React Faster and Better: New Approaches for Advanced Incident Response

Version 1.4

Released: April 15, 2011

Securosis, L.L.C. 515 E. Carefree Blvd. Suite #766 Phoenix, AZ 85085 **T** 602-412-3051 info@securosis.com www.securosis.com

Author's Note

The content in this report was developed independently of any sponsors. It is based on material originally posted on the Securosis blog, but has been enhanced, reviewed, and professionally edited.

Special thanks to Chris Pepper for editing and content support.

Licensed by NetWitness

NetWitness® is a revolutionary network monitoring platform that provides enterprises a precise and actionable understanding of everything happening on the network. NetWitness solutions are deployed in customer environments to solve a wide range of challenging information security problems including: insider threats, zero-day exploits and targeted malware, advanced persistent threats, fraud, espionage, data leakage, and continuous monitoring of security controls. NetWitness is a privately held, venture backed company that has achieved and maintained profitability since mid-2008. Our customers include enterprises in the Global 1000 representing sectors such as financial services, power and energy, telecommunications, retail, and high-tech, as well as government agencies in defense, homeland security, law enforcement, and intelligence.

Copyright

Table of Contents

Introduction

New Responses for New Threats

Incident response is near and dear to our philosophy of security — it's impossible to prevent everything (we see examples of this in the press every week), so you must be prepared to respond. *The sad fact is that you will be breached.* Maybe not today or tomorrow, but it will happen. We have made this point many times before (and it has even happened to us, indirectly). So response is more important than any specific control. But it's horrifying how unsophisticated most organizations are about response.

This is compounded by the reality of an evolving attack space, which means even if you do incident response well today, it won't be good enough for tomorrow. We don't want to bog this paper down on the basics (covered more fully in our Incident Response Fundamentals series), so let's review those quickly because they are still an essential foundation.

Organization and Process

First and foremost, you need to have an organization that provides the right structure for response. That means you must have a clear reporting structure, focus on objectives, and be flexible (since you never know where any investigation will lead). You need to make a fairly significant investment in specialists (either in-house or external) to ensure you have the right skill sets on call when you need them. Finally you need to make sure all these teams have the tools to be successful, which means providing the communications systems and investigation tools they'll need to quickly find root causes and contain damage.

Data Collection

Even with the right organization in place, without an organizational commitment to systematic data collection, much of your effort will be for naught. You want to build a data collection environment to keep as much as you can, from both infrastructure and applications/data. Yes, this is a discipline itself, and we have done considerable research into these topics (check out Understanding/Selecting SIEM and Log Management and Monitoring up the Stack). But even with a lot of data, there isn't as much useful information as we need to pinpoint what happened and figure out why.

Before, During, and after the Attack

We also spent some time in the Fundamentals series on what to do before the attack, which involves analyzing the data you are collecting to figure out if and when you have a situation. We then moved to next steps, which involve triggering your response process and figuring out what kind of situation you face. Once you have sized up the problem, you must move to contain the damage and perform a broad investigation to understand the extent of the issue. Then it is critical to revisit the response in order to optimize your process – unfortunately this aspect of response is often forgotten.

It's Not Enough

Yes, there is a lot to do. Yes, we wrote 10 discrete blog posts that barely cover the fundamentals. And that's great, but for high-risk organizations it's still not enough. And within the planning horizon (3-5 years), we expect even these fundamentals will be insufficient to deal with the attacks we anticipate. The standard way we practice incident response just isn't effective or efficient enough for emerging attack methods. If you don't understand what is possible, spend a few minutes reading about how Stuxnet seems to really work, and you'll see what we mean. While the process of incident response still works, our implementations need to change.

Incident Response Gaps

Today's incident responders are challenged on multiple fronts. First, attacks are significantly more sophisticated and carefully tuned than commonly discussed. We can't even say this is a recent trend – advanced attacks have existed for many years – but we do see them affecting a wider range of organizations, with greater of specificity and finer targeting than ever before. It's no longer merely the defense industry and large financial institutions that need to worry about determined persistent attackers. In the midst of this onslaught, the businesses we protect are using a wider range of technology – including consumer tools – in far more distributed environments. Finally, responders face the dual-edged sword of a plethora of tools; some of them are highly effective, while others aggravate information overload.

> Let's be clear — attackers *have* a beachhead. *Whether you know about it is another matter.*

Before we dig into the gaps we need a bit of context. First, keep in mind that we are focusing on larger organizations with dedicated incident response resources. Practically speaking, this probably means at least a few thousand employees and a dedicated IT security staff. Smaller organizations should still find this series useful, but probably don't have resources to implement many of these recommendations.

