CNE Presence in CT[10] Status Report

Authors:

[redacted], SNIP Intern

[redacted], 1LT

(U//FOUO) Executive Summary

(S//SI//REL USA, FVEY) This paper discusses the work completed as of 1 April 2008 on the inclusion of CNE data into the CT[10] Pilot Project. CT10 engaged two fronts in this effort. First, was the acquisition and alerting of active SIGINT into CT[10]. Second, was the design of an extension to the current CT[10] data model to support further analytic development of CNE data. As a result of this work, CT[10] now supports alerting of the establishment of new TAO points of presence on the Internet with respect to certain exploitation methodologies. Additionally, this paved the road with internal and external partners for the continued development of the CT[10] CNE data model extension.

(S//SI//REL USA, FVEY) Specific conclusions and recommendations are contained at the close of this paper. However, three key points underpin a majority of the information in this paper:

- (U//FOUO) Priority support to CT[10] data acquisition efforts.
- (U//FOUO) Further building a DNI/CNE analytic and development team.
- (S//SI//REL USA, FVEY) Development of a strategic vision for use of CNE data in CT[10].

(U//FOUO) TAO Data Flow Acquisition

(S//SI//REL USA, FVEY) The first step in processing CNE data is to identify and acquire CNE data flows in order to route this information into CT[10]. This work presents its own unique challenges,, as described in the next section.

(U) Current State

Derived From: NSA/CSSM 1-52
Dated: 20070108
Declassify On: 20291123

(S//SI//REL USA, FVEY) CT[10] is ingesting a single flow of CNE data, specifically FOXACID log files. This was identified as a top priority by CT[10] customers GHOSTWolf, therefore it was chosen to be the first CNE data flow to pursue. After mapping the flow of FOXACID data throughout NSA, CT[10] established a flow of this data from Protocol Exploitation starting in the middle of February 2008.

(S//SI//REL USA, FVEY) Because the CNE data model extension had not been implemented at this time, the only way to process FOXACID logs was to fit the information into the current CT[10] data model. This was possible primarily because of the natural flexibility in the current model. It was easy to define a new "event type" within the CT[10] ingestion system, and populate events with all selectors pulled from the each FOXACID event. Another benefit of the current model design was that it became easy to define a new selector type for an implant ID and easily integrate it into the system. This allows implant IDs to be added to watch lists and as target selectors, so that CT[10] users can track any events that are associated with a defined set of implants. Therefore, if CT[10] acquired a data flow containing implant callback information, it would be easy to add this into the current data model, giving analysts additional capability to follow implants deployed through FOXACID..

(S//SI//REL USA, FVEY) As a result of this work, the CT[10] system is now alerting on FOXACID exploitation activity through a suite of alerting technologies. This includes geo-locational tipping in Geo[T], and alerting through Agent Logic (IRC and e-mails) and the iSpace dashboard.

(S//SI//REL USA, FVEY) One conclusion reached from developing the FOXACID ingestion code is that the current CT[10] data model can support more CNE data than previously believed. As new CNE flows are introduced into the system, it may be possible to easily add them to the current data model if the extension is still not implemented.

(U//FOUO) Near and Long Term Goals for TAO Data Flow Acquisition

(S//SI//REL USA, FVEY) The near term goal regarding CNE data flow into CT[10] is to identify additional CNE data flows that build off of the FOXACID data. Based on the requirements of GHOSTWolf, the immediate next step would be to acquire implant callback data flows. This would allow analysts to track when their implants are calling back to listening posts for tasking and exfiltration. The implants in question are VALIDATOR, OLYUMPUS, UNITEDRAKE, and STRAITBIZZARE.

(S//SI//REL USA, FVEY) After establishing this data flow, the next step would be to ingest collected data from the implants. This includes any exfiltrated files, and implant plug-in results. With each new data flow, CT[10] analysts will have a more complete picture of CNE activity related to their targets. However, these types of data do not naturally fit into the current data model, which is why the proposed extension is so important.

(S//SI//REL USA, FVEY) Additionally, there are other CNE datasets not created by TAO that would be useful to CT analysts. GOLLUM, a partner implant, is one such data set that GHOSTWolf has indicated that would be useful to ingest into CT[10]. RADIUS logs (ISP dial up customer records) are also an excellent source of information, and create a natural link between DNR and DNI datasets.

(U) Challenges / Way Ahead

(S//SI//REL USA, FVEY) There are a number of challenges for securing further CNE data from TAO. First, given the complexity of TAO data flow, locating the ultimate data owner is challenging. Though everyone in TAO CT[10] has interacted with on data flow issues has been extremely helpful, no new CNE flows have been established. This is largely due to the fact that CT[10] is not imbedded within the requirements generation process in TAO. Additional transparency on TAO data flow requirements will greatly ease this issue.

(S//SI//REL USA, FVEY) One other issue that would greatly help other systems in using TAO data, is a universal marking and classification of CNE data. In the case of FOXACID, there are no native classification markings on the files, and compartmented targets as well as FISA data have to be stripped out by TAO before data is sent on to external customers such as CT[10]. As CT[10] becomes PL3 accredited, it becomes critical that CT[10] ingests full FOXACID records. Unfortunately, the lack of a unique marking and classification for each file creates additional challenges that complicate the oversight, classification and policy surrounding CNE data. Furthermore, additional discussions with TAO have indicated that this may be an issue for other data flows as well. The addition of classification markings to TAO data feeds would greatly assist all consumers of TAO data with maintaining the proper security procedures.

(U//FOUO) TAO engineers have worked towards creating an RDF database system that would house all of their information, and make it available to external customers. As this capability is developed, CT[10] could benefit greatly from becoming a consumer of this data. This system would address a number of issues, including the marking and classification challenge addressed in the previous paragraph.

(S//SI//REL USA, FVEY) In summary, the way ahead is to address these existing challenges in order to effectively push more TAO data flows out to analysts in near real time. CT10 should also remained synched with TAO's effort to develop and deploy the RDF database. Finally, CT10 needs to implement the data model extension and move any currently ingested TAO data into the new model.