilF2FPresentations.zip

Day 1 – April 30, 2019

# Welcome, Introduction, and Broad Event Framing

Jessica Fitzgerald-McKay (NSA), David Waltermire (NIST)

### Summary

This presentation welcomed the community to the SCAP v2 workshop, highlighted key challenge areas with respect to endpoint assessment, and identified the goals for the SCAP v2 workshop.

### Background

* Cyberattacks keep happening and are often the result of a failure to perform basic cybersecurity hygiene. Basic steps for cybersecurity hygiene include understanding what endpoints are on the network; understanding the context in which those endpoints are operating on the network; understanding how to mitigate risks to those endpoints.
* The SCAP v2 architecture introduces several key components/services and notional interfaces to facilitate automation of the basic steps for cybersecurity hygiene. These components/services include:
  + *Posture Collector*: Collect endpoint data from endpoints and store it in the Configuration Management Database (CMDB)
  + *SCAP Content Repository*: Store and distribute instructions that drive the collection and evaluation of endpoint data
  + *CMDB*: Store and provide endpoint data for evaluation
  + *Posture Evaluator*: Evaluate endpoint data
* SCAP v2 aims to support a variety of endpoint types including, but not limited to, traditional endpoints, network devices, mobile devices, and IoT devices.
* The goals for the workshop are:
  + Discuss a broad range of topics that may be of interest to the community
  + Identify 3 - 5 work areas that are critical to the community and work them to a solution
  + Create a subgroup for each work area and identify those in the community willing to lead or contribute to subgroups
* The goals for SCAP v2 are:
  + Make SCAP v2 more accessible than SCAP v1
  + Improve visibility into vulnerable endpoints
  + Support event-based security content data collection
  + Collect data once, use many times

### Key Discussion Points

* There is a need for SCAP education and outreach for new users. While SCAP education and outreach has typically been a NIST activity, there is a role for vendors and integrators in this effort to reach out to potential users and to help provide education.
* It was noted that not all SCAP tools provide full SCAP support; some only provide what is necessary for validation. Differences in SCAP support hurts usage by organizations depending on the quality of their tool. The key to addressing this issue is to increase the value that SCAP provides so that tool vendors see market value to implement full SCAP support.

## SCAP Content Metadata, Structure, and Interface

Charles Schmidt (MITRE)

### Summary

This presentation defined what SCAP content metadata is, how it is used in SCAP v1, and the vision for expanding its use in SCAP v2.

### Background

* SCAP content metadata is defined as a set of values that is used to describe and categorize content. Examples include title, tags, content author, date created, last modified date, type of check, and applicable software among others.
* Currently in SCAP v1, only limited metadata is provided by repositories and there is no standardized way of acquiring that metadata. Furthermore, there is no standardized mechanism to obtain alerts about changes to metadata in content.
* For SCAP v2, there is a desire to better understand what content is available in a repository, to improve the ability to search for applicable content, and to support organizationally-defined metadata.

## National Vulnerability Database (NVD) Metadata Wishlist

Dave Waltermire (NIST)

### Summary

There is a desire for better software identification and metadata from vendors for extended vulnerability information as well as for stronger configuration checklists and metadata from suppliers. There’s a need to host data using a common protocol for automated distribution.

### Key Discussion Points

* There was interest in the idea of content creation as a service which would allow producers to advertise the types of content they create to potential consumers.
* It was noted that there are different types of content needed by consumers and content developers, which may require different metadata and a distribution protocol that would enable the support of a variety of consumers.
* It was noted that metadata for SCAP results is critical so that it is possible to understand the context in which the results were derived. Use cases were cited of where users would go back to sources to gather metadata to better understand the meaning of their assessment results. As such, metadata is not just a "front-end" feature to find content to use, but can be utilized throughout the entire assessment process.
* There was discussion around why vendors are contributing to NVD. It was noted that motivations vary depending on the type of information being contributed. For example, there are CVE CNAs who are interested in providing their own CVSS scores and software identifiers. This gives them more control over the information being provided and gives them credit for the work that they are doing. Getting vendors to provide their own checklists and CPEs is much more challenging. The group should discuss how a federated model may better incentivize the creation of other types of content.

## Repository Metadata

David Ries (Joval)

### Summary

Current repositories don’t provide automation-friendly metadata. Repositories need to provide a manifest in a common, automation-friendly format at a predictable location.

### Key Discussion Points

* One member mentioned that they put together a strawman based on XML, but, that JSON would be fine too. The strawman script programmatically generates whatever can be found in the XCCDF content (identifier, title version, contributor, publisher, checksum, etc.). It uses CPEs since they already there and generated. It was noted that the group should be able to come to a consensus on what the common fields are and all agree that something is better than nothing.
* Two SCAP commercial entities that performed content distributors expressed interest in this and NIST also expressed interest and has an implementation of something similar using ROLIE. It was agreed that it would be good to get the teams together and sync up on ideas.
* It was noted that different repositories manage content using a wide variety of technologies and that a content manifest would represent a small step that all could benefit from.