Second, these issues and recommendations are based on discussions with real incident response teams. Not everyone has the same issues – especially across large organizations – nor the same strengths. So don't get upset when we start pointing out problems or making recommendations that don't apply to you – as with any research, we generalize to address a broad audience.

Across the organizations we work with, some common incident response gaps emerge:

- **Too much reliance on prevention at the expense of monitoring and response.** We still find even large organizations that rely too heavily on their defensive security tools rather than balancing prevention with monitoring and detection. This resource imbalance leads to gaps in the monitoring and alerting infrastructure, with inadequate resources for response. All organizations are eventually breached, and targeted organizations always have some kind of attacker presence. Always.

- **Too much of the wrong kinds of information too early in the process.** While you do need extensive auditing, logging, and monitoring data, you can't kick off your process on every alert, or analyze all available data in the initial investigation. And expecting to correlate all these disparate data sources as an ongoing practice would be ludicrous. Effective prioritization and filtering is key.
- **Too little of the right kinds of information too early (or late) in the process.** You shouldn't need to jump right from an alert into manually crawling log files. By the same token, after you've handled the initial incident you shouldn't need to rely exclusively on SIEM for forensics investigation and root cause analysis. This again goes back to filtering and prioritization, along with sufficient collection. This requires two levels of collection for key device types – first what you can do continuously, and second the much more detailed information needed to pinpoint root cause or perform post-mortem analysis.
- **Poor alert filtering and prioritization.** We talk constantly about false positives because they are most visible, but the problem is less that an alert triggered and more how to determine its importance in context. This ties directly to the previous two gaps, and requires a balance between alerting, continuing collection of information for initial response, and gathering more granular information for after-action investigation.
- **Poorly structured escalation options.** One of the most important concepts in incident response is building a smooth escalation path to the right resources. Your incident response process and internal organization must take this into account. You just cannot effectively escalate with a flat response structure — tiering, based on multiple factors such as geography and expertise, is key. And this process must be determined well in advance of any incident. Escalation failure during response is a serious problem.
- **Response whack-a-mole.** Responding without the necessary insight and intelligence leads to an ongoing battle, where the organization is always a step behind the attacker. You can't wait to complete a full forensic investigation before clamping down on an incident to contain damage, but you need enough information to make informed and coordinated decisions that actually stop the attack – not merely one symptom. So balancing hair-trigger response against analysis/paralysis is critical to minimize damage and potential data loss.
- ***Your goal in incident response is to detect and contain attacks as quickly as possible – limiting the damage by constraining the window within which the attacker operates.*** To pull this off you need an effective process with graceful escalation to the right resources, to collect the right amounts of the right kinds of information to streamline your process, to perform ongoing analysis to identify problems earlier, and to coordinate your response to kill the threat instead of just a symptom.

Too often we see flat response structures, too much of the wrong information early in the process, too little of the right information late in the process, and a lack of coordination and focus that allow the bad guys to operate with near impunity once they establish their first beachhead. Let's be clear — attackers *have* a beachhead. Whether you know about it is another matter.

In this paper we'll focus on pushing the concepts of incident response past the basics and addressing these gaps. Dealing with advanced threats requires advanced tools. React Faster and Better is about taking a much broader and more effective approach on dealing with attacks – from what data you collect, to how you trigger higher-quality alerts, to the mechanics of response/escalation, and ultimately to remediation and cleaning activities. This is not your grandpappy's incident response.

New Data for New Attacks

As we discussed, we tend to gather too much of the wrong kinds of information, and use it too early in the process. To clarify, we are still fans of *collecting* as much data as you can, because once you miss the opportunity to collect something you'll never get another chance. But there is a tendency to try to boil the ocean with analysis of all sorts of data, which can cause failure and has plagued technologies like SIEM, because customers try to do too much too soon.

The operational objective is to react faster, which means discovering as quickly as possible that you have an issue, and then engaging your incident response process. But merely responding quickly isn't useful if your response is inefficient or ineffective, which is why the other objective is to react better.

Collecting the Right Data at the Right Time

Balancing all the data collection sources available today is like walking a high wire, in a stiff breeze, after knocking a few back at the local bar. We definitely don't lack potential information sources, but many organizations find themselves either overloaded with data or missing key information when it's time to investigate. You need three kinds of data:

- Data to support continuous monitoring and incident alerts/triggers. This is the stuff you look at on a daily basis to figure out when to trigger an incident.
- Data to support your initial response. Once an incident triggers, these are the first data sources you consult to figure out what's going on. This is a subset of all data sources. Keep in mind that not all incidents will tie directly to one of these sources, so sometimes you'll still need to dive into the ocean of lower-priority data.
- Data to support post-incident investigation and root cause analysis. This is a much larger volume, some of it archived, used during the full in-depth investigation.

