

# VAST 2013 Mini-Challenge 3 Solution Description for Reviewers and Committee

Updated 7/11/2013

## Introduction

The 2013 VAST Mini-Challenge 3 focuses on situation awareness of operations for the fictitious international marketing company, Big Marketing. Contestants are provided with background information that provides the setting and context for the problem and solution; three questions that they must answer in their submission concerning the scenario, a dataset consisting of three different types of data, and a standard answer form. The Challenge description provided to contestants can be found at the VA Community Wiki Site, under the listing [VAST Challenge 2013: Mini-Challenge 3](#). Before reviewing solution packets, reviewers may want to familiarize themselves with the general background provided for Big Marketing.

The primary goal of the contestants is stated in the Challenge Description. To wit:

“You work as the Big Marketing computer network manager, ensuring that Big Marketing networks are up and running for both the Internet-facing web services and the internal workforce. This responsibility encompasses the full range of maintaining current operations, planning for future needs, and securing and defending network assets against threats. So, although this mini-challenge has a forensics theme, it is still concerned with situation awareness.

...your job is to understand events taking place on your networks over a two week period. To support your mission, your choice of visual analytics should support near real-time situation awareness. In other words, as network manager, your goal for your department is to notice network events as quickly as possible.”

The three questions to be answered for Mini-Challenge 3 are:

**MC3.1** – Provide a timeline (i.e., events organized in chronological order) of the notable events that occur in Big Marketing’s computer networks for the supplied data. Use all data at your disposal to identify up to twelve events and describe them to the extent possible. Your answer should be no more than 1000 words long and may contain up to twelve images.

**MC3.2** – Speculate on one or more narratives that describe the events on the network. Provide a list of analytic hypotheses and/or unanswered questions about the notable events. In other words, if you were to hand off your timeline to an analyst who will conduct further investigation, what confirmations and/or answers would you like to see in their report back to you? Your answer should be no more than 300 words long and may contain up to three additional images.

**MC3.3** – Describe the role that your visual analytics played in enabling discovery of the notable

events in MC3.1. Describe whether your visual analytics play a role in formulating the questions in MC3.2. Your answer should be no more than 300 words long and may contain up to three additional images.

Contestants have the option to ask up to five questions of the VAST Challenge team to get more context for the signals they see in the dataset. Contestants who have asked their questions wisely will be able to give more complete and nuanced answers to the challenge questions.

## Ground Truth

(Please note - the scenario and all organizations are entirely fictitious.)

A group calling themselves the “Butterfly Warriors” is attacking Big Marketing over the two week period. Big Marketing is helping Total Crop Protection Services roll out a marketing campaign for “Butterfly 2.0”, an altered butterfly that will eventually lead to the extinction of natural butterflies. Prior to the dates covered by the dataset, the Butterfly Warriors send a threatening letter to Big Marketing. (This letter is only provided to the participants if they ask about it.)

In Week 1, the Butterfly Warriors begin by doing reconnaissance on the network. Then they hit the network with denial of service attacks (also known as DOS or DDOS attacks, depending on how they are conducted) on successive days. These attacks cause the [www.bigmkt2.com](http://www.bigmkt2.com) server to crash, but [www.bigmkt3.com](http://www.bigmkt3.com) uses load balancing across two servers, so it is better able to withstand the attack and does not crash.

Simultaneously with but independent of the last denial of service attack, the Butterfly Warriors implant malicious code on one of the Big Marketing web externally-facing sites; this infection event does not leave any traces in the data. For the remainder of Week 1, visitors to this affected web site (both Big Marketing staff and external customers) are immediately redirected to a malicious web server, where they are also infected with malicious code. Clues to this infection are very subtle: session durations for visitors on the infected Big Marketing web site are now very short, and infected Big Marketing computers are seen visiting a new web site not previously visited.