## Usage of Metadata in Siemens’ Scapolite Format

Bernd Grobauer (Siemens)

### Summary

Every document in Scapolite format is a YAML document. The desired approach for metadata change tracking is that each rule carries metadata about changes.

### Key Discussion Points

* It was asked what was the granularity of the changes (e.g., was it the recommendation/purpose for the rule, low-level registry or filesystem checks to test for the rule, both, etc.). It was explained how if someone changes a typo, you don’t need to know about it, but there are key words (created, revised, modified, etc.) that help track it. The changes discussed were more for the human-readable part of content and OVAL would likely be harder.
* It was noted that in tracking SP 800-53, there is a table of equivalences between ISO 27001 and what we have to comply with here. It was asked if the look-up table from NIST is helpful in complying with German requirements. There was interest in more granular rules about different settings. It’s less about an ISO requirement versus a NIST requirement versus a SP 800-53 requirement, but having mappings from high-level rules to ISO rules would make certification easier. Another member asked if they were suggesting that the XCCDF model needs to be updated to include additional metadata as well as define best practices for things like identifiers not changing. It was confirmed that this is what was suggested.
* A few pieces of metadata (manager who is in charge of it, etc.) were mentioned. It was asked if more information could be given on how the metadata is used by systems and internal processes since in previous calls consumers didn’t indicate that they inspect the content deeply (e.g., rely on some party to say they are using the latest version). It was stated that the metadata is what would be desired from a producer such that the consumer can build on that content. One of the biggest challenges discussed was version changes because using straight-from-the-vendor content often time requires modifications that go beyond what SCAP tailoring provides. The metadata helps track documents to minimize time spent on updates and enables people to filter rules based on system, environment, and role to determine responsibilities in relation to service providers and classification of the system. Given all this, when rules change, people are given an evaluation period to update to the latest rules.
* It was asked if there was a need for similar metadata in OVAL content. It was stated that they don’t use OVAL much right now, but, they may have thoughts if OVAL becomes more relevant to them.
* Getting back to diversity of endpoint types and IoT, there is typically an XCCDF benchmark where you have to take whatever OVAL is out there. It was asked if Siemens had created its own XCCDF and gone from the bottom up? It was stated that there are different OVAL extensions and they might not support that, but, it would be interesting to see how OVAL evolves in that area especially as potentially leveraging other collection mechanisms was discussed earlier. In some cases, it may be easier to leverage these mechanisms rather than installing an OVAL scanner on the device.
* One member pointed out that there is a middle ground as they are trying to use SCAP in an ICS setting where robotic manufacturing applications and controllers are running Ubuntu Linux and a popular Ubuntu library for controlling robots. Looking at the canonical repository there are thousands of definitions and it might be a good opportunity for standardizing metadata. The speaker asked if others had a similar need - it was noted that it makes sense, but little has been done so far.
* It was asked how widespread SCAP was used in Germany. Bernd said he didn’t know, but the CIS membership shows some German companies and it would be great to see more get involved.
* One member recapped some interesting points discussed throughout the content metadata sessions.
  + There is a critical need to be able to consistently generate metadata labels in a federated fashion. This was a problem with CPE and required a single authority to ensure consistency.
  + There is value in understanding check results in the context of the content run (e.g., where did it come from, what does it mean, etc.).
  + It’s critical that the metadata isn’t just a list of fields. We need to know software, authors, when it was created, but, we have to have consistency behind the process (e.g., when do we change identifiers, under what situations, when are things related or not related, etc.) and uniformity in semantic meaning. We need to make sure we have a good understanding of how this will work under the federated model and the fact that the teams aren’t going to be directly working with each other.
  + Generating profiles based on metadata seems like a promising idea. When you are filtering content on a repository this would be the next step beyond just selecting applicable content from a set list.
  + There is a need to address overlapping profiles (e.g., Windows Domain Controller CAT 1 and Member Server CAT 1) and better understand what controls were run and show people what the results means. Whether they are combined or both just applied will need to be thought about.
  + Multiple members in the community expressed interest in exploring this issue together.

# Endpoint Data Collection

Jessica Fitzgerald-McKay (NSA)

### Summary

This presentation discussed the challenges associated with collecting data from an endpoint and communicating it to a posture collection server/service in a standardized way.

### Background

* Endpoint data collection is challenging because tools do not generally use standardized mechanisms, and these tools do not make the data available to other tools. Furthermore, infrequent collection means that changes may go unnoticed for a long period of time.
* For SCAP v2, the goal is to standardize collection mechanisms to avoid redundant collection of data and consolidate tools as well as to improve the freshness of data by leveraging event-based collection.

## Internet Engineering Task Force (IETF) Security Automation and Continuous Monitoring (SACM) Working Group (WG) Overview

Bill Munyan (CIS)

### Summary

SACM, a working group within the IETF, consists of a community of security automation subject matter experts who work to define requirements, specification, interactions, and data models. The group has made progress in standardizing architecture, collection, and evaluation.

