A National Early Warning Capability Based on a Network of Distributed Honeypots

Cristine Hoepers, Klaus Steding-Jessen, Luiz E. R. Cordeiro, Marcelo H. P. C. Chaves
NIC BR Security Office – NBSO
Brazilian Computer Emergency Response Team
{cristine,jessen,cordeiro,mhp}@nic.br

Abstract

We present here the work developed by NBSO/Brazilian CERT, in the "Brazilian Honeypots Alliance – Distributed Honeypots Project", to centralize the data gathered in several honeypots and to process this data to be used for early warning and incident response. We shortly describe how the honeypots are deployed and how the data is centralized, then focus on how the data is being used by NBSO to generate statistics and to notify networks potentially compromised or infected.

1 The Distributed Honeypots Project

Honeypots have been pointed as tools that can play a significant role in gaining early warning of attacks [1], although most of the effort made by the community has focused on using honeypots and honeynets to observe the motives and methods of the attackers [2,3,4].

With the objective of using low-interaction honeypots [2] as tools for early warning and trend analysis, the NBSO/Brazilian CERT and the CenPRA Research Center deployed the "Brazilian Honeypots Alliance – Distributed Honeypots Project"[1]. The main objective of this project is to increase the capacity of early warning, event correlation and trend analysis in the Brazilian Internet space.

The Brazilian Honeypots Alliance is coordinated by NBSO and CenPRA, and has around 30 partner institutions from academia, government and private sector, as we can see in Figure 1 and Table 1. Each institution is responsible for the maintenance of at least one honeypot and for providing a network range[2] to be used by the honeypot.

All honeypots run Honeyd[3] on the OpenBSD operating system, and are configured according to the project's standards. The goals of these standards are to guarantee that all honeypots are hardened to minimize the chances of a compromise, and to facilitate the maintenance. These honeypots generate Honeyd and `pf` [5] firewall logs. The firewall is configured to log the payload of the packets. These logs are then collected to a central location, as described in section 2.

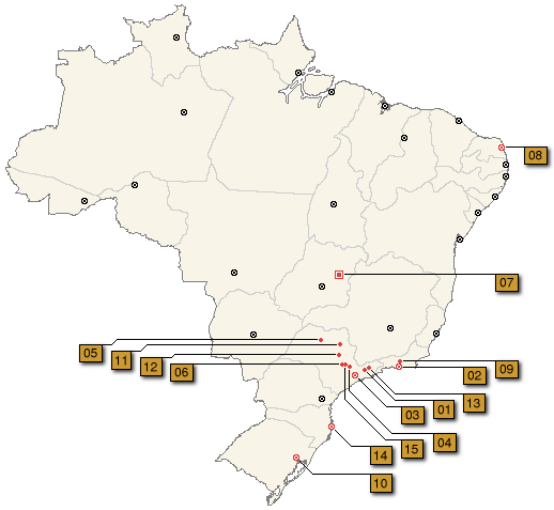

Figure 1: Brazilian map showing the cities where there are honeypots deployed, as of April 2005. See details in Table 1.

[1] Project site: `http://www.honeypots-alliance.org.br/`
[2] Usually a `/24` per institution, but this can vary.
[3] `http://www.honeyd.org/`

#	City	Alliance Members
01	S. José dos Campos	INPE, ITA
02	Rio de Janeiro	CBPF, Fiocruz, PUC-RIO, RedeRio, UFRJ
03	São Paulo	ANSP, Diveo, Durand, NBSO/Brazilian CERT, UNESP, USP
04	Campinas	CenPRA, HP Brasil, UNICAMP
05	S. José do Rio Preto	UNESP
06	Piracicaba	USP
07	Brasília	Brasil Telecom, TCU, UNB LabRedes
08	Natal	UFRN
09	Petrópolis	LNCC
10	Porto Alegre	CERT-RS
11	Ribeirão Preto	USP
12	São Carlos	USP
13	Taubaté	UNITAU
14	Florianópolis	UFSC DAS
15	Americana	VIVAX AMR

Table 1: Members of the Brazilian Honeypots Alliance grouped by city, as of April 2005.

2 Data Collection and Status Check

NBSO has developed several tools that collect the data gathered by each honeypot. The honeypots maintain their logs locally for 24 hours. Every day the central server connects to each honeypot and transfers the data using `ssh`.

