



IP Dossier

Paul N. Krystosek, Ph.D.
CERT/NetSA
January 2008



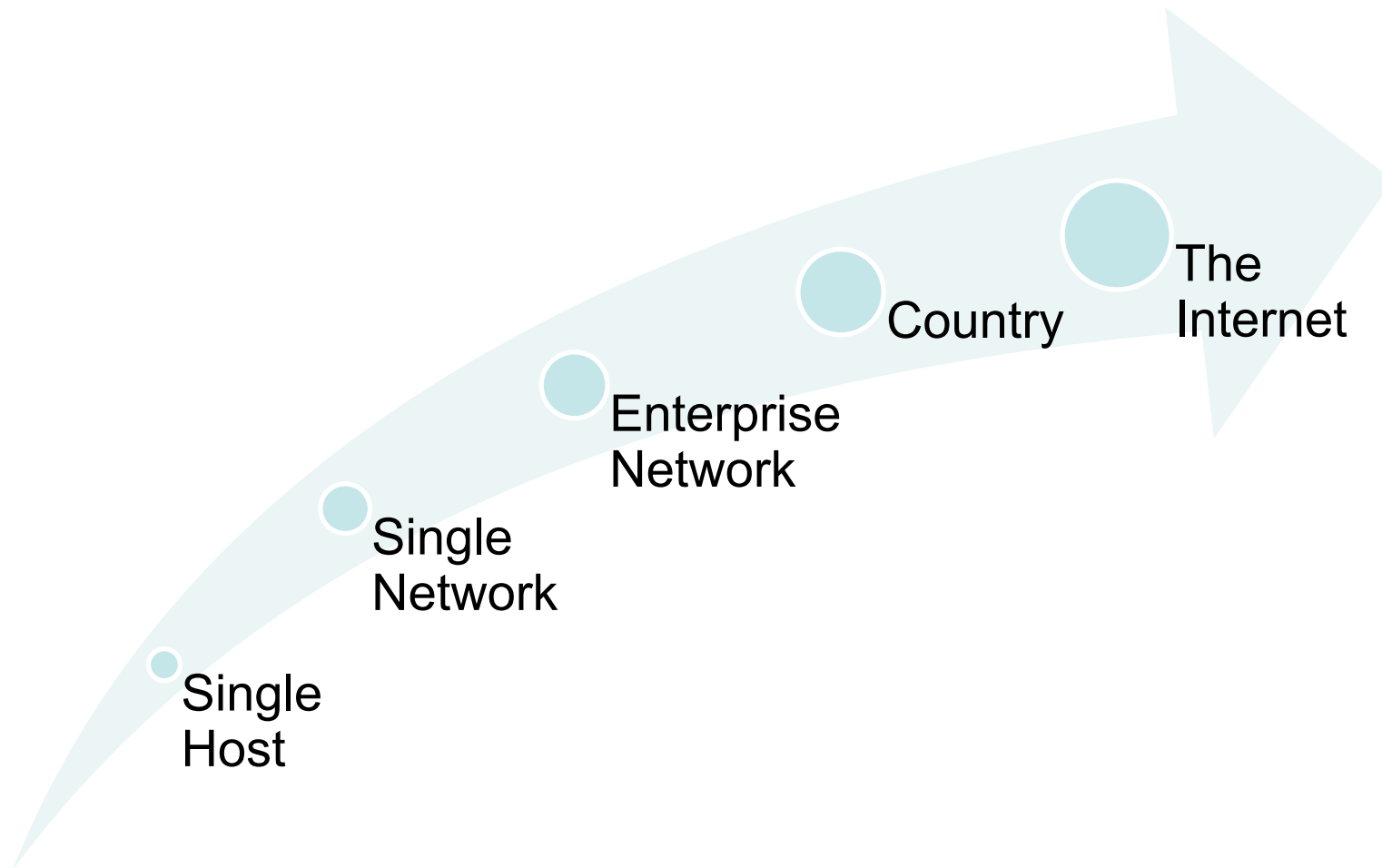
SEI CERT NetSA

Mission

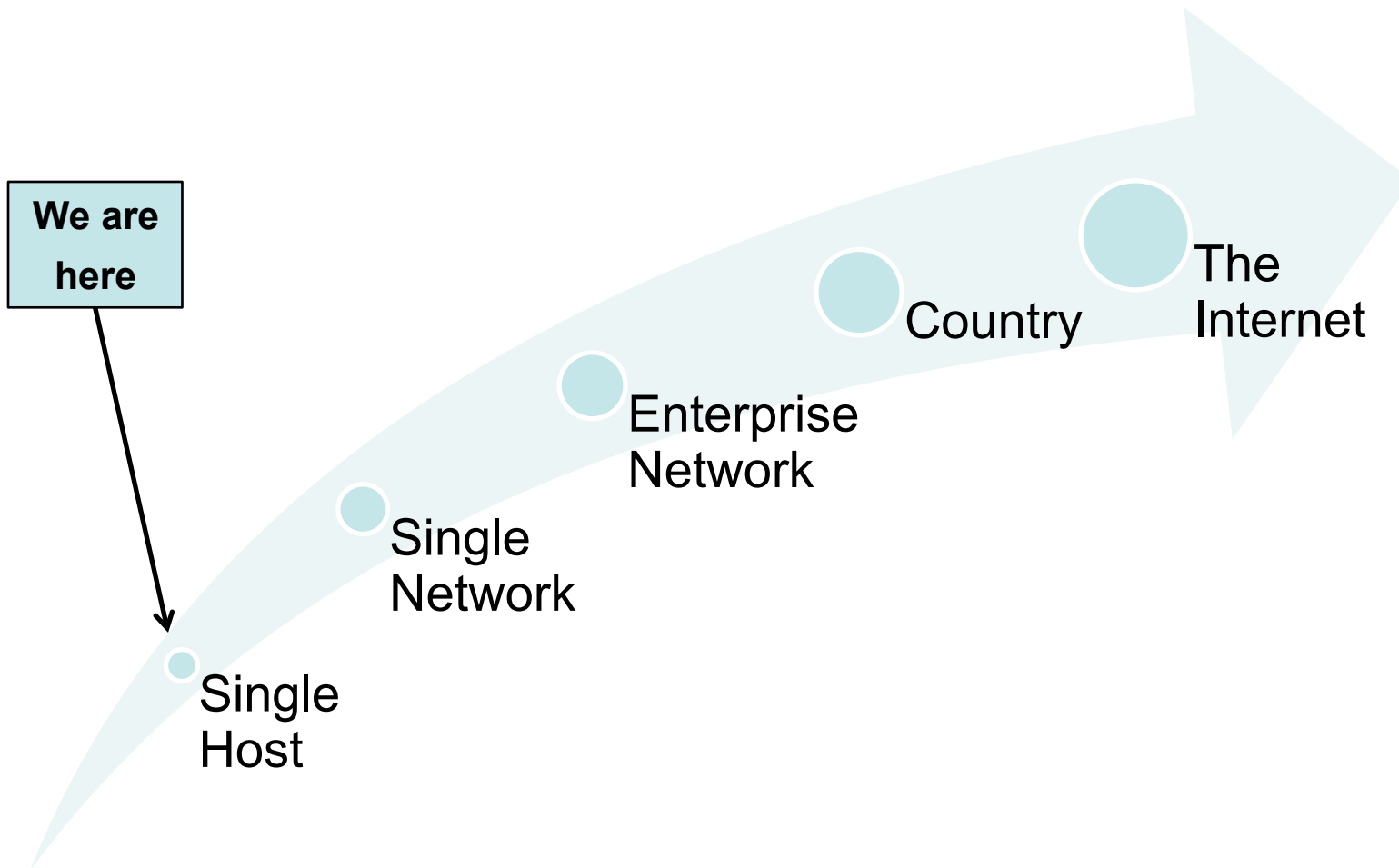
The mission of the Network Situational Awareness (NetSA) team is to enable and provide situational awareness to a broad constituency through applied research and engineering approaches. [Situational Awareness](#) (SA) attempts to quantitatively characterize threats and intruder activity in order to provide network operators tractable views and actionable insight into their network; improve and confirm best-practices; and inform technology design and implementation.

As an auxiliary goal, the NetSA team explicitly tries to foster a community of analysts spanning organizations.

The Range of Network Situational Awareness



The Range of Network Situational Awareness



Outline

Introduction

Motivation

The Task

Method

Examples

Redesign

Conclusion

Introduction

This is a work in progress

It does not exist as a cohesive product

It is not done

It needs input from your expertise

Motivation

Why are we doing this?