### Key Discussion Points

* It was asked whether or not the IETF has a tool validation process similar to NIST’s National Voluntary Laboratory Accreditation Program (NVLAP). It was explained how IETF specifications undergo significant review by the WG and IETF Area Groups prior to becoming an RFC.
* Not all endpoints can be provisioned with an agent, and a question was raised about how things like Network Access Control (NAC) would fit in and how ports and protocols could be utilized if an agent could not be installed on the assessment target. It was noted that IETF YANG data models and the NETCONF, RESTCONF, and YANG Push protocols could be leveraged to get that type of information and that not all these technologies require an agent.
* The NIST/NSA vision for SCAP v2 includes acquiring data from endpoints both when you have and don’t have authenticated access.
* There was discussion around whether or not UUIDs were going to be supported in SCAP v2. It was stated that nothing in SCAP prohibits the use of UUIDs.

## Experiences Linking Many Sources with Many Analytics

Gordy Scott (MITRE)

### Summary

MITRE’s Data Center Infrastructure Management covers an extensive array of products with a wide range of features and functions. They fielded a home-grown solution that consolidates asset databases, provides visualization of data centers and assets, has advanced reporting capabilities, and provides real time monitoring.

### Key Discussion Points

* It was asked if a centralized database was used for all endpoints or if different ones were used and consolidated with a single application. A single database was used because the tool focused on collecting what is needed and not everything that could be collected.
* There was discussion around a gold standard for collected data and characteristics, but for this solution, a single management tool which provides hardware, software, and configuration was used as the gold standard and is linked by property number.
* There is a need to know both what you can collect, what you need to collect, and how to collect it.
* It was noted that standardizing the endpoint data collection interface isn’t an all or nothing proposition. The group should consider how much data standardization is necessary to make this useful.
* There was a concern that having a specific way to interact with a certain endpoint could dictate implementation given there are so many different ways for systems, tools, and vendors to get information about devices. It was noted that OVAL already supports the need for a generic way to interact with endpoints and a protocol isn’t necessarily needed, rather it is a spectrum, and if there is a standard way to describe endpoint data we can just use tools that provide the data (e.g., MDM, agent, scanner, etc.).
* It was suggested that, the group needs to determine what is the minimal level of standardization needed for endpoint collection. There was a concern that given the variety of endpoints everything could be special cases for this interface making it difficult to standardize on. Furthermore, it was noted that if vendors are just asked to provide a posture collection service, they might be more willing to support it rather than if we say they must do it a certain way, in which case they just won’t support SCAP. Interoperating with other tools and APIs is not the challenge, it’s having a format in which to communicate the data. With that said, the group acknowledged a need for some standardization and that architectural components could be thought of as servers or services (and there may be more than one instance).
* There is a need to determine what endpoints are on the network, what attributes are available, and a way to ask a posture collection service/server to acquire that data with minimal disruption to the network. It was noted that not asking vendors to re-architect their solution for SCAP v2 would be great. Similarly, it was recognized that authenticated access would not be possible for certain endpoint types, however, if you can acquire other information from the network, it may be possible to get good information about those endpoints.
* Several people expressed interest in defining the requirements for an endpoint posture collection service.

# Configuration Management Database (CMDB)

Dave Waltermire (NIST)

### Summary

A goal of SCAP v2 is to eliminate redundant data collection. There is no need to collect all data when you only need a subset of data. The CMDB will provide that capability.

### Background

* A primary goal of SCAP v2 is to eliminate redundant data collection and enable that data to be reused across evaluations. In order to satisfy this, the data must be collected, stored, and accessible to tools for later use.
* CMDBs provide several benefits, including the elimination of redundant collection. A CMDB also serves as a single data repository for all evaluation tools, which gives a consistent picture of the enterprise and eliminates the need for evaluation tools to have permissions to communicate with every endpoint on the network.
* Some challenges associated with CMDBs include protecting the data stored within them, identifying the types of data required by different evaluation tools, and the need for standards.

### Key Discussion Points

* Multiple people agreed that the goal of endpoint data collection and storage in a CMDB should be to only collect the data necessary for security assessments rather than to collect all available data.
* It was noted that redundant data is more an issue with compliance checks rather than configuration items. For example, password length can be determined by one setting, but, PII, HIPAA, and RMF all required different password lengths. This highlights the need to separate collection and evaluation (collect-once-use-many-times).
* There was a question about the difference between a CMDB and an asset management database. It was explained that, from a technical perspective, the CMDB should hold hardware and software inventory information, whether an endpoint is patched, and software configured properly. Other information includes characteristics assigned to an endpoint such as criticality, role, owner, managing organization, etc. A CMDB should encompass both types of information.
* There was a question as to the purpose of the SCAP v2 components. Is it to define the interfaces? Or, is it to just go over the use cases? It was stated that, from a minimal perspective, the group can talk about characteristics that are desirable and what formats can be provided over these interfaces. The group can also talk about things at the data level or at the transport level, be data agnostic, and move data across the network regardless of what it is. The group needs to figure out what the initial phase of work is and then work toward it. It was noted that if the group can focus only on data models it can be very productive whereas also focusing on components and protocols could make achieving results difficult.
* It was noted that the group should determine key use cases, see what standards are available, and enable some minimal-level of capability for the CMDB.
* The group was asked who has read the Asset Identification specification (about 8 people raised their hands). The group was also asked if anyone would be willing to update it. It was noted that something like AI is needed, but AI doesn’t need to be the sole solution.
* The group agreed that once the scope for SCAP v2 is better defined, it should look at the specifications in SCAP v1 as well as other standards bodies and determine which ones would best address the SCAP v2 use cases.