Fundamentally we believe in data collection and identifying potential incidents from the data. The goal is to collect (almost) everything, and then analyze the *right* material to draw reasonable conclusions that can be investigated using a structured and efficient process.

Collection is fairly simple. You can generate a tremendous amount of data, but with the log management tools available today scale is generally not an issue. Analysis of that data, on the other hand, is still very problematic — when we mention too much of the wrong kinds of information, that's what we are talking about. To address this we advocate segmenting your network into vaults, and analyzing traffic and events within the critical vaults at a deep level.

So collect all you can within the limits of reason and practicality, then analyze the right information sources for early indications of problems, so you can engage the incident response process. Start with a set of sources to support your continuous monitoring and analysis, followed by a set of prioritized data to support initial incident management, and finish with a massive prioritized archive of available data sources.

Continuous Monitoring

We have done a lot of research into SIEM and Log Management, as well as advanced monitoring (Monitoring up the Stack). That's the kind of information to use in your ongoing operational analysis. For those vaults (trust zones) you deem critical, you want to monitor and analyze:

- **Perimeter networks and devices:** The bad guys tend to be *out there*, so they need to cross the perimeter to get to the good stuff. So look for issues on those devices in between them and their targets.
- **Identity:** *Who* is as important as *what*, so analyze access to specific resources – especially within a privileged user context.
- **Servers:** We are big fans of anomaly detection and white listing, on critical servers such as domain controllers and app servers, so you can be alerted to funky stuff at the server level – which usually warrants investigation.
- **Database:** Likewise, correlating database anomalies against other types of traffic (such as reconnaissance and network exfiltration) can indicate a breach in progress. Better to know that early, before your credit card brand notifies you.
- **File Integrity:** Most attacks change key system files, so by monitoring their integrity you can pinpoint when an attacker is trying to make changes. You can even block these attacks using technology like HIPS, but that's another story for another day.
- **Application:** Finally, you should be able to profile normal transactions and user interactions for your key applications (those accessing protected data) and watch for non-standard activities. Again, they don't necessarily indicate a problem, but do help prioritize investigation.

Focus on your most important zones, keeping in mind that you need at least baseline monitoring of everything. The two most common baselines we see are network monitoring and endpoint & server logs (or whatever security tools you use on those systems).

Full Packet Capture Sandwich

One emerging advanced monitoring capability — the most interesting to us — is full packet capture. These devices basically capture all traffic on a given network segment. Why? The only way you can really piece together exactly what happened is to use the actual traffic. In a forensic investigation this is absolutely crucial, providing detail you cannot get from log records.

Going back to a concept we call the Data Breach Triangle, you need three components for a real breach: an attack vector, something to steal, and a way to exfiltrate it. It's impossible to stop all potential attacks, and you can't simply delete all your data, so we advocate heavy perimeter egress filtering and monitoring, to (hopefully) prevent valuable data from escaping your network. Capturing *all* network traffic isn't really practical for any organization of scale, but perimeter traffic should be feasible.

Along with vaulting, we recommend organizations look to deploy full packet capture on the most critical internal segments as well. That's what attackers will go for, so if you capture data from key internal networks as well as perimeter traffic (which is why we call this the sandwich), you have a better chance to piece together what happened.

> You need three components for a real breach: an attack vector, something to steal, and a way to exfiltrate it.

Monitoring these sources for the critical vaults and integrating full packet capture are key parts of your security operational processes. What about less-critical internal zones? You can probably minimize analysis by focusing on things like IDS alerts and NetFlow output, which should be enough to pinpoint egregious issues for investigation. But you probably can't analyze and/or capture all traffic on all your networks all the time – this is likely to be well past the point of diminishing returns.

Sources and Sizing

The good news is that you are likely already collecting most of the data you need, which tends to be log records. Regardless of how deeply you analyze the data, collect as much as is feasible. That means pulling logs from pretty much all the devices you can. Depending on your data collection platform, you can implement a set of less sophisticated log aggregators and only send data from critical segments upstream to a SIEM for analysis. We described this ring architecture on pages 23-26 of Understanding and Selecting SIEM/Log Management.

In addition to logs, mine your identity system, the configurations of your key devices, and network flow data. Many security analysis platforms can gather data from all sorts of sources, making that less of a constraint.

For sizing, it's important to be able to analyze log data over at least a 90-day period. Today's attackers are patient and persistent, meaning they aren't just trying to do smash-and-grabs – they stretch attack timelines to 30, 60, and even 90 days. So you have two vectors for sizing your system: the number of critical segments to analyze, and how long to keep the data. We prefer greater retention across more critical resources, rather than retaining and analyzing everything quickly (before the data is flushed in favor of newer records).