- Automate a complex task
- Transition from “one off” to “everyday” operation
- Establish “organizational memory”
- Refine the subtasks that make up the whole

The Task

Find everything about the activity of a host given an IP address (and perhaps a time range)

Primarily from NetFlow data

Present it in an understandable fashion

To at least two levels of personnel

- Manager
- Analyst

Why is *The Task* Important?

A common task in computer security incident response is to fulfill a request such as:

- Tell us everything there is know about host with IP address a.b.c.d
- Host w.x.y.z compromised a system at agency blah look for it at your agency and report back... now
- This host at your agency just scanned our agency, what are you going to do about it?

How to Organize it?

Administrative information

External information

Top level or Volumetric information

Connection oriented

Service and protocol oriented

Applications

Functional

Behavioral

OS or IP stack specific

Administrative & External Information

Administrative Information

- Nslookup
- Whois

External information (not flow based)

- Watchlists
- UV

Top Level and Connection Oriented

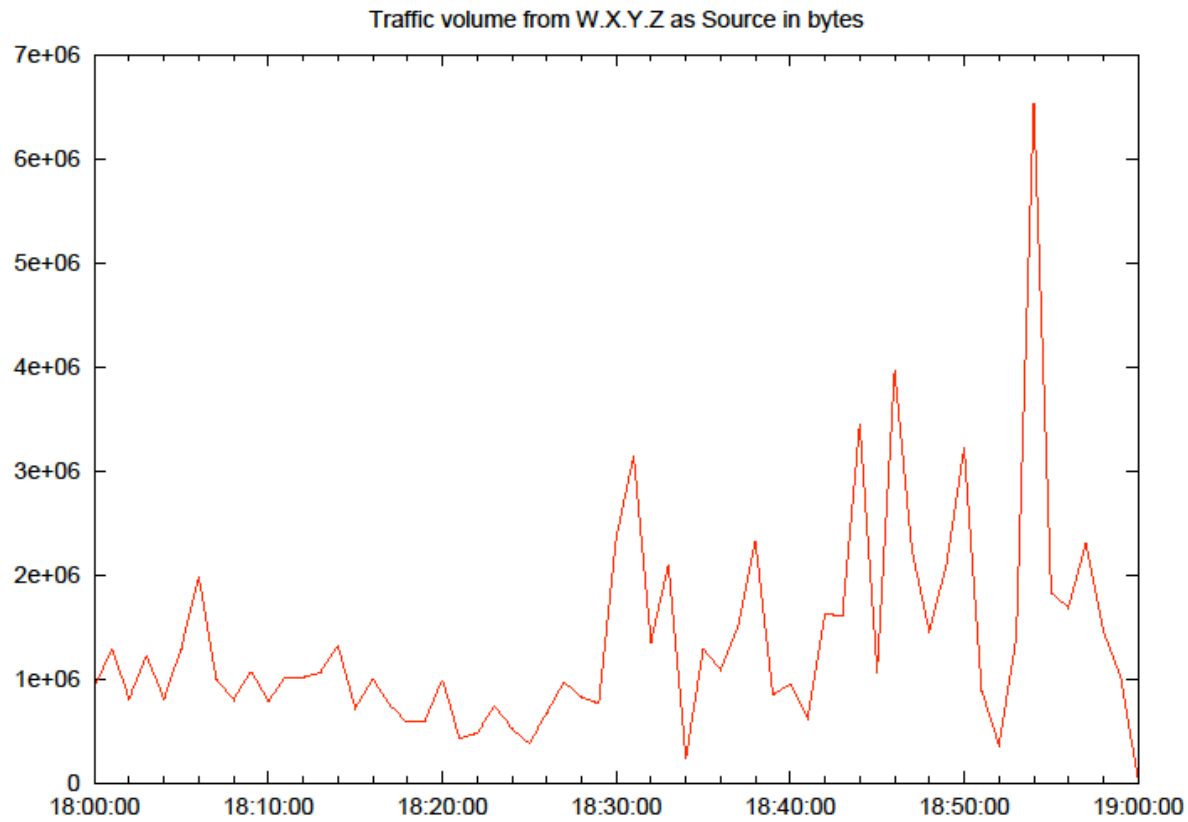
Top Level

- Look at volume and direction of traffic
- Diurnal cycle
- Continuous operation

Connection Oriented

- Who did it talk to?