# Content Distribution Protocols

Danny Haynes (MITRE)

### Summary

This presentation provided an overview of ROLIE (RFC 8322) which is a protocol for the publication, discovery, and sharing of security automation content.

### Background

* One of the benefits of a standardized format is that enterprises can acquire content from a variety of producers. Unfortunately, it can be difficult to find content and learn of updates.
* Content is currently distributed in many ways, but, it can be difficult for users to find content and learn of updates to content. However, there are no standardized protocols for acquiring available content.

## Resource-Oriented Lightweight Information Exchange (ROLIE)

Stephen Banghart (NIST)

### Summary

ROLIE is a potential content distribution solution. ROLIE provides a generalized information approach with a data format and transport protocol for publishing, organizing, and sharing computer security information.

### Key Discussion Points

* Given that ROLIE leverages the DNS model, it was asked if implementers would need to set up root and cache servers and transfers that need to be tracked. Since ROLIE uses DNS-SD, the implementer only needs to set the zone file and then use the DNS query to do it. It was noted that ROLIE is built on the Atom Syndication Format and that most browsers already have support for Atom. For example, it’s possible to browse the ROLIE server in a web browser and see the types of information it uses.
* A vendor explained how there were three use cases that they cared about with respect to content distribution.
  + 1) A content publisher that is also consumer of lots of content. ROLIE seems like a great fit.
  + 2) A solution being run by an enterprise and subset of content being pulled in as part of that solution. ROLIE could be useful.
  + 3) Not taking content from a variety of sources and not publishing a lot of content. ROLIE could be too heavy of a lift and a lighter manifest format could be valuable. It was noted that ROLIE could be implemented to leverage a flat file and HTTP GET.
* Given several specifications are using XHTML and incorporating Bash and Puppet remediation content, it was asked if other specifications were migrating this way as well. It was noted that ROLIE limits where you put XHTML and is not intending to drive remediation, but the security content it distributes could be remediation content.
* One vendor asked if SWID tags should be submitted to NVD for publication. It was explained that it would be ideal if vendors set up their own ROLIE servers to host their content. Then, NVD could just subscribe to the vendors feed. NIST offered to help anyone that wants to stand up a ROLIE server.
* It was asked how someone on a disconnected network would pull down a ROLIE feed. To do this, a local ROLIE server that is accessible on that network could be set up and populated manually.
* There was a concern that if there was a lot of content then pulling ROLIE information down with a web browser could be difficult. To address this, a user could get the content length and tell the server to round it down.
* It was asked if ROLIE could filter content by content type (e.g., XCCDF, OVAL, etc.). It was confirmed that ROLIE can do that.

Day 2 – May 1, 2019

# Day 1 Highlights

### Summary

This presentation discussed the key points made during the previous day’s discussion.

### Key Discussion Points

* It was mentioned that in the commercial space, there is a lot of SCAP content through curated sources (e.g., McAfee, etc.), GRC vendors where multiple compliance regulations impend on an asset, and rationalization of controls and optimization of collection. It was then asked if there is a role or any recognition in the roadmap for things like ROLIE to consider how do we bring curation into the process and how do we track provenance across that. Yes, it was confirmed that the SCAP v2 ecosystem needs to be dynamic and encourage people to update content rather than just being bodies of dead documents where consumers look for the closest match.

# Evolution of OVAL

### Summary

This presentation discussed the current state of OVAL, how OVAL aligns with SCAP v2, and proposals and challenges associated with OVAL that were brought up on past community teleconferences.

### Background

* OVAL is the automated checking engine for SCAP v1 and is moderated by CIS, the OVAL Board, and the OVAL Community.
* While OVAL satisfies some goals of SCAP v2, there are key areas where it falls short and needs work (e.g., ease of content authoring, event-based collection, reliable applicability, etc.).

## OVAL for Specialists

Matěj Týč (Red Hat)

### Summary

OVAL doesn’t have supportive tooling making it hard to have everything in one place. Supportive tooling would provide test environment scenarios, visualize results, and debug OVAL evaluation.

### Key Discussion Points

* One concern raised was the lack of money to fund SCAP and OVAL development. Otherwise, it is only a coalition of the willing and things get done as time allows.
* It was asked if there are other checking engines that could be used other than OVAL. It was noted that there is the Ansible project, which is an agentless python solution, but from a design perspective, it is not something OVAL could be inspired by.
* One member of the group re-iterated that SCAP/OVAL is becoming less and less a part of their assessments because it’s difficult to create. Furthermore, it is too verbose and difficult to debug.

