

Incorporating Flow-Based Behavioral Analysis Inside Agency Networks

by Frank Doane

Abstract

The US Government is embarking on an ambitious endeavor that will substantially change the way government networks communicate with the outside world and how they are protected from external threats. The reduction of Internet gateways through the Trusted Internet Connection (TIC) initiative, and the standardization of threat management and incident response through development of core requirements for Trusted Internet Connection Access Provider (TICAP) capabilities and the management of in-cloud security by the United States Computer Security Readiness Team (US-CERT) under the EINSTEIN program, ushers in a new era of unprecedented inter-agency collaboration and network consolidation for our government's federated networks.

One of the keys to success for this effort is for information technology (IT) security professionals at all levels of the enterprise (within the CERT and TICAP down to the individual owners of systems) to be able to detect the presence of malicious users connecting to government systems, to communicate details about the nature of the attack and its source to the organizations responsible for carrying on further investigations of the event, and ultimately to be able to respond to bulletins notifying them of events that have had an impact on the networks they are responsible for protecting despite having missed the attack in the first place.

To meet this challenge in the current world where a zero-day threat is an everyday reality, agencies need an always-on technology that constantly monitors the activities of their own users and outsiders touching their networks. This can be used to detect anomalous and out-of-policy system activities to spot the presence of an attack early and make sure a record of each communication is preserved to support the investigation of events that are only

of the traffic they observe back to a collector for analysis. Network Behavioral Analysis systems are expert systems that process flow records into conversation-level log files that record the fact of a networked conversation and write the records to a flow table analogous to a telephone bill—showing who talked to whom, when, through which ports and protocols, and how much traffic passed during the conversation. These records are

To meet this challenge in the current world where a zero-day threat is an everyday reality, agencies need an always-on technology that constantly monitors the activities of their own users and outsiders touching their networks.

apparent to the community protecting the network after the fact. Such capabilities can easily be added at all layers of the government enterprise by bringing flow-based network behavioral analysis into the technology architecture of the agency.

Flow works by leveraging the switching and routing infrastructure as a virtual surveillance grid. When remote flow collection technologies such as NetFlow™ and sFlow® are turned on, infrastructure components report records

read over by an engine that builds a baseline of normal behavior for each system observed to have touched the monitored network. The system generates alarms on anomalous activities, keeps a record of all activities of each connecting host, and provides a graphical user interface to speed the time analysts require to isolate information about conversations relevant to an investigation.

Other components are utilized to provide the additional information



necessary to speed investigation. This can include username and Media Access Control (MAC) address or hostname translation to IP address through integration with directory stores and DHCP servers; communication via SNMP and syslog with other reporting systems to pull in additional data for correlation and gain more information about the network and hosts being monitored; and integration upstream to Security Information and Event Management Systems (SIEM) such as ArcSight or Enterprise Network Management (ENM) such as HP OpenView to leverage Behavioral Systems ability to compress flow records and isolate relevant communications to aid the use of data collected in a broader Service Oriented Architecture (SOA).

This article surveys the capabilities of both the current tools used to protect government networks leveraging external monitoring programs (EINSTEIN and Centaur) and the capabilities of enterprise Network Behavioral Analysis (NBA) systems; and explores areas where use of a NBA, such as Lancope's StealthWatch, can enhance the ability of IT Security teams to detect suspicious behavior/zero-day threats and speed the incident response process in conjunction with EINSTEIN and Centaur to improve the Federal government's security posture under the emerging world of the National Strategy to Secure Cyberspace. The intended audience is anyone involved in operational security within the Federal government, to include

US-CERT, Joint Task Force—Global Network Operations (JTF-GNO), TICAPs, Computer Network Defense Service Providers (CNDSPs), and agency-level network security and network operations personnel as well as those responsible for providing resources, technical guidance, and oversight to these communities.

What additional benefits would deploying NBA systems inside agency networks offer over and above the current capabilities for threat monitoring and response provided by EINSTEIN and Centaur?