In order to ensure that all the honeypots are operating correctly, a remote status checking mechanism was developed. By running this several times a day from a centralized location, it is possible to quickly determine that all honeypots are running according to the project's standards.

The tests are performed using a ssh connection to each honeypot. Some of the items checked are listed below:

- connectivity;
- processes running;
- amount of free disk space;
- clock synchronization, using NTP;
- uptime.

3 Data Analysis

Every day, project members receive a sanitized summary of the activities observed in all honeypots. This summary is distributed only to the members, through an encrypted mailing list, and includes information about ports, protocols, IPs and Source OSs that generated the major part of the activities observed.

Besides these internal summaries, we also maintain public daily statistics based on network flow data directed to the honeypots of the Brazilian Honeypots Alliance.

We maintain statistics for the categories below:

- Destination TCP Ports (bytes/s)
- Destination TCP Ports (packets/s)
- Destination UDP Ports (bytes/s)
- Destination UDP Ports (packets/s)
- Source Country Codes (bytes/s)
- Source Country Codes (packets/s)
- Source Operating Systems (bytes/s)
- Source Operating Systems (packets/s)

To explain how the graphics are generated, we will use the "Destination TCP Ports" category as an example. We extract all the destination TCP ports seen in the network flow data and compute the quantity of bytes or packets received by each port. The top 10 TCP ports related to the highest quantities of bytes or packets are selected to be displayed on the graphic. The quantities of bytes or packets of the remaining ports are summed in a data set called "Others". Then, this information is used to plot the packet or byte rate per second for the top 10 TCP ports and the "Others" data set. The other categories follow the same criteria.

Each image presented in the statistics contains a stack area graphic, which means that each data set of the graphic is stacked on top of the previous one.

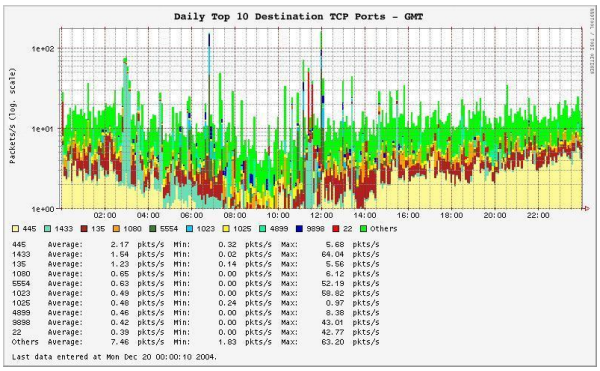

Figure 2: Top TCP ports (packets).

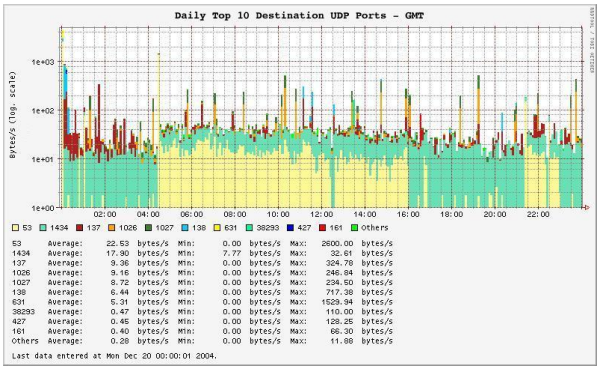

Figure 3: Top UDP ports (bytes).

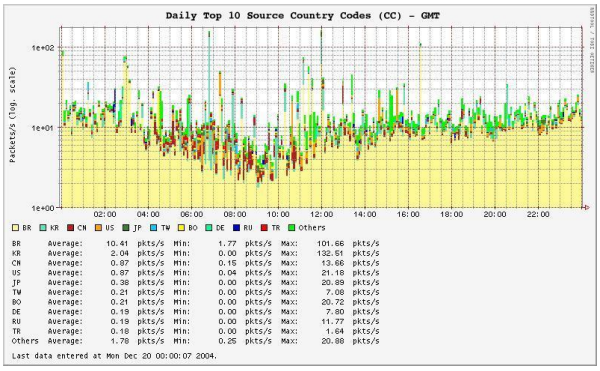

Figure 4: Top Country Codes (packets).