One of the infected computers belongs to the system administrator for Big Marketing. The Butterfly Warriors use this vulnerability to open up all of the protected ports on the network. On the weekend, the Butterfly Warriors hack in through the system administrator’s computer. They exfiltrate a couple of high value files from the Big Marketing network. It is not clear from the data what these files are, but if contestants ask, they will find out that the exfiltrated files are

- A file containing Big Marketing’s private client information
- A recording of a video conference between Big Marketing and Total Crop Protection Services discussing the marketing plan and the likely consequences of Butterfly 2.0.

Once the system administrator discovers that important files are being exfiltrated, he pulls Big Marketing off the internet in order to investigate and add an Intrusion Protection System (IPS). This

results in a three day gap in the data collection. In week 2 when Big Marketing reconnects to the internet, the Butterfly Warriors try many of the same attacks that they used successfully in the first week, including denial of service attacks and FTP exfiltration, their attacks are stopped by the IPS. However, the Butterfly Warriors have a more effective trick -- they post Big Marketing's exfiltrated customer information on the internet. Big Marketing's representatives are inundated with angry calls from their customers who have seen their private data on the internet, so the marketing reps all go out to the web site to see what was posted. There, they pick up yet more malware -- this time, they are turned into slaves to a botnet controller. By the end of Week 2, the Butterfly Warriors are using the Big Marketing computers to stage denial of service attacks on another company's networks.

## Network and Data Characteristics

The Big Marketing Network consists of three separate sites, each with its own domain controller, email server, web servers, and user workstations. The network is outfitted with a network flow collector which captures all of the traffic between Big Marketing and the (fictitious) internet used in this challenge, as well as a small portion of the internal Big Marketing traffic. In Week 2 of the data, the network is augmented with an Intrusion Protection System as well. The network diagrams and detailed descriptions of the network for both weeks are available as downloads on the reviewer site.

The Big Marketing web sites use addresses in the 172.x.x.x space internally. The "internet" in this scenario uses IPs in the 10.x.x.x address space. When Big Marketing traffic goes out to the "internet at large" -- it passes through a Network Address Translation (NAT) layer and receives a 10.X.X.X based address. Contestants can interpret the Big Brother and netflow data correctly using just the internal (172.x.x.x) Big Marketing addresses, but must also use the NAT addresses when using the IPS data.

- Site 1 - "bigmkt1" uses addresses in the range 172.10.0.1-255. Servers with NAT addresses are 10.0.2.2-8.
- Site 2 - "bigmkt2" uses addresses in the range 172.20.0.1-255. Servers with NAT addresses are 10.0.3.2-8 and 15.
- Site 3 - "bigmkt3" uses addresses in the range 172.30.0.1-255. Servers with NAT addresses are 10.0.4.2-8.

The data for Week 1 begins at 7:30 a.m. on April 1. The normal traffic includes web browsing by Big Marketing staff, by customers browsing Big Marketing web sites, email traffic and FTP of files into Big Marketing.

Three types of data are provided for the two week period of Mini-Challenge 3:

- **Network flow (netflow) data.** Netflow is a network traffic collecting tool on the Big Marketing network. As traffic passes through the tool, highly condensed metadata about the traffic is captured and logged. The logged data includes information such as the IP addresses that connected and the duration of the connection. The resulting log data was provided to the participants in comma-separated values (CSV) format.

- **Big Brother data.** Big Brother is a network health monitoring tool. Servers and workstations run a client program to monitor resources and report several types of status to the Big Brother server every few minutes. This data was logged, and some of the relevant fields were extracted in order to make it easier for participants to analyze.
- **Intrusion Prevention System (IPS).** An IPS will provide an additional layer of protection against external intrusions. An IPS blocks detects and blocks certain types of traffic; it creates log entries to document what it found. An IPS was used in Week 2 only. It was configured to permit regular web and email traffic, but to stop threats including denial of service attacks, and ftp from inside Big Marketing to outside Big Marketing.