Example Volume Plot



Service and Protocol Oriented

By looking at ports and protocols

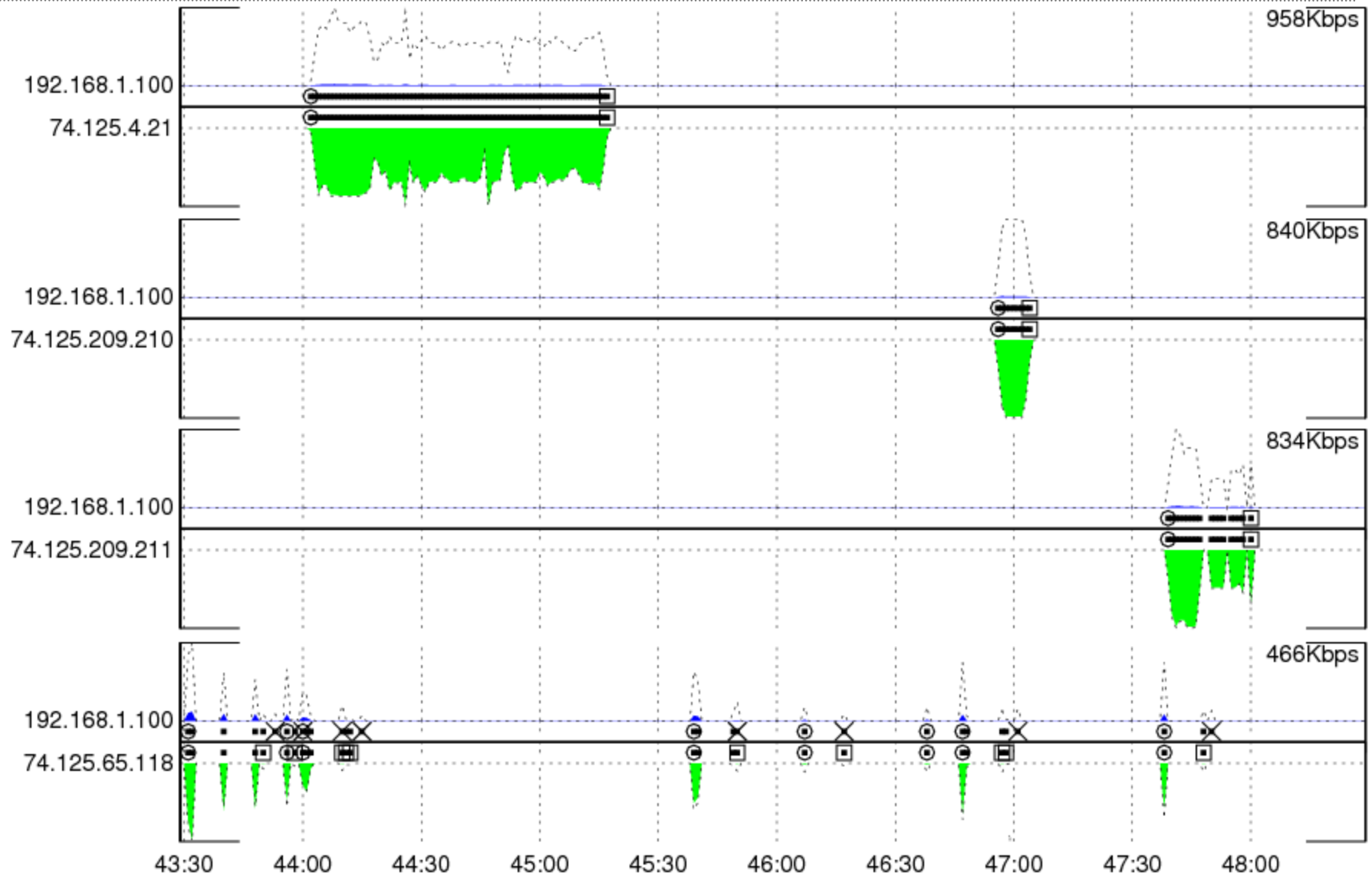
- Is it a client, server or both
- Do we see its DNS queries
 - If so, is it a client or server?
- Same for mail, web ...