The attacks we are discussing are primarily leveraging the involuntary recruitment of systems across the Internet via the infection of hosts who are invited or tricked into downloading exploit code—which in turn enables a controller to remotely access the owned system and delegate it to carry out tasks to accomplish the missions of the larger attack without detection. So the activities of both the coordinating controller and the infected systems will ultimately traverse the Internet for the success of the overall attack. Are there any advantages to watching inside the perimeter of the private network to find these attacks?

Yes. The reasons for adding NBA inside the agency networks as a defense-in-depth strategy are vital to the overall objectives of securing Federal government networks from current threats, and include—

1. Providing coverage for all communications traversing internal or private network space; eliminating missed attacks leveraging networked backdoors; and providing a comprehensive forensic record that shows not only the full extent of the compromise by picking up the presence of other recruits inside the network, but contains a record of any file transfers or other exfiltration events necessary to understand the impact of the event.
2. Earlier attack detection to leverage the predictability of host behavior as viewable only from the inside of the network to tune behavioral anomaly alarms that detect a single instance of internal host compromise and act as a “spotter scope” for teams in higher tiers conducting a comprehensive forensic analysis of widespread attacks. In addition to anomaly-based alarms, NBA systems offer customizable policy enforcement that acts as a trip wire to catch the presence of compromised systems by enabling agency analysts to define rules of the road for network hosts. These rules are often violated when a compromised system is remotely controlled by a party unfamiliar with established policies.
3. Increased speed of incident response by providing internal agency personnel responsible for