Participants can detect events by identifying specific trends in the data, and in some cases, by identifying the absence of records. For example, when a machine becomes too busy or falls off the network, it ceases to report Big Brother records.

## Answers to Specific Mini-Challenge Questions

MC3.1: Participants should describe up to twelve events using either a list or a visualization view. There are many more than twelve different events present in the data. It is hoped that participants will aggregate fine-grained details into summarized events in order to keep their responses to twelve events. Contestants are expected to report events they see in the data, of course. However, if they also ask good questions and identify more of the explanation for why the attacks are occurring, they should also include this information in their answers. Contestants are required to provide their interpretations in MC3.2 but might also provide some of this analysis in MC3.1

Reviewers are not expected to evaluate the submission for accuracy, which would be time-consuming and require detailed knowledge of the data. Accuracy reviews are being performed by a separate team. However, reviewers are asked to consider whether the submitted solution supports the detection of both obvious and subtle events in the data. Because of the way this data is generated, there are almost certainly additional patterns detectable in the data that were not embedded intentionally.

The following table provides a detailed list of items that may be included in the timeline. The column “Aggregate Event” reflects a logical aggregation of the events, although the submitter may choose to aggregate them differently. Note that the table shows fourteen aggregate events, while contestants have been asked to report no more than twelve. It is expected that some of the very subtle events will not be identified by the contestants. The column “Degree of Subtlety” identifies whether an item is fairly obvious from the data, whether the data signals are subtle, or whether the participant would know about this item only from asking the appropriate questions of the VAST Challenge team.

Date and Time	Aggregate Event	Degree of Subtlety	Event
03/01/2013	Videoconference	Questions only	Videoconference between Big Marketing and Total Crop Protection Services
03/15/2013	Threatening Letter #1	Questions only	Letter sent to Big Marketing from Butterfly Warriors
04/01/2013 11:30	Port scans	Subtle	Port scans of computers by an attacker (10.6.6.6) occur against Site 3.
04/02/2013 05:22	Denial of Service #1	Obvious	Part 1: Denial of service attack attempted. <a href="http://www.bigmkt3.com">www.bigmkt3.com</a> is subjected to a denial of service attack from 10 attackers.
04/02/2013 07:00	Denial of Service #1	Subtle	Part 2: server crash due to denial of service attack. <a href="http://www.bigmkt3.com">www.bigmkt3.com</a> becomes unresponsive from the attack temporarily.
04/02/2013 13:25	Port scans	Subtle	Port scans occur against Site 3 from a single attacker (10.6.6.6)
04/03/2013 09:30 - 11:48	Denial of Service #2	Obvious	Part 1: Denial of service attack attempted. 10 attackers use denial of service on <a href="http://www.bigmkt3.com">www.bigmkt3.com</a> . 5 attackers use a denial of service to attack <a href="http://www.bigmkt2.com">www.bigmkt2.com</a> .
04/03/2013 11:27	Denial of Service #2	Subtle	Part 2: server crash due to denial of service attack <a href="http://www.bigmkt3.com">www.bigmkt3.com</a> crashes again from the denial of service attack.
04/04/2013	Malware on Big Marketing web site	Subtle	Sessions that connect to <a href="http://www.bigmkt2.com">www.bigmkt2.com</a> will exhibit shorter session durations and reduced packet counts in the netflow traffic because malware has been inserted onto a Big Marketing web site. This malware redirects visitors to a new and malicious web site, where the visiting computers also become infected with malware.
04/05/2013	Denial of Service #2	Subtle	<a href="http://www.bigmkt3.com">www.bigmkt3.com</a> returns to the network.
04/05/2013	Admin infection	Subtle	An administrator machine hits <a href="http://www.bigmkt2.com">www.bigmkt2.com</a> and gets infected.
04/06/2013 10:36	Exfiltration	Obvious	The administrator's computer being compromised allows an FTP exfiltration for a large file (~100 MB) to 10.7.5.5.
04/07/2013 07:00	Exfiltration	Obvious	FTP exfiltration of ~650MB file to 10.7.5.5.
04/07/2013	Threatening letter #2	Questions only	Big Marketing received follow up threat from Butterfly Warriors
4/7/2013 9:10 to 4/10/2013 6:50	Network Down	Obvious	Big Marketing network is down while the administrator investigates the security issues and installs an IPS.
04/10/2013 12:20	Port scans	Obvious	Multiple port scans against all Big Marketing sites