Applications

Sid Faber showed us how to find

- Web
- Video
- Audio
- News feed
- FTP

Watch Three YouTube Videos



Functional

Is it a

- Client
- Server
- Proxy
- Network device

FloViz looks like it will help out here

Functional continued

John Gerth from Stanford has a useful and simple technique to categorize hosts he calls **Local Role**

- Servers are given a positive number
- Clients are given a negative number
- 1 and -1 indicate packets in one direction only
- 2 and -2 indicate packets in both directions
- 3 and -3 indicate data in both directions
- 0 indicates backscatter

Behavioral

What is it doing?

- Scanning
- Beaconsing
- Ordinary user/server

OS or IP Stack specific

- OS by ephemeral port behavior

Format

The example I'll show is in Word

But a LaTeX template producing a PDF would be a better choice

I was able to ~~fake~~ prototype it in Word by hand fairly easily, but it won't scale well

Small stuff as tables

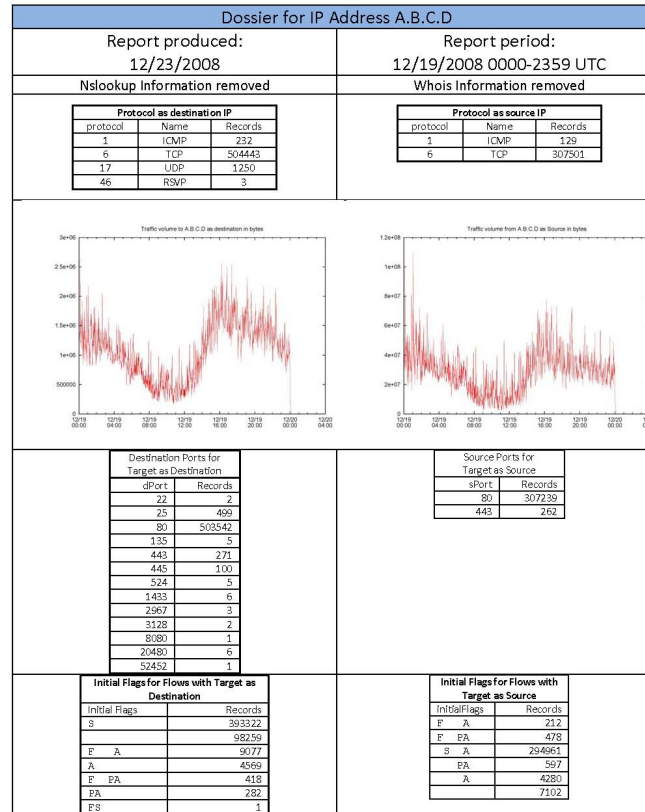
Large stuff as plots

Connection diagram useful in some instances

What Might the Result Look Like?

Dossier for IP Address: 1.2.3.4	
Report Date	Date/time range analyzed
nslookup information	whois information
Watchlist	UV Status
Inbound byte volume plot	Outbound byte volume plot
Protocol Information	Connection Partners
Ephemeral Port plot	Service Port Information
“As client”	“As server”
Out bound scan info	In bound scan info
Out bound flag usage	In bound flag usage
Analyst commentary	

Dossier of a Server



Server Dossier continued

Session Flags for flows with target as destination		Session Flags for flows with target as source	
Session Flags	Records	sessionFlags	Records
F PA	164289	FS PA	1826
RPA	164578	F A	1647
S	88626	SRPA	4
F RPA	21363	F PA	251181
PA	13913	S A	1433
	12888	PA	36398
R A	11390	SR A	5
F A	4098	A	1474
FS PA	1650	F RPA	9
S	975	R	7284
SRPA	727	R A	22
F R A	517	FS A	51
A	318	S PA	575
FSRPA	235	RPA	176
S PA	232		5545
SR	48		
FS A	45		
SR A	18		
Other	18		

Analyst Commentary
This would seem to be a Web server with clients in North America that access it during business hours. It is not scanned excessively. And does not provide other services to the Internet. It uses internal DNS services.

