

# Using 'Big Data' Tools to Understand Your Cyber Environment

BE Rhodes | November 19th, 2020



# Outline

- WHOIS: Brad Rhodes
- Why Are We Here?
- Cyber is Big Data!
- Big Data Tools (Free-ish is Good)
- Normalizing Data (Garbage In = Garbage Out)
- Visualizing an Environment (Wireshark, Power BI, Tableau)
- Resources for You!
- Questions & Contact Info

# WHOIS: Brad Rhodes

zvelo

- WHOIS: Brad Rhodes
- TLDR:
  - ✓ Head of Cybersecurity at zvelo
  - ✓ LTC, Cyber (17A) Colorado Army National Guard & Cyber Shield Planner
  - ✓ Military Cyber Professionals Association, HammerCon Co-Lead
  - ✓ Speaker, Author, Professor, Coach
  - ✓ #toomany Pro-Certs, highlights: CISSP-ISSEP, CISM, PMP, CEH, GMON, GCIH, RHCSA, CCNA Cyber Ops, CySA+



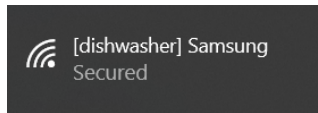
Recent Podcast Appearance (Colorado = Security, Episode 159):

<https://www.colorado-security.com/podcast>

# Why Are We Here?

You can put just about **EVERYTHING** on the Internet today.

- **Traditionals** - Laptops, servers, phones, network hardware, etc.
- **Internet of Things** - Smart speakers, thermostats, fridges, crockpots, and more.
- **Everything Else** - ICS/SCADA, sensors, cars, and others



The Numi toilet combines unmatched design and technology to bring you the finest in personal comfort and cleansing. Kohler's most advanced toilet now offers personalized settings that let you fine-tune every option to your exact preferences, from ambient colored lighting to wireless Bluetooth® music sync capability to the heated seat and foot warmer. Play your favorite music and podcasts - simply stream wirelessly with any device enabled with Bluetooth technology, store MP3 files to the SD card, or plug in your device using the auxiliary cable. Other upgrades include Power-Save mode for energy efficiency, emergency flush for power outages, and an intuitive touch-screen remote. From its striking form to its exceptional water efficiency, the Numi toilet marks a new standard of excellence in the bathroom. [Read More](#)



➡ Do you really know what is in your Cyber Operating Environment?

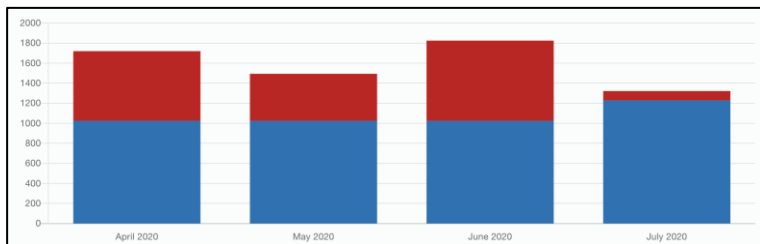
➡ Are you **VULNERABLE** (due to the 'smart' thing your employee BYOD'd)?

➡ Is your **DATA LEAKING** right now (via something you thought you could trust)?