For full packet capture – depending on the size of your perimeter, key network segments, and storage capabilities – it may not be realistic to capture more than 4-7 days of traffic. If so, you'll need to continue traditional log monitoring and analysis on your perimeter and critical segments to catch the low and slow attacks, with full packet capture for the last week to dig deeper once you've identified a pattern.

This is the first leg of the data collection environment: what you need to do on an ongoing basis. Another possible incident response gap is not having enough of the right data later in the process. That means you must be able to dig deep for forensic analysis, to ensure you understand the attack appropriately. Next we'll dig into what that means and how your tactics must evolve given the new types of persistent attacks in play.

Knowing They Are There

Alerts & Triggers

In the last section we considered types of data to systematically collect, through both log record aggregation and full packet capture. As we have said many times, data isn't the issue – it's the lack of actionable information for prioritizing our efforts. We must more effectively automate data analysis to learn what is at risk and what isn't.

Automation Means Tools

We prefer to start with process because that's where most security professionals fail, but automation is really about tools. And plenty of tools are available to alert you when something is funky. You have firewalls, IDS/IPS devices, network monitors, server monitors, performance monitors, DLP, email and web filtering gateways … and that's just the beginning. In fact there is a way to monitor everything in your environment. Twice. And many organizations pump all this data into some kind of SIEM to analyze it, but this continues to underscore that we have too much of the wrong kind of data — at least for incident response.

So let's table the tools discussion long enough to figure out what we are really looking for…

Threat Modeling

Regardless of the tool being used to fire alerts, you need to 1) know what you are trying to protect, 2) know what an attack on it looks like, and 3) be able to prioritize those attacks. Alerts are easy. Relevant alerts are hard. That's why we need to focus considerable effort early in the process on figuring out what is at risk and how it can be attacked.

So we take a page from Security 101 and spend some time building threat models. We delved into this process in gory detail in our Network Security Operations Quant research, so we won't repeat it all here, but the key steps are:

- **Define what's important:** First figure out what critical information/applications would create the biggest issues if compromised.
- **Model how it can be attacked:** It's always fun to think like a hacker, so put on your proverbial black hat and think about ways to exploit and compromise the most important stuff you just identified.
- **Determine the data those attacks would generate:** Those attacks will result in specific data patterns that you can look for using your analysis tools. This isn't always an attack signature – it might be the effect of the attack, such as excessive data egress or bandwidth usage.
- **Set alert thresholds:** Once you establish the patterns, figure out when to actually trigger an alert. This is an art, and most organizations start with fairly broad thresholds, knowing they will result in more alerts initially.
- **Optimize thresholds:** Once your systems start hammering you with alerts, you'll be able to tighten the thresholds to focus on real alerts and increase the signal-to-noise ratio.

- **Repeat for next critical system/data:** Each critical information source/application will have its own set of attacks to deal with. Once you've modeled one, go back and repeat the process. You can't do everything at once, so don't even try. Start with the most critical stuff, get a quick win, and then expand the system.

Keep in mind that the larger your environment, the more intractable modeling everything becomes. You will never know where all the sensitive stuff is. Nor can you build a threat model for every known attack. That's why behind all our research is the idea of determining what's really important and working hard to protect those resources.

Once we have threat models implemented in our monitoring tool(s) – which include element managers, analysis tools like SIEM, and even content monitoring tools like DLP – these products can (and should) be configured to alert based on scenarios in the threat model.

More Distant Early Warning

We wish the threat models could be comprehensive, but understand that inevitably you'll miss something. And there are other places to glean useful intelligence, which can be factored into your analysis and might show attacks not factored into the threat models.

- **Baselines:** Depending on the depth of monitoring, you can and should be looking at establishing baselines for your critical assets. That could mean network activity on protected segments (using NetFlow), or perhaps transaction types (SQL queries on a key database), but you need some way to define 'normal' for your environment. Then you can start by alerting on activities you recognize as abnormal.
- **Vendor feeds:** These come from your vendors – mostly IDS/IPS – because these vendors have research teams tasked with staying on top of emerging attack plans. This is reactive to known attacks, but the vendors spend significant resources making sure their tools remain current. Keep in mind that you'll want to tailor these signatures to your organization and industry – obviously you don't need to look for SCADA attacks if you don't have those control systems.
- **Intelligence sharing:** Larger organizations see a wide variety of stuff, mostly because they are frequently targeted and have the staff to identify attack patterns. Many of these folks cooperate a bit and participate in sharing groups (like FS-ISAC) to leverage each other's experience. This could be a formal deal or just informal conversations over beers every couple weeks. Either way, it's good to know what peer organizations are seeing.