Client Dossier

Dossier for IP Address W.X.Y.Z																									
Report produced: 12/23/2008	Report period: 12/19/2008 0000-2359 UTC																								
Nslookup Information removed	Whois Information removed																								
Target does not appear on any watchlists																									
<table border="1"> <thead> <tr> <th colspan="3">Protocol as destination IP</th> </tr> <tr> <th>protocol</th> <th>Name</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>1</td> <td>ICMP</td> <td>10</td> </tr> <tr> <td>6</td> <td>TCP</td> <td>389</td> </tr> <tr> <td>17</td> <td>UDP</td> <td>11</td> </tr> </tbody> </table>	Protocol as destination IP			protocol	Name	Records	1	ICMP	10	6	TCP	389	17	UDP	11	<table border="1"> <thead> <tr> <th colspan="3">Protocol as source IP</th> </tr> <tr> <th>protocol</th> <th>Name</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>6</td> <td>TCP</td> <td>416</td> </tr> </tbody> </table>	Protocol as source IP			protocol	Name	Records	6	TCP	416
Protocol as destination IP																									
protocol	Name	Records																							
1	ICMP	10																							
6	TCP	389																							
17	UDP	11																							
Protocol as source IP																									
protocol	Name	Records																							
6	TCP	416																							
	<table border="1"> <thead> <tr> <th colspan="2">Destination Ports for Target as Source</th> </tr> <tr> <th>sPort</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>80</td> <td>408</td> </tr> <tr> <td>443</td> <td>8</td> </tr> </tbody> </table>	Destination Ports for Target as Source		sPort	Records	80	408	443	8																
Destination Ports for Target as Source																									
sPort	Records																								
80	408																								
443	8																								
No apparent outbound scans	21 apparent inbound SYN scans from 13 IP addresses																								

Client Dossier, continued

<table border="1"> <thead> <tr> <th colspan="2">Initial Flags for Flows with Target as Source</th> </tr> <tr> <th>initialFlags</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>S A</td> <td>322</td> </tr> <tr> <td></td> <td>60</td> </tr> <tr> <td>S</td> <td>21</td> </tr> <tr> <td>F A</td> <td>4</td> </tr> <tr> <td>A</td> <td>3</td> </tr> </tbody> </table>	Initial Flags for Flows with Target as Source		initialFlags	Records	S A	322		60	S	21	F A	4	A	3	<table border="1"> <thead> <tr> <th colspan="2">Initial Flags for Flows with Target as Source</th> </tr> <tr> <th>initialFlags</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>S</td> <td>321</td> </tr> <tr> <td></td> <td>74</td> </tr> <tr> <td>A</td> <td>17</td> </tr> <tr> <td>F A</td> <td>3</td> </tr> <tr> <td>PA</td> <td>1</td> </tr> </tbody> </table>	Initial Flags for Flows with Target as Source		initialFlags	Records	S	321		74	A	17	F A	3	PA	1																						
Initial Flags for Flows with Target as Source																																																			
initialFlags	Records																																																		
S A	322																																																		
	60																																																		
S	21																																																		
F A	4																																																		
A	3																																																		
Initial Flags for Flows with Target as Source																																																			
initialFlags	Records																																																		
S	321																																																		
	74																																																		
A	17																																																		
F A	3																																																		
PA	1																																																		
<table border="1"> <thead> <tr> <th colspan="2">Session Flags for Flows with Target as Destination</th> </tr> <tr> <th>Session Flags</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>F PA</td> <td>250</td> </tr> <tr> <td>R</td> <td>83</td> </tr> <tr> <td>RPA</td> <td>21</td> </tr> <tr> <td>PA</td> <td>20</td> </tr> <tr> <td>F RPA</td> <td>20</td> </tr> <tr> <td>F A</td> <td>17</td> </tr> <tr> <td>A</td> <td>2</td> </tr> <tr> <td></td> <td>1</td> </tr> <tr> <td>SRPA</td> <td>1</td> </tr> <tr> <td>F R A</td> <td>1</td> </tr> </tbody> </table>	Session Flags for Flows with Target as Destination		Session Flags	Records	F PA	250	R	83	RPA	21	PA	20	F RPA	20	F A	17	A	2		1	SRPA	1	F R A	1	<table border="1"> <thead> <tr> <th colspan="2">Session Flags for Flows with Target as Destination</th> </tr> <tr> <th>sessionFlags</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>F PA</td> <td>262</td> </tr> <tr> <td></td> <td>51</td> </tr> <tr> <td>R</td> <td>30</td> </tr> <tr> <td>PA</td> <td>23</td> </tr> <tr> <td>R A</td> <td>20</td> </tr> <tr> <td>A</td> <td>10</td> </tr> <tr> <td>RPA</td> <td>6</td> </tr> <tr> <td>S</td> <td>3</td> </tr> <tr> <td>F A</td> <td>3</td> </tr> <tr> <td>F RPA</td> <td>1</td> </tr> <tr> <td>S PA</td> <td>1</td> </tr> </tbody> </table>	Session Flags for Flows with Target as Destination		sessionFlags	Records	F PA	262		51	R	30	PA	23	R A	20	A	10	RPA	6	S	3	F A	3	F RPA	1	S PA	1
Session Flags for Flows with Target as Destination																																																			
Session Flags	Records																																																		
F PA	250																																																		
R	83																																																		
RPA	21																																																		
PA	20																																																		
F RPA	20																																																		
F A	17																																																		
A	2																																																		
	1																																																		
SRPA	1																																																		
F R A	1																																																		
Session Flags for Flows with Target as Destination																																																			
sessionFlags	Records																																																		
F PA	262																																																		
	51																																																		
R	30																																																		
PA	23																																																		
R A	20																																																		
A	10																																																		
RPA	6																																																		
S	3																																																		
F A	3																																																		
F RPA	1																																																		
S PA	1																																																		
Analyst Commentary																																																			
<p>This would seem to be a client in North America active during business hours in the Central or Eastern Time Zone. It is not scanned excessively. The regular activity that occurs every 80 minutes or so is with Google.</p>																																																			