## Programmatic OVAL

David Ries (Joval)

### Summary

The OVAL format is complex and difficult to work with. Script-based DSL for assertions should be considered, as they would retain core benefits of OVAL while leveraging tools and techniques developed for code.

### Key Discussion Points

* One vendor indicated that a key driver of adoption is the availability of content and that they are seeing lots of adoption in the commercial space because of it.
* It was noted that OVAL and other SCAP specifications have challenges associated with being relational and all the jumping around and whereas scripting is easier to write and use, although people are uncomfortable running other people’s scripts. As a result, there may be some benefit to keeping the assertion-based aspects of OVAL and supporting the programmatic constructs provided by scripting for generating content. Another person noted that not everyone creating OVAL content are developers.
* It was noted that this could work as vendors use libraries without necessarily knowing how it works. However, a concern was raised indicating that while libraries are great for automating and obscuring complexity, they may make it difficult to understand the context of how and why a check failed.

## Protocol-Oriented OVAL Schemas

David Solin (Joval)

### Summary

OVAL has focused on authenticated endpoint scans, but there may be considerations for unauthenticated scanning to which OVAL can be applied with the right new schemas. OVAL support has been added for network devices. There are additional protocols that need to be considered.

### Key Discussion Points

* It was noted that while OVAL historically has focused around authenticated, on-the-endpoint scans; support has been added for network devices (Cisco, Juniper, VMware ESXi, and NETCONF). There are also unmet use cases for vulnerability detection, which could leverage protocol-based schemas.
* It was asked how long it took to extend OVAL to support this. It was explained that the majority of the work is in the time and research to read the protocol specifications, but after that, writing the extensions is not the most difficult aspect. The bigger hinderance is getting the extensions into OVAL, then into SCAP, and then implemented by tools (1 – 5 years out). With that said, there have been improvements to governance models that should help alleviate this.
  + SCAP 1.3 introduced an annex so component specifications could be incremented as needed.
  + OVAL introduced a more community-based model for approving and getting extensions into the specification.
* These schemas will be submitted to the OVAL community to test the new governance process and to get more feedback.

## InSpec Overview

Aaron Lippold (MITRE)

### Summary

InSpec is an open-source, community-developed compliance validation framework. It’s cross-platform, integrates into multiple continuous monitoring tools, and is easy to create, validate, and read content.

### Key Discussion Points

* A question was asked about how InSpec handles a configuration file that might be set incorrectly and corrected later on. It was stated that InSpec will ask the SSH daemon for the correct file and use that. More specifically, InSpec can do passive and active testing and check different locations all of which is abstracted away in resources. It was also noted that the expansion of InSpec to support new capabilities is easily done.

# Evolution of XCCDF

Charles Schmidt (MITRE)

### Summary

This presentation discussed the current state of XCCDF, how XCCDF aligns with SCAP v2, and proposals and challenges associated with OVAL that were brought up on past community teleconferences.

### Background

* XCCDF is the format in SCAP v1 for expressing checklists and linking to checking instructions. However, it suffers from being rather verbose, and tailoring capabilities tend to make content and results more difficult to interpret.
* There are several other specifications in SCAP (OCIL, TMSAD, AI, and ARF). It needs to be determined how much these specifications are used and how they may need to evolve if leveraged in SCAP v2.

## Creating SCAP Content like a Book

Marek Haicman (Red Hat)

### Summary

Compliance-As-Code is a framework to build configuration compliance content with a number of contributors. It is focused on many products and expands beyond SCAP. It has evolved over the years, making results more maintainable.

### Key Discussion Points

* It was asked how you would create YAML-based XCCDF from XML-based XCCDF. It was noted that the YAML format was created to make content authoring easier; not to change the current XCCDF specification. The XML format used by XCCDF is not a big deal if it is just a machine reading it.
* It was asked if it would be better to have one format rather than multiple formats for generating content and whether it adds complexity since you now have various scripts equivalent to Schematron to generate XCCDF content. It was also asked if the content was validated prior to being transformed into XML. It was noted that validation is only performed on the XML and not the YAML as it is just an intermediate format. There was agreement among the group that having a single approach to generating content based on YAML would be beneficial.

## Modernizing SCAP

Gabriel Alford (Red Hat)

### Summary

SCAP is moving into the modern world by incorporating YAML for authoring and JSON for final machine language format. There’s greater flexibility in building SCAP content with smaller file sizes, there’s faster development, and it’s easier to understand and edit.

