Endpoint Data Collection teleconference Aug 29, 2019

Attendees: Bill Munyan (CIS), Adam Montvale (CIS), Charles Schmidt (MITRE), Joe Sain (MITRE), Christine Argenio (MITRE)

**Discussion of responses to the whitepaper Jessica posted to the mailing list.**

Charles started off by summarizing prior discussions from the mailing list:

* Jessica had posted a whitepaper that captured an early vision for components that would support endpoint data collection.
* David had expressed concern that some elements were overly prescriptive, especially with regard to the role of endpoints and their relation to Posture Collection Servers.

Charles opined that both parties were expressing compatible visions, but that the role of the Posture Collection Server was underdefined and the confusion came from assuming a more limited restricted role for this component than was intended. In particular, Charles felt the Posture Collection Service (PCS) role was simply intended to be the interface that other parties queried to initiate a collection of data from endpoints, but that this request could be fulfilled in many different ways (including but not limited to direct interaction with specific endpoints).

Bill shared many of David's concerns, but was happy with the idea of more flexibility in collecting data. He suggested that there could be multiple PCSs, some on endpoint and some separate from the, all of which provided a channel by which endpoint information could be gathered and sent to a CMDB. Bill did not like the idea of a Posture Collection Engine (PCE), feeling that it was to bound to the idea of endpoint agents, and instead the idea should be subsumed by the PCS.

Charles said the one thing he wanted to tease out is the invocation of a collection vs. the actual collection and deliver of the data. If PCS is distributed with services on endpoints, cloud, etc, the invocation needs more info because they need to know which to contact. The vision in white paper architecture had PCS as a single point for all requests.

Bill suggested that a Posture Collection Manager (PCM) could serve as the single point of query, orchestrating what info gets communicated to PCS which in turn holds subscriptions to the data being collected. One would query the PCM, which would contact the appropriate PCSs, which would gather the data and deliver it to a CMDB.

Bill said the document Jess provided came out of a discussion between him and Adam at IETF. There was cutting from a different document so it’s a rough draft. Within there, the graphics were a little off but there were references to other components that were being replaced. PCS communicating to PCE on endpoint, that’s where the point of contention came up – difference of opinion.

Charles said he knows the concept of a standardized agent on endpoints to collect data is important to Jessica. He felt that we shouldn't eliminate the PCE because there’s a desire to try to unify the disparate agents on endpoints providing dribs and drabs. The key thing to note, however, is that the PCE is not the ONLY source of information and other components, potentially not located on the endpoint, would also be supported. Thus it would be possible to pursue Jessica's vision of a standardized endpoint agent that could support many vendors' tools, but doing so would not preclude other services for data collection.

Charles said nothing incompatible of these visions, just need to work on terminology. Bill said it’s probably just a matter of adding a few extra diagrams. You could have a diagram showing here’s how it would work in agentless capacity, or using an agent, or aggregating info from other services. It would clarify different people’s perspective on what that service should look like. Charles agreed and asked if Bill could something together that captures that? Bill said he would try.

**Concerns about trust in data collection.**

Charles said his concern remains as to what extent vendor A would be willing for their tools to do evaluations and give recommendations (some of which might lead to changes to system settings) based on data collected by Vendor B? He noted that, in SCAP v1 discussions, some security tool vendors stated that they were unwilling to run SCAP content they didn’t create or validate because if something went wrong, users would blame the tool rather than the content.

Adam agreed that this was a point of concern, but also noted that he had been talking to vendors and that some vendors are already ingesting data from third party tools. He noted this was happening even in cases where there wasn't a formal relationship between the data-gathering vendor and the data-consuming vendor. As such, he felt that it might not be too hard to convince vendors to adopt this model of separating collection and evaluation.

**Validation**

Bill noted that previous validations of SCAP have focused on validation of correctly ingesting content and producing the correct output. He asked if this was going to be kept around, noting that SCAP v2 incorporates both data collection and data evaluation pieces. In particular, he was interested in this given that the content ingestion and content evaluation are now effectively happening in different components in the architecture. The decoupling of components creates new challenges that were not present in SCAP v1 when this separation was not envisioned. Adam wondered if the solution was to support validation of components separately, but Bill noted that there might be a need for more architectural-level validation as well.

Charles said I don’t know what a NIST validation would look like. He did note that the SCAP 1 validation program quickly learned that it was beneficial to subdivide the validations to more precisely align with component roles and he felt that component-level validation seemed like a reasonable approach.

Charles also noted that the new SCAP v2 architecture has a far greater dependency on common understandings of semantics. He noted that in SCAP v1, since a single vendor extracted and evaluated data, they could tweak this process in unique ways as long as the output was consistent with language expectations. With the SCAP v2 separation of roles, we’ll need a greater degree of uniformity so different tools can correctly interoperate in what was previously a single tool's role.

**Workshop preparation**

The group was asked what would be good topics for the sub-team's 90 minute slot on day 2 of the upcoming SCAP workshop. Charles suggest that that the updated flows that Bill was assembling would be a good topic to get broader input on.

Adam noted that these flows are very similar to the ones described in the IETF SACM architecture. He noted that we probably don't want to duplicate or compete between the two groups. He asked if there were thoughts on how best to coordinate. Charles was not opposed to focusing to having the work done in the IETF venue but noted that it probably won't be practical to move all work to either venue – shifting entirely to either venue would likely lose inputs. Charles wondered, if the main work effort was moved to IETF, would it be possible for the SCAP community to meet independently and provide input that would propose for inclusion in an IETF document, or would that be seen as a violation of IETF transparency rules. Adam suggested that, as long as it was transparent where suggestions came from, he felt t would probably be ok.

Charles suggested that, as a first step in this discussion, it could be useful to present a mapping between the current SACM architecture status and current thoughts on the SCAP architecture. This would help the group decide if closer collaboration makes sense. Bill said that he would put together some notes on this mapping. Charles would then turn this into a presentation and brief it at the workshop.

**=== ACTION ITEMS ===**

Bill – Create a set of expanded flow diagrams for SCAP data collection to outline the multiple possible flows from query to upload of data into a CMDB.

Bill – Create a rough mapping of SCAP architectural elements/flows/terms to corresponding aspects of the SACM architecture.

Charles – Turn Bill's mapping into a presentation for the SCAP workshop in September.

Jessica – Prepare and deliver the sub-team summary briefing for day 1 of the SCAP September workshop.