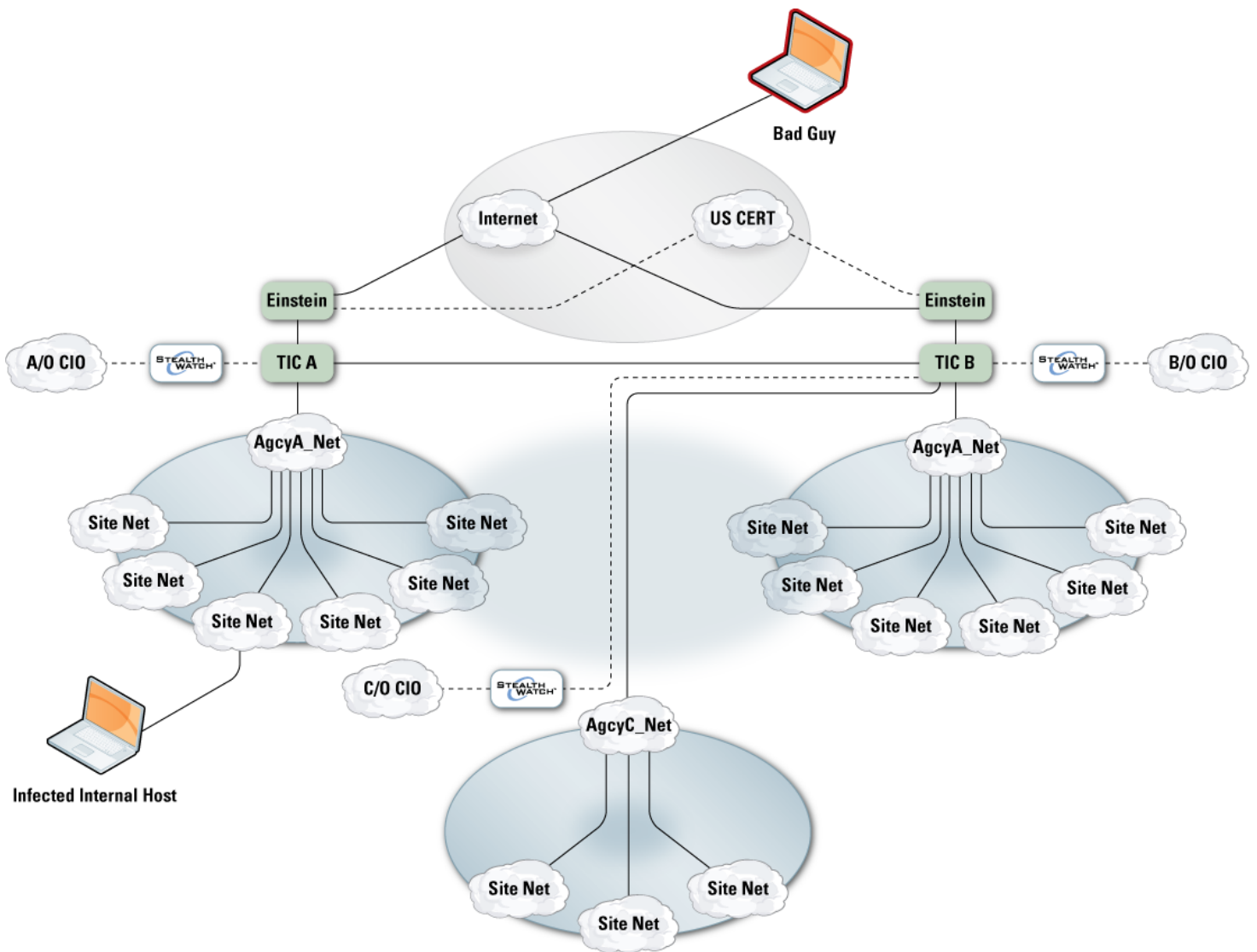


Figure 1

coordinating with the TICAP or CNDSP or the US-CERT or JTF GNO the ability to perform their own flow-based analysis of an attack touching inside the network to pinpoint the location and identity of the host actually responsible—a task that is impossible given the obscurity that NATing firewalls (firewalls performing Network Address Translation) and proxying devices create for higher-tier analysts.

Expanding Coverage Down to the User

The drawing below shows a high-level view of proposed coverage for the entire system with areas of responsibility broken out between agencies, TICAPs, and US-CERT. A similar model applies in the

Department of Defense (DoD), with JTF-GNO, CNDSPs, and component agency networks acting at each level.

The drawing shows flow-based monitoring responsibility broken into two broad categories—above the TICAP inside the Internet, and below the TICAP inside government networks. Other monitoring systems provide visibility beneath the TICAP, and many mandated security controls ensure that systems attached to these networks meet a minimal standard for risk. What is missing when flow-based analysis is not applied at this lower level of infrastructure is visibility into host behaviors required to understand the normal operating conditions of the network—who connects to which systems how frequently, and the ability to leverage

this knowledge to understand when a threat is present on an authorized system that has somehow managed to hide from the various systems protecting the hosts and networks on which they reside.

Adding flow-based monitoring at this level would provide a pervasive forensic trail to track any behavior of an internal system involved in a threat event and fully scope the compromise. It would also serve as an early warning system for threats that evade current controls, such as the walk-in threat where a system is compromised by physical access either through addition of a rogue system or the porting of malicious code through a removable drive; the mobile threat of laptops, personal digital assistants (PDA), or other systems that do not always reside on the protected network

and are thus susceptible to compromise while detached; and holes created by misconfigured systems such as networked backdoors, open wireless connections, bridging of network segments, rogue applications, or firewall ports left open.

Operationally, once flow is collected at this level the flow records are easily shared with analysts at higher levels. The key is to collect flow once and provide interfaces into the data to as many people as may require access, and establish the ability to limit access to only the data those users have reason to access. NBA systems, such as StealthWatch, accomplish this through role-based access permissions that limit analysts' access to data through functional roles defining what they can see of the data collected and data roles that control which hosts or network segments they have purview over. In addition, NBA systems include the ability to pull flow records from lower-level systems into aggregating systems such as SIEM or other reporting tools *via* a north-facing Simple Object Access Protocol (SOAP), Extensible Markup Language (XML)-based Application Programming Interface (API). This enables data to reside in the agency system until it is needed to support an investigation by a TICAP, CNDSP, US-CERT, or JTF-GNO, at which point only records relevant to the investigation would be pulled. Having this just-in-time pull capability provides the ability for systems to retain local storage of flow data until required for analysis, which mitigates some of the impact of bandwidth overhead required to support flow analysis throughout the infrastructure. It also mitigates privacy issues, and concerns over who owns the role of protecting the local network segments, the agency, or the service provider.

