Examining Intrusion Prevention System Events from Worldwide Networks

Sathya Chandran*
Kansas State University
Manhattan, KS, USA
sathya@ksu.edu

Sandeep Bhatt
HP Labs
Princeton, NJ, USA
sandeep.bhatt@hp.com

Marc R. Eisenbarth
HP TippingPoint
Austin, TX, USA
marc.r.eisenbarth@hp.com

Abstract

HP TippingPoint IPS devices are installed in over 1,000 customer networks, deployed in every part of the network, from the perimeter to the network core, and see a very wide variety of attacks, from common everyday attacks such as Cross Site Scripting (XSS) to the more sophisticated attacks against Microsoft RPC bugs being launched within a network by a malicious insider. The IPS devices inspect traffic in real-time and enforce security policies at network speeds approaching 10 Gbps with over 6,000 filters deployed. The policies are maintained as part of the TippingPoint Reputation Digital Vaccine (DV) service. Figure 1 shows a typical deployment of a TippingPoint IPS in a customer's network.

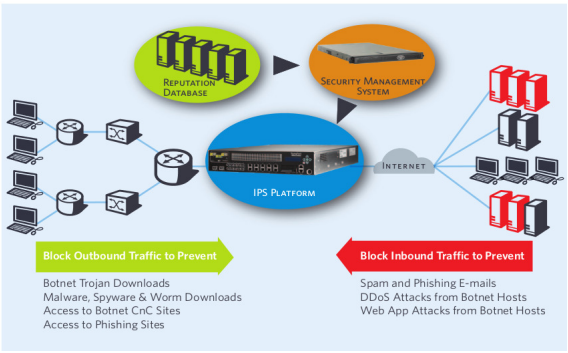

Figure 1: TippingPoint IPS deployment

IPS devices use signatures to flag inbound and outbound traffic that is known to be bad or that violates policy. As new vulnerabilities become known, new signatures are pushed out to IPS devices. When an incoming or outbound packet triggers a filter, the filter can either block, rate-limit or allow the traffic to pass through. TippingPoint devices are configured to block all traffic that triggers critical and high severity alerts. In all cases, an alert is recorded by the device; among other fields, each alert contains the source and destination IP addresses, port numbers, filter ID, a hit-count which represents the number of times the filter was triggered within a one-minute interval, and timestamp.

In contrast to Intrusion Detection Systems (IDS), IPS filters which block traffic must be guaranteed to have very low false positive rates. This generally comes at the expense of potentially higher false negative rates. Low false positive rates are achieved by designing filters to block post-compromise events such as a connection back to a command and control server or traffic intended to propagate the infection to other machines. Furthermore, the filter set focuses on blocking vulnerabilities rather than specific incarnations, often called exploits. This difference is important as well in that a generic signature for, say, the MS08-067 vulnerability will block any exploit variant that produces the network traffic on the wire needed to exploit this vulnerability, as we see for example with the numerous Conficker variants. In other words, the IPS serves to block the behavior common to all exploits that leverage a given vulnerability, not necessarily the exploits themselves.

Clearly, no single mitigation mechanism can address the myriad of complex attack scenarios that challenge computer security today. Often times, NIPS must pass on complex malware analysis, thereby allowing an endpoint to be infected and instead focus on detecting and preventing any resulting outgoing communication for this host to command and control servers. With the complexities associated with compound documents such as PDF and Microsoft Office documents, and the plethora of vulnerabilities against these code bases, the most reliable defense is to identify machines post-compromise for remediation and excise traffic associated with compromise emanating from these machines.

TippingPoint customers can opt-in to allow the aggregated alert data to be sent to a centralized ThreatLinQ server which provides aggregate statistics and common lists of bad IP addresses and domain names. This ThreatLinQ dataset we have started to analyze has been collected over a 5 year period, between 2007 and early 2012. Besides its large size, processing the data required non-trivial effort because of the complexity of interpreting, analyzing and parsing the accompanying filter metadata and correlating it with the raw alert data.

Our goal in examining this dataset, beyond reporting aggregate statistics, is to understand the nature of attackers, the attacks launched, and the customers targeted. For example, are certain customers, or groups of customers, more likely to be early targets of many attacks? Are there relationships between attacks and their targets? Do attacks occur in clusters? This extended abstract reports our initial findings. As we continue to mine the dataset we expect to develop furher insights into the nature of attacks.

*Sathya Chandran is a Ph.D., student in the Computing and Information Sciences department at Kansas State University, Manhattan, KS, USA. This work was done as part of his summer (2012) internship at HP Labs, Princeton, NJ, USA.