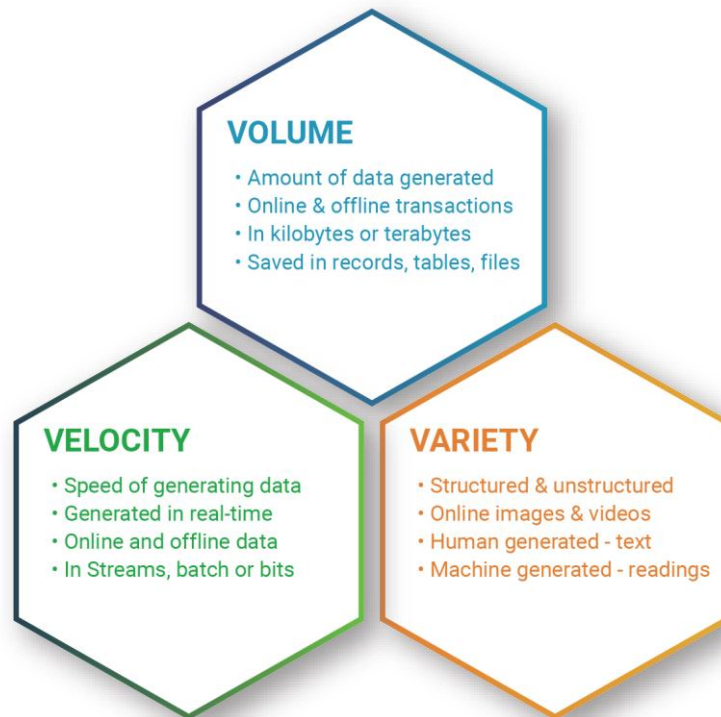
If any of these questions keep you up at night, then this talk is for you! Let's explore **'FREE'** (other than your time) Big Data tools, tactical techniques, and concepts you can use right now to begin to understand the cyber stuff in your environment! And hopefully sleep a little better!

# Cyber is Big Data!

- **VOLUME:** Zettabytes (~1 Billion Terrabytes) of data on the internet (source: Cisco)
- **VOLUME and VARIETY:** 5G — support for 1,000,000 devices per km<sup>2</sup> (source: Rogers Communications)
- **VARIETY:** 500 Billion '*things*' on the internet by 2030 (source: Cisco)
- **VOLUME and VARIETY:** Small Home Network has 40-50 devices and millions of data points daily
- **VELOCITY:** 1Gbps typical network speeds



## THE 3VS OF BIG DATA



# Big Data Tools (Free-ish is Good)

## Microsoft Power BI (Desktop): Capabilities & Limitations



The image displays three key components of the Microsoft Power BI Desktop interface:

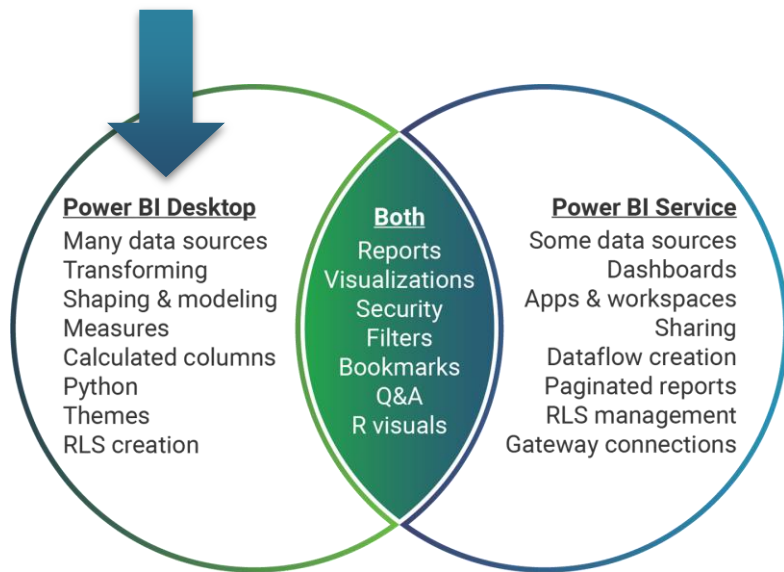
- Get Data:** A screenshot of the 'Get Data' pane showing a list of data sources. A blue callout bubble with the text "70+ Sources!" points to the list. The sources include Excel, Text/CSV, XML, JSON, Folder, PDF, SharePoint folder, SQL Server database, Access database, SQL Server Analysis Services database, Oracle database, IBM Db2 database, IBM Informix database (Beta), IBM Netezza, MySQL database, and PostgreSQL database. A 'Connect' button is visible at the bottom.
- Visualizations:** A screenshot of the 'Visualizations' pane showing a grid of visualization icons. A green callout bubble with the text "Download More!" points to the grid. Below the grid is a search bar with the text "Ask a question about your data".
- Filters:** A screenshot of the 'Filters' pane showing a search bar and two sections: "Filters on this page" and "Filters on all pages". Each section has a button labeled "Add data fields here". An orange callout bubble with the text "Slice & Dice" points to the "Filters on all pages" section.