The benefits of flow to quickly pull together the facts of an event and isolate the individual hosts involved should be brought into agency networks and shared up to the TICAP analysts responsible for protecting these networks. Without this data, much of the work involved in investigating threats to validate that they are indeed events and extrapolate the full nature of the compromise involves

arduous collection of data from the owners of various systems logging some part of the event—manually culling through logs and relying on guesswork for the parts of the picture left incomplete. NBA systems are ideal for this mission, as shown in the next section.

Integrating EINSTEIN and Centaur Outside Agency Enterprise Networks with Commercial NBA Inside to Turbo-Charge Government Detection and Response Capabilities

The Spotter Scope—Speeding the Time Needed to Detect the Existence of the Threat in the First Instance

NBA systems detect threats through deviations from normal behavior by networked hosts and through the detection of network policy violations. The “trip wire” policy enforcement capabilities of an NBA system monitoring internal networks allows agency security architects to design a set of “rules of the road” to which networked hosts must adhere when using agency resources. Violations of these rules will often flag the presence of a threat, since they often serve as a check of control systems such as firewalls to prevent known bad actions from occurring. For example, a firewall policy disallows communications of an internal host to a server farm asset across a certain port. A network administrator opens a bridged connection to the server farm subnet from the user subnet, and leaves the connection open after completing his work. The internal user gets “owned” through the exploits of a hacker whose reconnaissance of the network uncovers the backdoor. The hacker uses the backdoor to communicate with the server, bypassing the corporate firewall. The fact that a two-way communication is established between a server in the server farm and a host on the user subnet is alarmed against by the NBA system as violating established policy—and during the incident response, the hacker’s presence on the network comes to the attention of security personnel.

Discovering this event *via* current controls would be difficult because the

firewall never saw the traffic; the routing and switching infrastructure is not designed to report these events; the server did not realize the policy violation had occurred; and whatever permissions were used to gain access were probably accredited. Leveraging the flow-logs of the NBA system, the security team could then report this event to US-CERT or JTF-GNO, or to the TICAP or CNDSP responsible for monitoring their specific network. Once the hacker’s actions have been validated at that level *via* a log file check for systems reaching external network systems, the full extent of the hacker’s activities could be quickly assessed through queries into the flow files stored and managed by the CERT and JTF-GNO.

In the case of anomaly events, the key role of the NBA system is again to find the original malicious actor and leverage it as a starting point at higher levels to investigate the full scope of the attack. NBA systems are particularly good at using non-deterministic means of detecting threats. Instead of applying deterministic rules as in the policy enforcement model, anomaly detection is accomplished through the NBA system profiling every host observed to have touched the network—establishing thresholds for certain traffic characteristics that are tracked as potentially security-relevant (*i.e.*, numbers of syn packets sent without a response, numbers of concurrent sessions established, *etc.*, over periods of time) and then leveraging algorithms that look for patterns in behavior indicative of a network attack or compromise by observing traffic that exhibits these characteristics once the threshold established as “normal” is exceeded for each host. The key to anomaly detection is its “fuzziness” or ability to point at something that appears odd, as opposed to signature-based detection that relies on deterministic coding of rules meant to detect threat events in progress. In the modern network threat arena, the ability to catch the actual attack has been greatly diminished by the use of “pull threats” such as phishing, redirecting traffic to websites through DNS exploits, and other means of

drawing or tricking users to visit a site where they unwittingly load exploit code to their systems. This is in contrast to the old model, where the hacker sought out victims through scanning, then “pushed” the exploit to vulnerable hosts.

Anomaly events look at general categories of behavior consistent with threats across the network, such as scanning, spamming, Distributed Denial of Service (DDoS) and high-volume Domain Name System (DNS) queries. More often than not, these systems are picking up not the actual attacker, but one of its minions of “bots”—infected end systems tasked to participate in the attack by a controller who earlier had malicious code injected into the system through phishing or other recruitment techniques. Systems such as Lancope’s StealthWatch have further reduced the simplicity of detecting the presence of compromised hosts by looking at the aggregate of bad behavior emanating from a particular host without reference to a pattern of attack through index-based alarms that point to net bad actors or net targets of attack.