There are many places to leverage data and generate alerts. No one information source can identify all emerging attacks. You're best served by using many, then establishing a system for prioritizing alerts which warrant investigation.

Visualization

Just about every organization – particularly large enterprises – generates more alerts than it has the capacity to investigate. If you don't there is a good chance you aren't alerting enough. So prioritization is a key skill that can determine success or failure. We advocate tiered response, where a first tier of analysts handles the initial alerts. Additional tiers of experts come into play as needed, depending on the sophistication and criticality of the attack.

> Just about every organization – particularly large enterprises – generates more alerts than it has the capacity to investigate... Prioritization is a key skill that can determine success or failure.

How do you figure out what warrants a response in the first place, and where to look for answers? One key tool we see for prioritizing is alert visualization. It could be a topology map to pinpoint areas of the network under attack, or alert categorization by device class (good for recognizing something like a weaponized Windows exploit making the rounds). The idea is to have a mechanism to detect patterns which would not be obvious simply by scanning the event/alert stream.

Critical Incident Data

Once the incident response process kicks in, too much data is rarely the problem, so let's dig deeper into the most *useful* data for the initial stages of incident response. When we don't yet know what we are dealing with, you need to confirm the issue with additional data sources to help isolate the root cause.

We assume that at this stage of investigation a relatively unsophisticated analyst is doing the work. So these investigation patterns can and should be somewhat standard and based on common tools. At this point the analyst is trying to figure out what is being attacked, how the attack is happening, how many devices are involved, and ultimately whether (and what kind of) escalation is required.

Once you understand the general concept behind the attack, you can dig a lot deeper with cool forensics tools. But at this point we are still trying to figure out where to dig. The best way to work through this is to focus on the initial alert, and then what kinds of data would validate the issue and provide the *what*, *how*, and *how many* answers we need at this stage. There are plenty of places we might see the first alert, so let's go through each one.

Network

If a network alert fires, what then? It becomes all about triangulating the data to pinpoint what devices are in play and what the attack is doing. This kind of process isn't comprehensive, but represents what kind of additional data you'd look for and why.

- **Attack path:** The first thing to check is the network map — is there a geographic or segment focus to the network alerts? Figure out what is under attack and how. Is this a targeted attack, where only specific addresses are generating the funky network traffic? Or is it reconnaissance that might indicate some kind of worm proliferating? Or is it command and control traffic which might indicate zombies or persistent attackers?
- **Device events/logs/configurations:** Once we know what IP addresses are in play, we can dig into specific devices to figure out what is happening and/or what changed. At this stage we are looking for obvious stuff. New accounts or executables, and configuration changes, are typical indications of some kind of issue with the device. For the sake of both automation and integrity, this data tends to be centrally stored in one or more system management platforms (SIEM, CMDB, Endpoint Protection Platform, Database Activity Monitor, etc.).
- **Egress path and data:** Finally, we want to figure out what information is leaving your network and (presumably) going into the hands of the bad guys, and how. We aren't concerned with a full analysis of every line item, but need a general sense of what's headed out the door and an understanding of how it's being exfiltrated.

Endpoint

The endpoint may alert first if it's some kind of drive-by download or targeted social engineering attack. You also can see endpoint activity first in the event of a mobile device doing something bad outside your network, then connecting to your internal network to wreak havoc.

- **Endpoint logs/configurations:** Once you receive an alert that something funky is happening on an endpoint, the first thing to do is investigate the device and figure out what's happening. Look for new executables on the device or a configuration change that indicates a compromise.
- **Network traffic:** Another place to look when you get an endpoint alert is the network traffic originating from and terminating on the device. Analyzing that traffic can give you an idea of the targeted. Is it a back-end data store? Other devices? How and from where is the device getting instructions/commands? Also be aware of exfiltration activities, which indicate not only a successful compromise, but also a data breach. The objective is to profile the attack and understand its objective and tactics.
- **Application targets:** Likewise, if it's obvious a back-end datastore is being targeted, you can look at the transaction stream to figure out the attacker's specific target and how widely the attack spread. You also need to know the target to figure out whether and how to remediate.

Upper Layers

If the first indication of an attack happens at the application layer (including databases, application servers, DLP, etc.) – which happens more and more due to the growing popularity and sophistication of application-oriented attacks – then it's about quickly understanding the degree of compromise and watching for data loss.

- **Network traffic:** Application attacks are often all about stealing data, so at the network layer you are looking primarily for signs of exfiltration. Secondarily, understanding the attack path will help discover which devices are compromised, and decide on short and long term remediation options.
- **Application changes:** Is your application functioning normally? Or is the bad guy inserting malware to compromise your customers? While you won't perform a full application assessment at this point, you need to look for key indicators of the bad guy's activities that might not show up through network monitoring.
- **Device events/logs/configurations:** As with the other scenarios, understanding to what degree the devices involved in the application stack are compromised is important for damage assessment.