Best Capabilities in Microsoft Power BI Desktop: Ingest, Build, Save, & **REFRESH!**

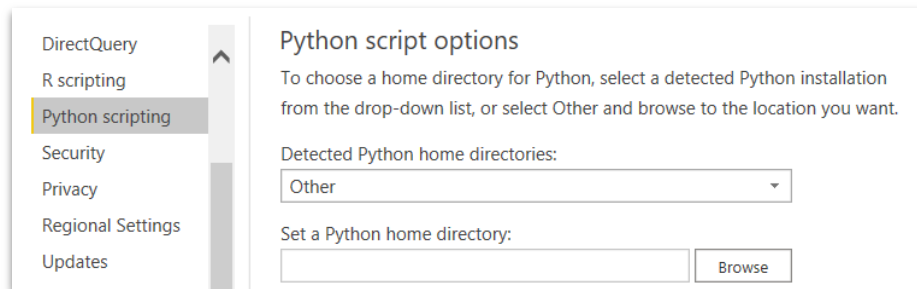
# Big Data Tools (Free-ish is Good)



## Microsoft Power BI (Desktop): Capabilities & Limitations



Ref: <https://docs.microsoft.com/en-us/power-bi/fundamentals/service-service-vs-desktop>

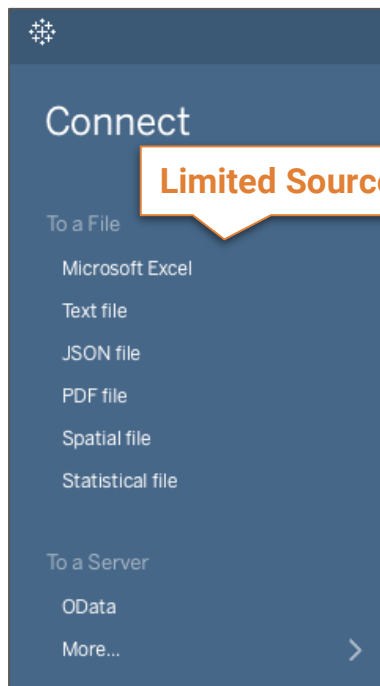


### Power BI Key Limitations:

- 1GB data ingest
- Records/rows limits depend on data source size
- Cannot save visualization directly (but you export as PDF!)
- Sharing visualizations requires receiver to have Power BI

# Big Data Tools (Free-ish is Good)

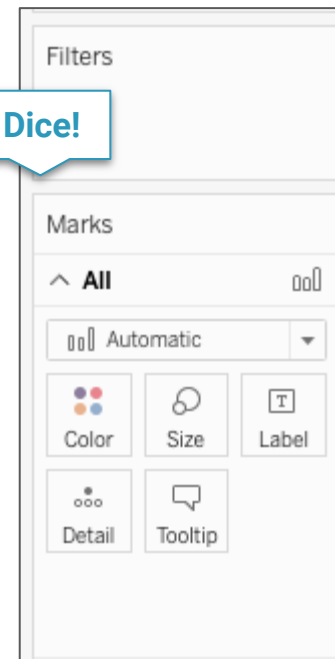
## Tableau Public (Desktop & Online): Capabilities and Limitations



WYSIWYG!



Slice & Dice!



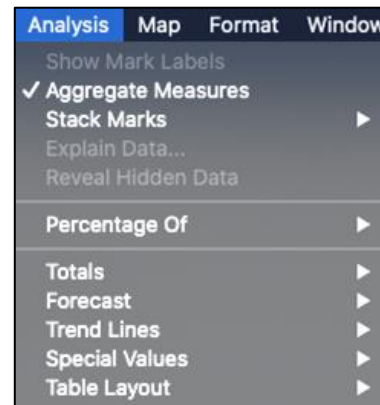
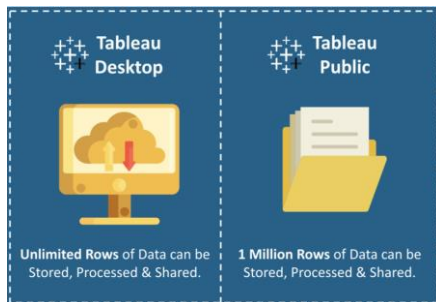
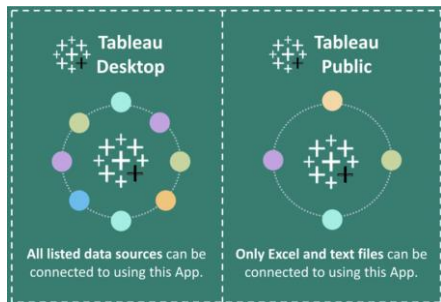
Best Capabilities in Tableau Public Desktop: JSON file handling & Sharing\*