Malware Dossier

Dossier for Target IP Address A.B.C.D																			
Report produced: 1/09/2008	Report period: 1/05/2008 0000-2359 UTC																		
Nslookup Information removed	Whois Information removed																		
Target appears on malware watchlist																			
<table border="1"> <thead> <tr> <th colspan="3">Protocol as destination IP</th> </tr> <tr> <th>protocol</th> <th>Name</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>6</td> <td>TCP</td> <td>363</td> </tr> </tbody> </table>	Protocol as destination IP			protocol	Name	Records	6	TCP	363	<table border="1"> <thead> <tr> <th colspan="3">Protocol as source IP</th> </tr> <tr> <th>protocol</th> <th>Name</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>6</td> <td>TCP</td> <td>360</td> </tr> </tbody> </table>	Protocol as source IP			protocol	Name	Records	6	TCP	360
Protocol as destination IP																			
protocol	Name	Records																	
6	TCP	363																	
Protocol as source IP																			
protocol	Name	Records																	
6	TCP	360																	
	<table border="1"> <thead> <tr> <th colspan="2">Destination Ports for Target as Destination</th> </tr> <tr> <th>sPort</th> <th>Records</th> </tr> </thead> <tbody> <tr> <td>80</td> <td>363</td> </tr> </tbody> </table>	Destination Ports for Target as Destination		sPort	Records	80	363												
Destination Ports for Target as Destination																			
sPort	Records																		
80	363																		
No apparent outbound scans	No apparent inbound scans																		

Malware Dossier, continued

	Session Flags for Flows with Target as Destination			Session Flags for Flows with Target as Destination	
	Session Flags	Records		sessionFlags	Records
	F'S PA	363		F'S PA	360
Analyst Commentary					
<p>This address is known to host malware. It is the target of beacons after an internal system has been infected. The beacon interval is approximately 5 minutes (295-310 seconds). Beacons typically continue for several hours. It is not known what data is transferred during the flows with larger numbers of bytes.</p>					

Things to Consider

We have a “dossier” on an IP address, now

- How long is it good for?
- When it “expires” and we run another, do we keep the old one?
- Who can we show it to?
- Have we drawn the correct conclusions?
- Is there a better way to store it than PDF?

Future Work

Better report format

TCP and UDP Work Weight

Uncleanliness Vector values

Entropy

Better volume plots

More comprehensive scan data

Think about how other disciplines coax surprising amounts of information out of raw data

Conclusion(s)

Flow data can provide a lot of information about a single host

It's useful to present all it in one place

How clever can we be to entice more information out of it?

Now let's automate the process

Still need analyst commentary

Remember the audience

Help Me Redesign It, Please

What should it look like?

What should it include?

How might you use it?



Thanks for sticking around

**Paul Krystosek
CERT Software Engineering
Institute
pnk@cert.org
<http://www.sei.cmu.edu/>
<http://www.cert.org/netsa/>**