- **Content monitors:** Given the data theft focus of most application attacks, you'll want to consult your content monitors (DLP, as well as outbound web and email filters) to gauge whether the attack has compromised data and to what degree. This information is critical for determining the amount of escalation required.

Incident Playbook

Obviously there are infinite combinations of data you can look at to figure out what is going on (and whether you need to investigate and/or escalate), but we recommend that the first steps in the process be scripted and somewhat standardized. The higher up the response pyramid you go, the more leeway analysts will need to do what they think is right. The only way to make sure the right information is provided to each succeeding level of escalation is to be very specific and clear what data is required before escalating.

Chain of Custody

Depending on the type and objective of the attack, you may want to consider prosecution, which entails a certain amount of care with data handling and integrity. This becomes particularly important at higher levels of the escalation process, but it's a good habit to make sure that any evidence gathered (even for escalation) is collected in a way that does not preclude prosecution. So part of your incident playbook and analyst training should specify how to gather forensically acceptable data:

- **Isolate the machines:** Depending on what you find on a device, you may want to take a clean image for law enforcement before continuing your investigation.
- **Investigation management:** Many larger organizations use some kind of case management tool to manage the investigation process within a workflow. The first-level response populates this tool and provides structure for the entire investigation. For smaller organizations this may be overkill, but it's worth defining how data will be collected in some detail, and where it will be stored to ensure proper handling.
- **Ensure log file and data integrity:** There are many rules about how to handle log records to ensure their integrity. NIST has a good guide (PDF) for what that typically involves, but ultimately your legal team will need to define the specifics for your organization. Your logging infrastructure must meet the integrity requirements.

Once you have this initial data collected, in a forensically acceptable manner, you need to get into actual investigation and analysis. Additional techniques and tools are required to do this correctly, especially given the nature of modern attacks.

React Faster and Better: In Action

The best way to understand this new approach to incident response is to actually run through a sample incident with some commentary to provide the context you need to apply the method to something tangible. To avoid bogging down in a lot of detail on organizational structure and the like, we'll refer you to our recommended model for an incident response organization and the responsibilities of the Tier 1, Tier 2, and Tier 3 levels of the response organization.

For brevity we will use an extremely simple high-level example of how the three response tiers typically evaluate, escalate, and manage incidents:

The Alert

It's Wednesday morning and the network analyst has already handled a dozen or so network/IDS/SIEM alerts. Most indicate probing from standard network script-kiddie tools and are quickly blocked and closed (often automatically). He handles those himself — just another day in the office.

The network monitoring tool pings an alert for an outbound request on a high port to an IP range located in a country known for intellectual property theft. The analyst needs to validate the origin of the packet, so he looks and sees the source IP is in Engineering.

The tier 1 analyst passes this information along to a tier 2 responder. Important intellectual property may be involved and he suspects malicious activity, so he also phones the on-call handler to confirm the potential seriousness of the incident. Tier 2 takes over, and the tier 1 analyst goes back to his normal duties.

This is the first indication that something may be funky. Probing is nothing new and tier 1 needs to handle that kind of activity itself. Yet the outbound request very well may indicate an exfiltration attempt. And tracing it back to a device that does have access to sensitive data means it's definitely something to investigate more closely. This kind of situation is why egress monitoring and filtering are so important. Monitoring is generally the only way you can tell if data is actually leaking. At this point the tier 1 analyst should know he is in deep water. He has confirmed the issue and pinpointed the device in question. Now it's time to hand it off to tier 2. Note that the tier 1 analyst follows up with a phone call to ensure the hand-off happens and that there is no confusion.

How Bad Is Bad?

- The tier 2 analyst opens an investigation and begins a full analysis of network communications from the system in question. The system is no longer actively leaking data, but she blocks any traffic to that destination on the perimeter firewall by submitting a high priority request to the firewall management team. After that change is made, she verifies that traffic is in fact being blocked.
- She sets an alert for any other network traffic from that system and calls or visits the user, who predictably denies knowing anything about it. Through that investigation she also learns that system normally doesn't have access to sensitive intellectual property, which normally would be good. But this also might indicate privilege escalation or that the machine stages data before exfiltration – both of which are bad signs. Endpoint protection platform (EPP) logs for that system don't indicate any known malware. Without conclusive data, the smoking gun is still lacking, so it's time to go even deeper.
- She notifies her tier 3 manager of the incident and begins a deeper investigation of previous network traffic from network forensics data. She also starts looking into system logs to begin isolating the root cause.
- Once the responder notices outbound requests to a similar destination from other systems on the same subnet, she informs incident response leadership that they may be experiencing a serious compromise.
- Then she finds that the system in question connected to a sensitive file server it normally doesn't access, and transferred/copied entire directories. It's going to be a long night.