# Big Data Tools (Free-ish is Good)



## Tableau Public (Desktop & Online): Capabilities and Limitations



### Tableau Public Key Limitations:

- 1M records/rows limits
- Limited data sources
- **\* RISK:** To save your visualizations, they are saved to the Tableau Public Cloud (you can restrict access).

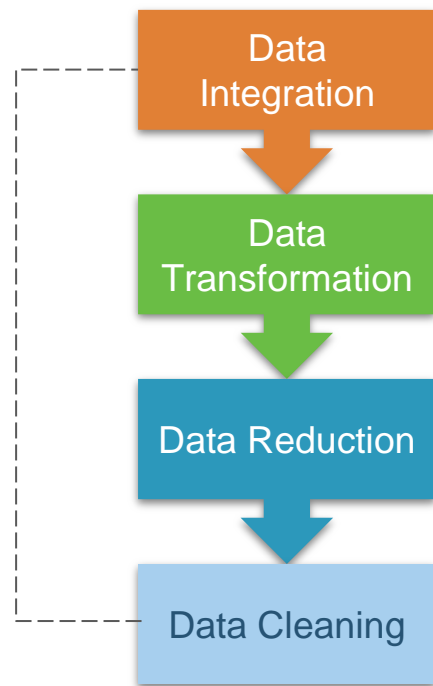
Ref: <https://www.edureka.co/blog/tableau-desktop-vs-tableau-public-vs-tableau-reader/>

# Big Data Tools (Free-ish is Good)



# Normalizing Data (Garbage In = Garbage Out)

- What are your sources?
- Are there **common** fields in your sources?
- If there are common fields, is the data **format** the same?
- Are there logical **pivot** points?
- Are there **duplicates** in the data? How do you dedupe?
- **Goal: Extract, Transform, Load (ETL)**



# Visualizing an Environment (Wireshark and Power BI)

## Data Collection - Architecture



# Visualizing an Environment (Wireshark and Power BI)

“Sensors” - the Tap!

Easy!

Passive!

1G Bandwidth!