### Key Discussion Points

* It was asked if we could update SCAP to use an easier format like YAML. It was noted that while a YAML-based approach helps content authors, it does re-create the vendor adoption problem because they would have to start from scratch. Instead, we should work to get enough user demand to drive vendor adoption. With that said, some customers have been asking for YAML.
* A concern was raised about unifying all the content (XCCDF, OVAL, etc.) because it was counterintuitive since XCCDF is platform independent and OVAL is platform-specific. It was noted that the SCAP datastream is essentially the same format, but this work needs further review because it was only tested on simple examples so far.
* It was noted that this approach seems to be a difference between the serialization (YAML, XML) of the same data model (i.e. SCAP) and maybe the group should just focus on the data model and let the serialization vary. There seemed to be agreement around the fact that the serialization of the content is independent of the data model and we should keep them decoupled.
* It was mentioned that there is an OASIS effort called Darwin Information Typing Architecture (DITA) which is XML-based, but now allows YAML and Markdown for content authoring.
* It was mentioned that the community should determine if XML is the correct ingest format for security tools and how we could enable a content generation pipeline. It was also noted that whatever the group does, it doesn’t need to be perfect rather we should aim to eliminate a large chunk of the effort.

## Siemens’ Experiences with ‘Scapolite’, a YAML+Markdown-Based Alternative to XCCDF

Bernd Grobauer (Siemens)

### Summary

There’s a demand for machine-readable security baselines but most organizations are not producing their own SCAP content. Authoring and maintaining content is almost impossible without SCAP v2, which will define standard formats.

### Key Discussion Points

* There was a question as to whether this was just for XCCDF or also for OVAL. It was noted that this was primarily for XCCDF and that Patrick Stockle’s work is more aimed at implementing checks.
* There is a desire to consolidate YAML formats for XCCDF content authoring, but to leave extension points so that organizations can handle their specific use cases.
* It was asked if there were any proposals out of the work to help manage identifiers internally and where XCCDF broke due to adding elements that aren’t included in the specification. To deal with this, when content was imported, the identifiers were changed to the Siemens namespace and unsupported extension points were added along with human-readable check descriptions. It was also expressed that in addition to improving content authoring, there is a need to be able to look at results, examine the original content, and understand what is exactly going on.
* The group was asked whether they represent content authors, tool developers, or results consumers (authors: ~half, developers: ~half, consumers: ~quarter). Given this, the group thought it might be helpful to get more input from results consumers.

## SCAP and AF Mission Planning

David Bricker (Leidos)

### Summary

SCAP is critical to measuring controls during the RMF process. It’s only part of a suite of tools but it’s critical because it provides automated compliance checking. Though there are some content challenges, improvement of scanning/automating compliance checking is a goal.

### Key Discussion Points

* It was noted that sometimes the source is more important than the content. For example, some organizations are restricted to using only DISA-published content, even if other content is available (although, once a source becomes trusted through testing, it might become usable).
* A suggestion was made that with respect to supply chain, vendors should help deliver content, but be able to update it as necessary. Content should also be signed. DISA is planning to start signing their content.
* The need for content creation tools was expressed and that eSCAPe was an attempt, but needed more work. It was suggested that thoughts on updates for this tool be sent to the mailing list and maybe someone can work on it.

# SCAP v2 Governance

### Summary

This presentation discussed the current processes used for SCAP v1 specifications, the challenges associated with these processes, and some improvements that have been made to address these challenges.

### Background

* SCAP v1 specifications were community-driven, but the USG or contractors held editorial control over them. Due to this model, there was often a significant delay between updates and usable technologies. Changes have been made to SCAP, CVE, and OVAL to start addressing these changes. The group needs to determine if similar improvements would be beneficial to the other efforts.

## OVAL Governance

Zack Port (CIS)

### Summary

There’s a need for speed and ease of the creation and maintenance of OVAL schemas. People are needed with domain expertise, knowledge of OVAL, and dedication to improving the language.

### Key Discussion Points

* It was asked if there was a way to know which version of OVAL a tool supports and if big vendors have given up on SCAP. It was stated that CIS has a list of OVAL adopters, but the version is not tracked. Someone mentioned that several vendors in the room all support OVAL 5.11.2, but the best thing to do is just to follow the OVAL GitHub repository to see what is going on and add issues as needed or talk to an OVAL schema supervisor. It was noted that some bigger vendors support integration with available SCAP engines.
* If members of the group know of vendors that might be interested or are in areas aligned with SCAP use cases, they should be encouraged to get involved.

# Governance Overlap, OCIL, and Compliance

David Oliva

### Summary

This presentation discussed the overlapping requirements of compliance regulations and how OCIL could be leveraged in the compliance use case.

### Key Discussion Points

* Given the discussion around the challenges and complexity of OVAL, it was asked if the group should consider similar updates to OCIL to simplify and improve the content authoring experience. It was noted that it might be possible to take a YAML approach since OCIL is a bit simpler than OVAL.
* It was asked if non-automatable (process, qualifications, etc.) checks are being performed with the same frequency as automated checks and whether or not we would just be providing the same information redundantly. It was pointed out that once an answer was provided it could be used to answer other questions although it may not perfectly apply given the nature of the question. It was also noted that you could just store the data in the CMDB and keep in mind that it is likely on the older side.

Day 3 – May 2, 2019

# Day 2 Highlights

Charles Schmidt (MITRE)

### Summary

This presentation discussed the key points made during the previous day’s discussion.