Once alerted to the presence of a bad actor, an NBA system provides a forensic record of all activities of the host over time whether or not suspicious events are associated with the particular communications involved. It is easy to track the source of infection back to a particular host on the outside merely by expanding the search of flow records back in time and examining suspicious “calls out” to external systems (*i.e.*, communications across IRC ports or at suspicious times of the day).

Once detected, the presence of a bad actor can be quickly validated and fully examined under the microscope of the larger flow-monitoring technologies of the Network Situational Awareness (NetSA) tool suite by US-CERT analysts or JTF-GNO. Having a starting point to launch an investigation and quickly pull together all points in the government impacted by the event is of critical benefit to the analysts at this level. While global event correlation systems such as the

NetSA tools can detect threats by seeing patterns emerge on a global scale and “connecting the dots” of multiple infected systems back to a controller, it helps them enormously to know where to start.

Marrying multiple complementary detection technologies produces greater context around a security incident. For example, a global correlation system will identify the presence of a botnet across a system when multiple hosts from multiple networks—the bot army—are exhibiting similar or dissimilar behaviors associated with threat events, DNS or scanning-based reconnaissance, SPAM production, high volumes of connections to other systems, *etc.*, across the network but are periodically communicating with a single host that they all share in common—the bot controller. This behavior is only viewable once the monitored space is enlarged to gather enough data points to see a certain number of bots. Conversely, an NBA system monitoring component enterprise networks within the system will be able to detect the instance of a single bot engaged in suspicious behavior such as scanning outside the enterprise network, doing DNS queries, sending a high volume of email traffic, talking on IRC channels, *etc.* Having different systems reporting both the big-picture view of the bot army as the whole and the smaller picture of the single infected system in many different instances increases the likelihood that an attack will be caught earlier and minimizes the possibility of the attack escaping detection entirely.

The Hunting Dog—Speed Time of Incident Response

After an event has been validated by US-CERT/JTF-GNO, the next step is generally to announce its presence to the community of security professionals responsible for monitoring agency networks. This has taken the form of bulletins specifying the location of the controller (including URL and/or IP address), details about how the attack was accomplished, and details about the agency systems involved. Once the agency

security personnel receive this bulletin, a new hunt begins. The easiest part is to reconfigure existing control systems such as proxies or firewalls to prevent further contact with the bad external actor. Discovering which internal end-systems are impacted, and isolating them on the network in the present, becomes an arduous task that involves the collection of data owned by multiple internal network professionals and reading through this data to understand where the impacted systems are today to affect remediation actions.

Looking at how this could be accomplished with an NBA system will show how much easier the incident response portion of the investigation can be made when NBA is brought into an agency to analyze internally collected flow. First, a policy enforcement rule can be established on the internal NBA system that looks for attempts to connect out to the controller to pick up systems that were compromised but had not previously reached out. This rule can be established to look for connection attempts instead of completed connections, since the connection will be presumably blocked at the firewall. Second, a quick flow analysis of the flow records for all internal systems can be run to look for any past connections to the malicious actor outside. When that list is returned, it can be further refined by looking at the URL to make sure the hosts involved were communicating with the website involved instead of another website residing on the same IP address, since many are in hosted environments. Once the communication has been validated, the internal system must be isolated. The quickest way to accomplish this is to use the NBA system’s ability to integrate with directory stores and Dynamic Host Configuration Protocol (DHCP) servers to look up the username, hostname, or MAC address associated with the IP address of all internal systems observed to have contacted the external host at whatever point in time their contact occurred. From there, the analyst needs only look up the current IP address for the same username or hostname/MAC address

to identify its current location on the network. NBA systems will often include information about infrastructure devices connecting the user to the network so the actual switch or router and interface information can be isolated given the IP address, and the user can be quickly taken off the network through reconfiguration of the infrastructure devices.

Anyone having experience with running down the information required by a CNDSP, JTF-GNO, or US-CERT today will understand how much time and effort is saved through the process outlined above. The current alternative involves hours or days of collecting log data and manually culling through it for answers, which often leaves the job half-completed. ■

About the Author

Frank Doane | is the Federal Sales Representative for Lancope, Inc. Mr. Doane is an attorney in the state of Virginia, a graduate *cum laude* of George Mason University School of Law and holds a BA in History from Knox College in Galesburg, IL.