# Visualizing an Environment (Wireshark)

## Wireshark View - WYSIWYG

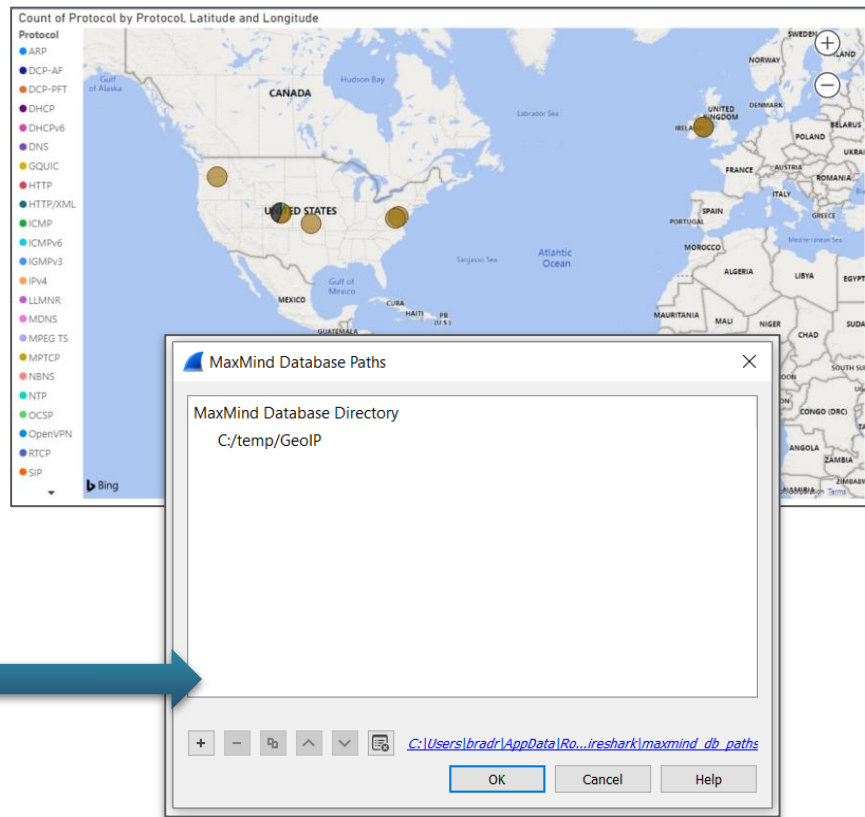
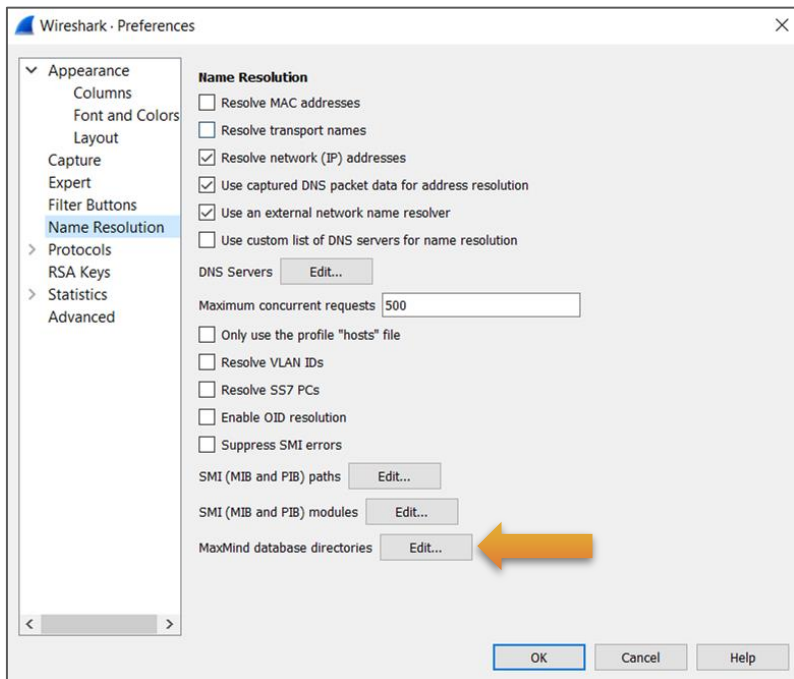
The screenshot displays the Wireshark network protocol analyzer interface. The main pane shows a list of captured packets with columns for No., Time, Source, Src Mac, Src Port, Destination, Dest Mac, Dest Port, Protocol, and Length. A blue arrow points to the 'Edit' menu, which is open, showing options like 'Copy', 'Find Packet...', 'Find Next', 'Find Previous', 'Mark/Unmark Packet', 'Mark All Displayed', 'Unmark All Displayed', 'Next Mark', 'Previous Mark', 'Ignore/Unignore Packet', 'Ignore All Displayed', 'Unignore All Displayed', 'Set/Unset Time Reference', 'Unset All Time References', 'Next Time Reference', 'Previous Time Reference', 'Time Shift...', 'Packet Comment...', 'Delete All Packet Comments', 'Configuration Profiles...', and 'Preferences...'. A green arrow points from the 'Preferences...' option to the 'Wireshark - Preferences' dialog box. The 'Columns' tab is selected in the dialog, showing a list of fields that can be displayed in the packet list pane. The 'Show displayed columns only' checkbox is checked.