### Key Discussion Points

* One member noted that over the past couple of days there have been multiple presentations discussing varying versions of XCCDF and OVAL YAML formats as well as programmatic APIs for the same languages and that we need to better understand where the programmatic APIs fit into the mix.
* Another member pointed out there were YAML implementations of SCAP capabilities and it’s probably best that we support what’s already in XML and have both efforts go forward in parallel and see if it makes sense to move away from XML in favor of YAML.
* It was mentioned that the tailoring aspects of XCCDF needs more guidance in SP 800-126. XCCDF provides many optional capabilities, but SP 800-126 does not say much about how to use them other than if you have an example like this it should work. Compliance to the specification is reduced to does the tool process the examples in the test suite and if it does then it is all set. The specification needs to say more about how the tailoring element works or re-think what it is for and how badly it is needed. Another member mentioned the complexity and challenges of tailoring were mentioned on previous calls and need to be on the radar moving forward.
* One member suggested that the group needs to consider the lack of extensibility in some of the SCAP languages such as XCCDF where it’s not terribly extensible. As part of the governance conversation, the group needs to consider how to incorporate extensions into the larger SCAP umbrella as well as where more extensibility could be useful. For example, some members of the community are trying to get remediation and configuration capabilities to work well in XCCDF and have done it, but have struggled a little bit with that process. Providing extensibility will help us work towards the future. It was also noted that the great balancing act for standards is extensibility to allow for innovation, but not make it overly complex and hurt interoperability.

# Software Asset Management

Dave Waltermire (NIST)

### Summary

This presentation discussed what software asset management is, how it is used in SCAP v1, and how it could be leveraged moving forward in SCAP v2.

### Background

* Software asset management is the effective management of software assets (e.g., inventory, patching, licensing, etc.). Currently, in SCAP v1, software inventory information is used to map to vulnerabilities with CVE, determine applicable security checklists, and check compliance against whitelists/blacklists. In SCAP v2, software information could also be used to support patch tracking and library tracking.
* SCAP v1 uses CPE to name software and OVAL to identify software installed on an endpoint.
* There are several concerns about SCAP v1 software asset management specifications including a single authority for CPE names which can be a bottleneck, CPE being unable to specify patches and libraries, CPE being unable to capture detailed metadata about software, and the mechanisms for checking installed software (OVAL) is separate from the naming structure (CPE).

## CoSWIDs and SWID Tags

Dave Waltermire (NIST)

### Summary

SWID Tags enable software metadata provided by vendors, platform-neutral, standardized software inventory, integration of data, and automation supporting risk-based management. SWID Tags provide identification for installable software releases and directly support software inventory.

### Key Discussion Points

* One challenge, noted by a vendor, is that submitting CPEs to NIST is difficult because it is a manual process (i.e., must submit one CPE at a time). It was noted that exposing APIs could help address this challenge.
* Another challenge is the fact that it is not easy for two different parties to consistently create the same CPEs for the same software indicating that it is not a good identifier.
* It was noted that integrating SWID tags into software development release cycles can be challenging and can take a significant amount of time. Also, that McAfee, IBM, Microsoft, and Red Hat are creating SWID tags.
* A question was asked if a SWID tag extensions could be used to indicate updated anti-virus definitions. They can be used if they are separate. It was also asked how SWID tags are handled for on-demand projects like Maven. It was explained that there are challenges, but could be done. PKI can also help verify SWID tags.
* Further consideration is needed on how SWID tags will play a role in virtualized environments where the software may be distributed (e.g., Office365 used by an endpoint). SWID tags can also help developers in the DevOps pipeline by allowing them to use SWID tag information to determine their vulnerability footprint.
* While SWID tags offer several improvements over CPE and OVAL, there was concern as to whether or not it was realistic for vendors to create SWIDs when they were not creating CPEs. It was noted that SWID tags support other use cases that might make them more appealing than CPEs. For example, SWID tags can help vendors and end users determine the exact version of the software when handling technical support issues.

# DevSecOps and SCAP v2

Dave Waltermire (NIST)

### Summary

This presentation discussed the current state of DevSecOps and how it fits into SCAP.

### Background

* Organizations plan and develop software and systems, test, deploy, and add security afterwards, but, are more recently moving towards building security into development, test, and deployment processes known as DevSecOps.
* SCAP can be leveraged in DevSecOps to validate the security posture of endpoints hosting applications as well as be used to create mappings between software libraries and vulnerabilities using SWID and CVE which could help inform library selection, identify vulnerabilities, and determine applicable content for detection and mitigations.

## DevSecOps Overview

Aaron Lippold (MITRE)

### Summary

Developers are trying to manage a lot of information. Identifiers need to get built into the process and get results in a standardized format. Development operation workers don’t have security as their top priority. Security needs to be taught as part of the job.

### Key Discussion Points

* Through the DevSecOps pipeline, developers are trying to manage lots of information which is changing very rapidly in applications and it gets more complex when you talk about the containers and platforms. It was asked if this would impact the CPE and build process. It was explained how these identifiers need to get built into the process and get results in a standardized format. Then, with this data and results, can judge risk.
* It was noted that the people working in development operations aren’t always keen on security and more has to be done to have security taught as part of general computer science curriculums. The group agreed with this and that it should really just be considered good development and engineering practices.