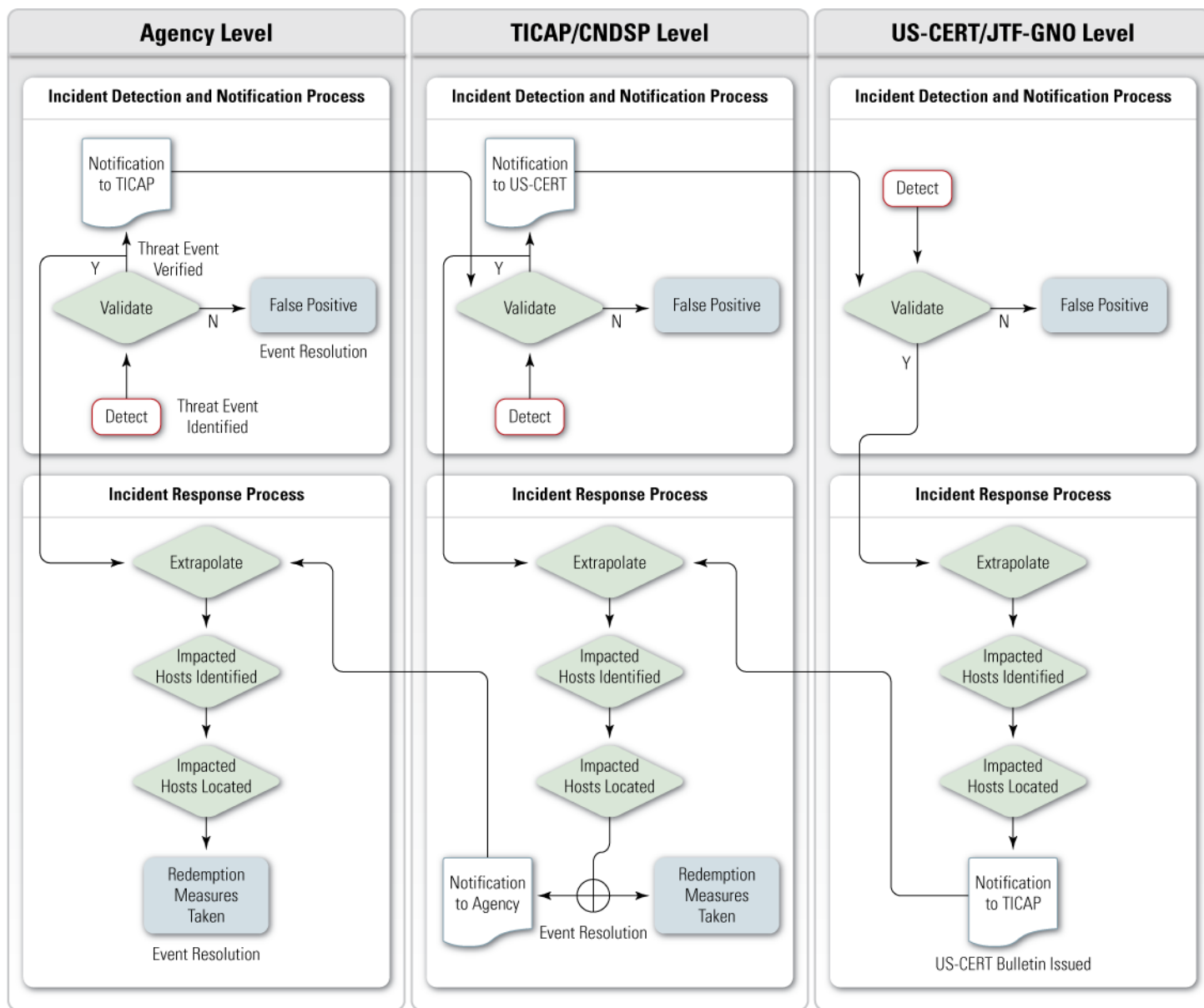


Figure 2 Putting it all together—proposed model for integration between US-CERT/JTF-GNO – TICAP/CNDSP and downstream agency security professionals leveraging NBA on agency networks