As we have mentioned, tier 2 tends to focus on network forensics because it's usually the quickest way to pinpoint attack proliferation and severity. The first step is to contain the issue, which entails blocking traffic to the external IP – this should temporarily eliminate any data leakage. Remember, you might not actually know the extent of the compromise, but that shouldn't stop you from taking decisive action to contain the damage as quickly as possible. At this point, tier 3 is notified – not necessarily to take action, but so they are aware there might be a more serious issue. This kind of proactive communication streamlines escalation.

Next, the tier 2 analyst needs to determine how much the issue has spread within the environment. So she searches through the logs and finds a similar source, which is not good. More than one compromised device might represent a major breach. Worst yet, she sees that at least one of the involved systems connected to a sensitive file store and grabbed a large chunk of content. So it's time to escalate and fully engage tier 3. Not that it hasn't been great so far, but now the fun really begins.

Bring in the Big Guns

- Tier 3 steps in and begins in-depth analysis of the involved endpoints and associated network activity. They detect custom malware that initially infected a user's system via drive-by download after clicking a phishing link. No wonder the user didn't know anything – they didn't have a chance against this kind of attack.
- An endpoint forensics analyst then discovers what appear to be the remains of an encrypted `.rar` file on one of the affected systems. The network analysis shows no evidence the file was transferred out. It seems they dodged a bullet and detected the command and control traffic before the data exfiltration took place.
- The decision is made to allow what appears to be encrypted command and control traffic over a non-standard port, while blocking all outbound file transfers (except those known to be part of normal business processes). Yes, they run the risk of blocking something legit, but senior management is now involved and has decided this is a worthwhile risk, given the breach in progress.
- To limit potential data loss through the C&C channels left open, they carefully monitor bandwidth usage. Due to the advanced nature of the attack they are trying to contain the problem without tipping off the attackers that they know

what's going on. Prior experience warns that prematurely cutting off communications will only escalate the attack before they can identify and clean involved systems.

- Sensitive data is slowly removed and replaced from the servers on that subnet. Forensics investigators turn an infected system over to the malware analysts to reverse engineer. The goal is to prepare a coordinated cleaning method and expulsion of the attacker, but to do this they need to fully understand the depth of compromise and identify all involved systems and malware variants.
- IDS/IPS specialists write a new network alert signature to identify similar traffic, and create a new malware signature for evaluating endpoints in the future.

Tier 3 immediately starts to analyze the attack and how it works. Once you have identified custom malware you know you aren't dealing with amateurs. So the decision to allow C&C traffic but block file transfers is unsurprising, though a bit risky. Until the malware analysts understand how to eliminate the threat it doesn't make sense to give any hints that you know about the attack.

At the same time outbound transfers are stopped, the response team acts decisively to remove sensitive data from the reach of these attackers. This again serves to contain the damage until the threat can be neutralized, which involves a set of custom network rules to block this particular attack.

To be clear, sophisticated attacks in the real world are rarely this cut and dried, but response team tactics are consistent. The objective is always to contain the damage while figuring out the extent of the compromise. Then you have options for how to remediate it.

Post-mortem

To reiterate: the key points in the scenario above are rapid identification of a serious issue (outbound IP exfiltration), quick escalation to tier 2 for scoping and initial investigation, and rapid coordinated investigation and response with high-level resources once it becomes clear this is a sophisticated and advanced attack. The initial handler did a good job of recognizing the problem and understanding he couldn't handle it himself. The second level responder didn't fall into the trap of focusing too much on the first device and missing the bigger picture. The containment plan provided breathing space for a full cleansing of the incident without tipping off the attackers to rush a deeper penetration, or allowing additional loss of important assets.

We need to React Faster and Better because we face new and sophisticated attacks. That makes detecting every attack before it happens a pipe dream. By focusing on shortening the window between attack and detection, and having a solid plan to contain and then remediate the attack, you give yourself the best chance to live to fight another day. That's one of the most significant epiphanies security folks can have. You cannot win, so success is about minimizing the damage. Yeah, that's crappy, but it is realistic.

To position yourself most effectively to RFaB, we advocate an institutional commitment to data collection at all levels of the computing stack. Given the usefulness of network-level data throughout the incident response process; monitoring tools such as full packet capture, Database Activity Monitoring, and Data Leak Prevention provide the best chance of being able to detect, contain, isolate, and remediate today's sophisticated attacks.

> To position yourself most effectively to React Faster and Better, we advocate an institutional commitment to data collection at all levels of the computing stack.