Some sample graphics of data collected in December 19th, 2004 are presented in Figures 2, 3, 4, and 5. Figure 2 shows the top 10 TCP ports that received the highest packet quantities. Figure 3 shows the top 10 UDP ports that received the highest byte quantities. Figure 4 shows the top 10 countries related to the majority of source network flows directed

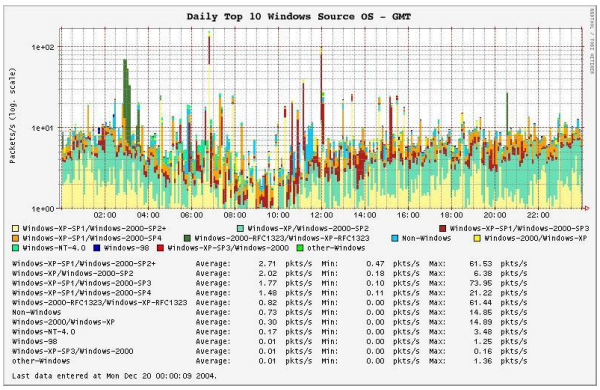

Figure 5: Top Windows Source Operating Systems (packets).

to the honeypots. The country code[4] values are obtained from RIR[5] statistics files. And, finally, Figure 5 shows the top 10 Windows Operating Systems related to the majority of source network flows directed to the honeypots. Source operating system guessing is performed by passive fingerprinting. Non-Windows Operating Systems statistics are also generated and available in the project's web site.

4 Use for Incident Response

All data collected is analyzed in order to identify signatures of well known malicious activities, for example: bots, worms and scans for ports known to run vulnerable services. Once these kind of activity is identified, we separate this data in 2 categories: Brazilian IP addresses and foreign IP addresses.

A portion of the sanitized data related to malicious activities coming from foreign IP addresses is donated to The Team Cymru[6], so the networks that use the data provided by their projects can benefit from the traffic seen in our honeypots.

NBSO processes all data related to malicious activities coming from Brazilian IP addresses and analyse the packets and their payloads looking for signatures of attacks.

We then find the contacts and/or CSIRTs for all the IPs and send to each one an email with the sani-

[4] ISO 3166 2-letter code.

[5] AfriNIC, APNIC, ARIN, LACNIC, and Ripe NCC.

[6] http://www.cymru.com/

tized logs and guidelines about how to deal with that information. The guidelines are usually technical tips to recover from a compromise or a worm infection.

Some categories of malicious activities identified are:

- generic scans
  - 21/TCP, 22/TCP, 111/TCP, etc
- worms/viruses
  - blaster, codered, dabber, mydoom, nimda, slammer, etc
- bots
- spam activity
  - open proxy scans
  - pop-up spam

Acknowledgments

This project is sponsored by both the NBSO/Brazilian CERT and the CenPRA Research Center.

Several other people and institutions have helped to set up the project. We would like to give our special thanks to all the Honeypots Alliance members[7], who maintain their own honeypots and allow us to collect data in their networks. We would also like to thank the Honeynet.BR Team[8] for their support and ideas.

References

[1] H. Lipson, "Tracking and Tracing Cyber-Attacks: Technical Challenges and Global Policy Issues," Tech. Rep. CMU/SEI-2002-SR-009, CERT Coordination Center, Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, November 2002.

[2] The Honeynet Project, *Know Your Enemy: Revealing the Security Tools, Tactics, and Motives of the Blackhat Community*. Addison-Wesley, 1st ed., August 2001. ISBN 0-201-74613-1.

[3] L. Spitzner, "Learning the Tools and the Tactics of the Enemy with Honeynets," in *Proceedings of the 12th Annual Computer Security Incident Handling Conference*, (Chicago, Illinois, USA), June 2000.

[4] C. Hoepers, K. Steding-Jessen, and A. Montes, "Honeynets Applied to the CSIRT Scenario," in *Proceedings of the 15th Annual Computer Security Incident Handling Conference*, (Ottawa, Canada), June 2003.

[5] D. Hartmeier, "Design and Performance of the OpenBSD Stateful Packet Filter (pf)," in *Proceedings of the FREENIX Track: 2002 USENIX Annual Technical Conference (FREENIX '02)*, (Monterey, California, USA), June 2002.

[7] http://www.honeypots-alliance.org.br/

[8] http://www.honeynet.org.br/