<b>04/11/2013 10:16</b>	Failed FTP exfiltration	Obvious	Attempted FTP blocked by IPS
<b>04/11/2013 10:35 - 11:21</b>	Port scans	Obvious	Port scan against Sites 1, 2, and 3
<b>04/11/2013 11:55 - 12:18</b>	Failed Denial of Service	Obvious	Failed denial of service attack against Big Marketing sites 1-3.
<b>04/11/2013 12:24 - 12:57</b>	Port scans	Obvious	All sites targeted for port scans
<b>04/12/2013 ~8:10</b>	Botnet infection	Subtle	Part 1: Origin of infection. Account managers web navigate to external site hosting exfiltrated data, become infected with botnet malware.
<b>04/12/2013 08:24</b>	Botnet infection	Obvious	Part 2: Ongoing infection. Big Marketing machines affected by a botnet start participating in a DDOS against external customer machines. Communicate every 10 minutes via SSH to command and control server(10.0.3.77)
<b>04/12/2013 11:23</b>	Port scans	Obvious	Port scans to sites 2 and 3
<b>04/13/2013 5:41</b>	Port scans	Obvious	Port scans on sites 1 and 3
<b>04/13/2013 6:51 – 7:51</b>	Botnet Attacks	Obvious	8 internal Big Marketing machines start a DDOS against an external machine
<b>04/14/2013 07:18-8:18</b>	Botnet Attacks	Obvious	8 internal Big Marketing machines start a DDOS against an external machine (10.1.0.100)
<b>04/14/2013 12:22</b>	Port scans	Obvious	Port scans against all 3 Big Marketing sites
<b>04/15/2013 7:45</b>	Port scans	Obvious	Port scans against all 3 Big Marketing sites

It is highly likely that contestants will identify events that are not on the timeline. Although credit should be given to plausible descriptions, the preferred answers should relate to items in the above timeline.

MC3.2: Contestants should propose reasonable analytic questions or hypotheses that reflect a pattern of successive attacks. The data are intended to simulate a set of targeted and persistent attacks. The Butterfly Warriors attack across as many fronts as possible. Their attack plan does not rely upon a fixed series of steps, but they opportunistically take advantage of any foothold they gain and exploit it. When Big Marketing goes offline and adds an IPS, previous attacks now fail but the attacking group comes up with new ways to wreak havoc.

Given that teams have the opportunity to ask up to five questions, they may know the name of the attacking group and the motivation for the attacks, as described in the Ground Truth section above.

Contestants will have to hypothesize possible causes for new behavior, such as:

- The attacks may be coordinated/targeted at Big Marketing (judging from the attacking IP addresses)
- The corporate machines have been recruited by the attackers to be part of a botnet, which they then use to attack other sites.
- Big Marketing needs to determine whether the threat is internal (“insider threat”). Some of the attack behavior seems to indicate an internal origin.
- Large outbound transfers suggest data exfiltration. Company resources should be diverted to investigating the severity of the information loss. Decisions are needed as to whether to contact law enforcement.
- The sudden increased attention to a previously unknown website (on 4/12) should be investigated as potentially connected to the recent problems. (If the submission is very good, they might be able to hypothesize a link between this website to the prior data exfiltration of customer data.

MC3.3: The submission should show that visual analytics techniques aided in identifying events, linking multiple events to detect that patterns, and inferring cause and effect.