But no collection of tools will ever replace a skilled team of incident handlers and investigators. **Get the right people, establish the right processes, and then give them the tools and support to do what they do best.**

About the Analysts

Rich Mogull, Analyst/CEO

Rich has twenty years experience in information security, physical security, and risk management. He specializes in data security, application security, emerging security technologies, and security management. Prior to founding Securosis, Rich was a Research Vice President at Gartner on the security team, where he also served as research co-chair for the Gartner Security Summit. Prior to his seven years at Gartner, Rich worked as an independent consultant, web application developer, software development manager at the University of Colorado, and systems and network administrator. Rich is the Security Editor of *TidBITS*, a monthly columnist for *Dark Reading*, and a frequent contributor to publications ranging from *Information Security Magazine* to *Macworld*. He is a frequent industry speaker at events including the RSA Security Conference and DefCon, and has spoken on every continent except Antarctica (where he's happy to speak for free — assuming travel is covered).

Prior to his technology career, Rich also worked as a security director for major events such as football games and concerts. He was a bouncer at the age of 19, weighing about 135 lbs (wet). Rich has worked or volunteered as a paramedic, firefighter, and ski patroller at a major resort (on a snowboard); and spent over a decade with Rocky Mountain Rescue. He currently serves on a federal disaster medicine and terrorism response team, where he mostly drives a truck and lifts heavy objects. He has a black belt, but does not play golf. Rich can be reached at rmogull (at) securosis (dot) com.

Mike Rothman, Analyst/President

Mike's bold perspectives and irreverent style are invaluable as companies determine effective strategies to grapple with the dynamic security threatscape. Mike specializes in the sexy aspects of security, such as protecting networks and endpoints, security management, and compliance. Mike is one of the most sought-after speakers and commentators in the security business, and brings a deep background in information security. After 20 years in and around security, he's one of the guys who "knows where the bodies are buried" in the space.

Starting his career as a programmer and a networking consultant, Mike joined META Group in 1993 and spearheaded META's initial foray into information security research. Mike left META in 1998 to found SHYM Technology, a pioneer in the PKI software market, and then held executive roles at CipherTrust and TruSecure. After getting fed up with vendor life, Mike started Security Incite in 2006 to provide a voice of reason in an over-hyped yet underwhelming security industry. After taking a short detour as Senior VP, Strategy at eIQnetworks to chase shiny objects in security and compliance management, Mike joined Securosis with a rejuvenated cynicism about the state of security and what it takes to survive as a security professional.

Mike published The Pragmatic CSO <http://www.pragmaticcso.com/> in 2007 to introduce technically oriented security professionals to the nuances of what is required to be a senior security professional. He also possesses a very expensive engineering degree in Operations Research and Industrial Engineering from Cornell University. His folks are overjoyed that he uses literally zero percent of his education on a daily basis. He can be reached at mrothman (at) securosis (dot) com.

About Securosis

Securosis, L.L.C. is an independent research and analysis firm dedicated to thought leadership, objectivity, and transparency. Our analysts have all held executive level positions and are dedicated to providing high-value, pragmatic advisory services.

Our services include:

- *Primary research publishing*: We currently release the vast majority of our research for free through our blog, and archive it in our Research Library. Most of these research documents can be sponsored for distribution on an annual basis. All published materials and presentations meet our strict objectivity requirements, and follow our Totally Transparent Research policy.
- *Research products and strategic advisory services for end users*: Securosis will be introducing a line of research products and inquiry-based subscription services designed to assist end user organizations in accelerating project and program success. Additional advisory projects are also available, including product selection assistance, technology and architecture strategy, education, security management evaluations, and risk assessments.
- *Retainer services for vendors*: Although we will accept briefings from anyone, some vendors opt for a tighter, ongoing relationship. We offer a number of flexible retainer packages. Example services available as part of a retainer package include market and product analysis and strategy, technology guidance, product evaluations, and merger and acquisition assessments. Even with paid clients, we maintain our strict objectivity and confidentiality requirements. More information on our retainer services (PDF) is available.
- *External speaking and editorial*: Securosis analysts frequently speak at industry events, give online presentations, and write and/or speak for a variety of publications and media.
- *Other expert services*: Securosis analysts are available for other services as well, including Strategic Advisory Days, Strategy Consulting engagements, and Investor Services. These services tend to be customized to meet a client's specific requirements.

Our clients range from stealth startups to some of the best known technology vendors and end users. Clients include large financial institutions, institutional investors, mid-sized enterprises, and major security vendors.

Additionally, Securosis partners with security testing labs to provide unique product evaluations that combine in-depth technical analysis with high-level product, architecture, and market analysis. For more information about Securosis, visit our website: http://securosis.com/.