## Towards Deriving Automated Implementation & Verification Mechanisms from a Single Machine-Readable Requirements Specification, using Windows Hardening as Proof-of-Concept

Patrick Stöckle (Technical University of Munich)

### Summary

Towards Deriving Automated Implementation and Verification Mechanisms from a Single Machine-Readable Requirements Specification, using Windows Hardening as Proof-of-Concept. Natural language processing is used to generate machine-readable checks as GPOs which can be passed back and forth. There’s a desire to have CIS/IASE specify required GPO settings in a machine-readable way and to use these machine-readable GPO settings.

### Key Discussion Point

* It was asked if they were using an intermediate language during transformation and what is the process for tailoring that content. It was noted that changes could be made through the GPO editor.
* It was asked, as teams evolve and consume this check content how is it shared and how are contributions managed. It was explained how natural language processing is used to generate the machine-readable checks as GPOs which can be passed back-and-forth. Otherwise, if people are copying-and-pasting code, inconsistencies and errors may arise.

# Using ROLIE to Advance SCAP Reporting

Tamelia Hutchinson (Pivotal)

### Summary

This presentation discussed the use of ROLIE for SCAP Reporting in a DevSecOps environment.

### Key Discussion Points

* A comment was made noting that based on the information being put in the repository, it sounds like a metadata service and CVEs are being decoupled. There could be differences between platforms for development versus production. It was asked if there was a place for this. It was explained that from an operating system perspective, it may not necessarily be applicable, but you would still need to provide application lists. It was noted that this was just an example and you could change the data to whatever you want.

# Applicability Language

Charles Schmidt (MITRE)

### Summary

This presentation discussed how applicability information is currently used, what it could be used, and some initial requirements for an applicability language.

### Background

* Enterprises often times must gather automation content (patches, checklists, advisories, etc.) then manually determine what content applies to endpoints deployed on the network or using CPE and OVAL.
* An applicability language could help automate the selection of endpoints based on standardized criteria and could serve as a common querying language for CMDBs.
* Initial requirements for an applicability language include:
  + Support for using arbitrary endpoint information
  + Support for different information formats (SWID, OVAL, CPE, etc.)
  + Standardized applicability information to enable automation
  + Manifest for a simple approach to determine applicability of content

# The Applicability and Query Abstraction (AQuA) Language

Stephen Banghart (NIST)

### Summary

There is a need for a standard applicability language so that a single party can write applicability statements that can be consumed across many tools. AQuA improves shortcomings of the CPE Applicability Language, supports SWID tags, and can be executed against an arbitrary CMDB.

### Key Discussion Points

* Given there seems to be some implicitness to the endpoint and its relationship to the software or container in virtualization, but also tends to give a version of software as the underlying operating system. It was asked if this could address the nested endpoint situation. It was answered that it could handle this situation given a good ontology which can be imported into applicability statements.
* Given you are collecting data and putting it in the CMDB and then using the applicability language to query it, you are decoupling collection and evaluation. It was asked if this meant OVAL wasn’t needed. It was explained that, yes it is one possibility, but the scope of this effort within SCAP needs to be determined.
* It was discussed how the applicability ontology could be vendor provided, third-party, etc. It was noted that during development, the team looked at SPARQL, SQL, and XQuery, but wanted something that was simpler that could be transformable into other query languages.
* It was noted that there is a design document, specifications, and a reference implementation available and the timeline for AQuA depends on where the community wants to go with the work. Several members of the group expressed interest in this work.
* It was mentioned how AQuA was reminiscent of Facebook’s OS Query where every system is modeled as a database table that you can query. This approach is good for some types of queries, but not for others where OVAL might be better. It was pointed out that the ability of using AQuA to perform these types of evaluations depends on if the correct data is in the right place. It was also noted that using AQuA to support queries for vulnerabilities for NVD would be fairly straightforward and maintainable whereas that might not be the case if you have to write your own queries.
* It was highlighted that the ontology is meant to be open-world and could be based on existing data models such as OVAL and YANG. Furthermore, the ontology enables spanning across other namespaces and nested virtualization.

## Toward a Hierarchical and Extensible Applicability Language

Bernd Grobauer (Siemens)

### Summary

Applicability metadata is a prime candidate for format extensions. The definition of an extension should include semantics of how applicability information on one level transfers to other levels and how different levels interact with each other.

### Key Discussion Points

* It was noted that organizational metadata about devices such as criticality, role, organizational owner, etc. needs to be considered. Furthermore, the group should consider making SCAP v2 more enterprise-focused rather than endpoint-focused and understanding the similarities between endpoint types.
* Further investigation should be done to determine if applicability should be modeled as a hierarchy, matrix, or graph. Once the model is determined the group must ensure the data can be pieced back together to get back to the endpoint case.
* While a CMDB is a good place to query this information, only certain organizations of a certain size and maturity will be able to have one.