No.	Time	Source	Src Mac	Src Port	Destination	Dest Mac	Dest Port	Protocol	Length
1	2020-05-22 13:23:11.055946	assets1.xboxlive.com.c.footprint.net	00:01:5c:66:9e:46	80	c-73-34-45-34.hsd1.co.com...	88:3d:24:a7:e4:ed	51637	HTTP	1514
2	2020-05-22 13:23:11.057536	c-73-34-45-34.hsd1.co.comcast.net	88:3d:24:a7:e4:ed	51635	auto.au.download.windowsu...	00:01:5c:66:9e:46	80	TCP	60
3	2020-05-22 13:23:11.059126	assets1.xboxlive.com.c.footprint.net	88:3d:24:a7:e4:ed	51635	auto.au.download.windowsu...	00:01:5c:66:9e:46	80	TCP	60
4	2020-05-22 13:23:11.060716	assets1.xboxlive.com.c.footprint.net	88:3d:24:a7:e4:ed	51635	auto.au.download.windowsu...	00:01:5c:66:9e:46	80	TCP	60
5	2020-05-22 13:23:11.062306	assets1.xboxlive.com.c.footprint.net	88:3d:24:a7:e4:ed	51635	auto.au.download.windowsu...	00:01:5c:66:9e:46	80	TCP	60
6	2020-05-22 13:23:11.063896	assets1.xboxlive.com.c.footprint.net	88:3d:24:a7:e4:ed	51635	auto.au.download.windowsu...	00:01:5c:66:9e:46	80	TCP	60
7	2020-05-22 13:23:11.065486	assets1.xboxlive.com.c.footprint.net	88:3d:24:a7:e4:ed	51635	auto.au.download.windowsu...	00:01:5c:66:9e:46	80	TCP	60

Displayed	Title	Type	Fields	Field Occurrence
<input checked="" type="checkbox"/>	No.	Number		
<input checked="" type="checkbox"/>	Time	Time (format as specified)		
<input checked="" type="checkbox"/>	Source	Source address		
<input checked="" type="checkbox"/>	Src Mac	Hw src addr (resolved)		
<input checked="" type="checkbox"/>	Src Port	Source port		
<input checked="" type="checkbox"/>	Destination	Destination address		
<input checked="" type="checkbox"/>	Dest Mac	Hw dest addr (resolved)		
<input checked="" type="checkbox"/>	Dest Port	Destination port		
<input checked="" type="checkbox"/>	Protocol	Protocol		
<input checked="" type="checkbox"/>	Length	Packet length (bytes)		
<input checked="" type="checkbox"/>	Info	Information		
<input checked="" type="checkbox"/>	City	Custom	ip.geoiip.city	0
<input checked="" type="checkbox"/>	Country	Custom	ip.geoiip.country	0
<input checked="" type="checkbox"/>	Latitude	Custom	ip.geoiip.lat	0
<input checked="" type="checkbox"/>	Longitude	Custom	ip.geoiip.lon	0

# Visualizing an Environment (Wireshark and Power BI)

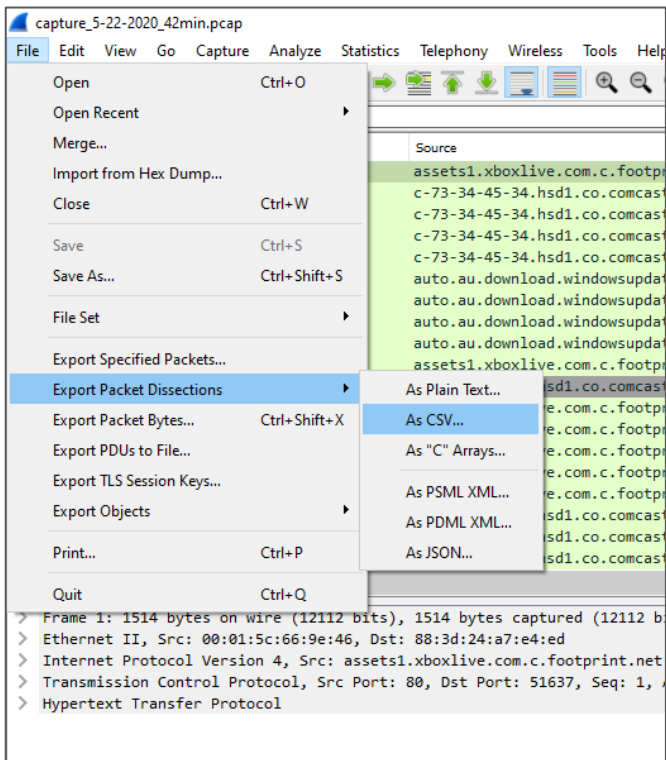
## Wireshark Settings / GeoIP



<https://www.maxmind.com/en/home>

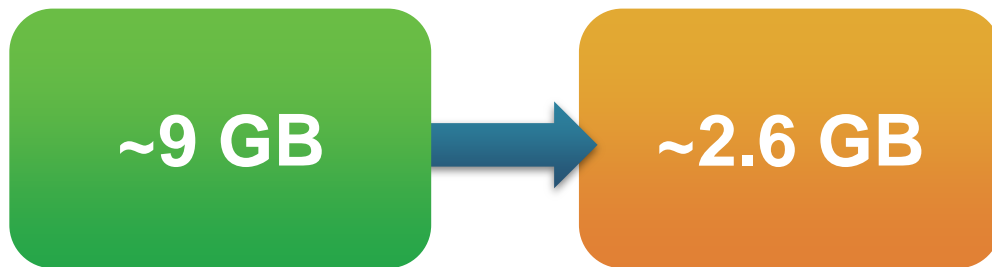
# Visualizing an Environment (Wireshark)

## Export to CSV



Packets: 8687501 · Displayed: 8687501 (100.0%)

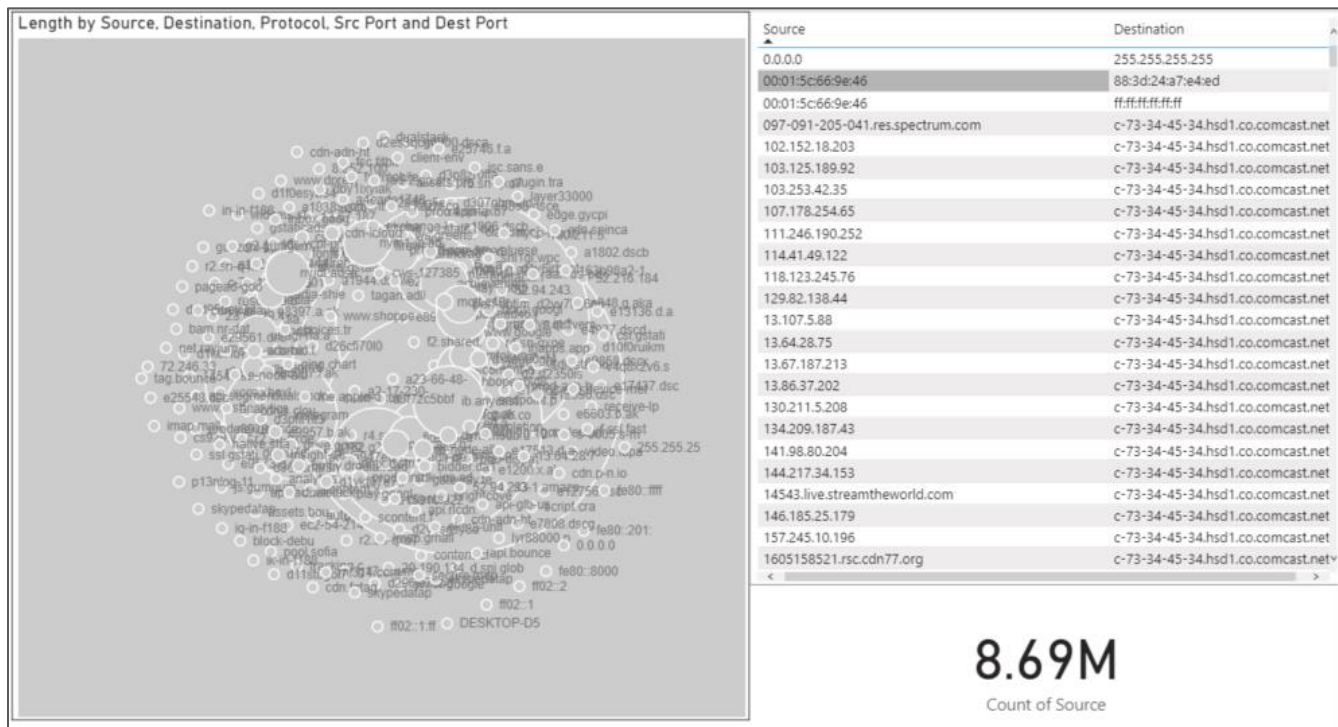
Name	Date modified	Type	Size
capture_5-22-2020_42min.pcap	5/22/2020 2:10 PM	Wireshark capture...	9,303,566 KB
homenet_5-22-2020_42min.csv	5/22/2020 5:37 PM	Microsoft Excel C...	2,774,413 KB





# Visualizing an Environment (Power BI)

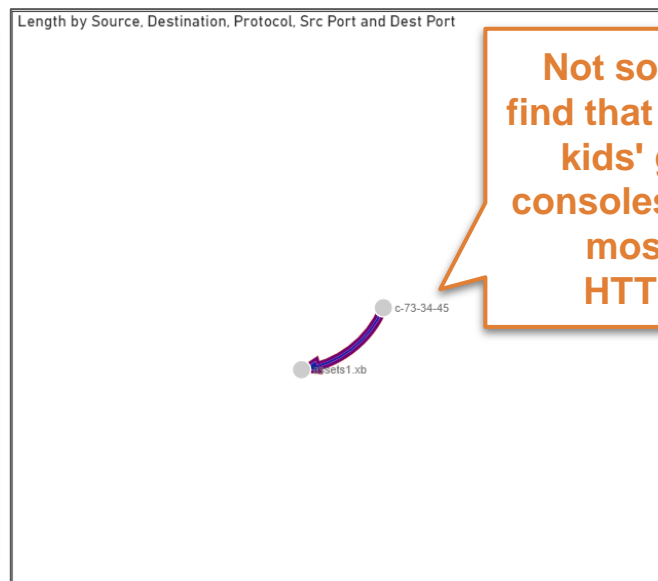
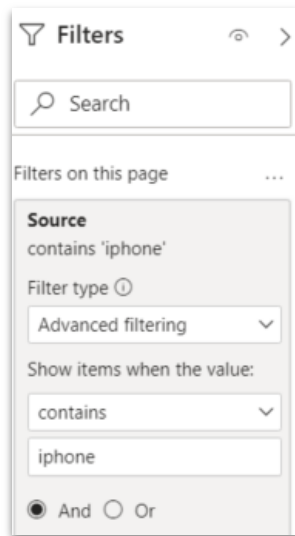
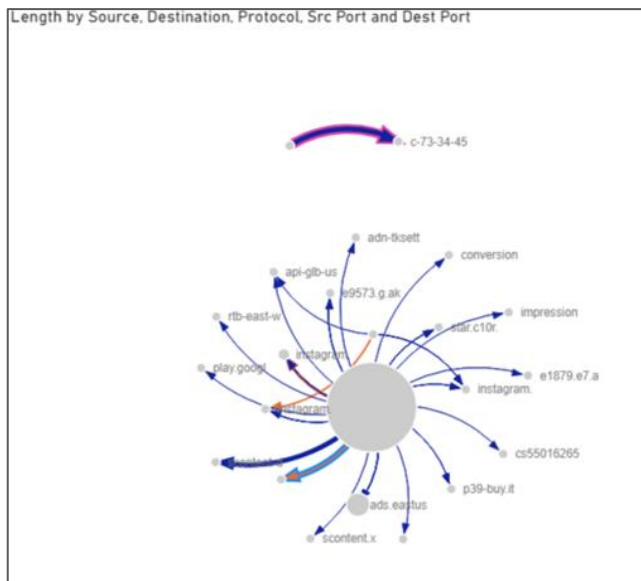
## Network Visualization in Power BI - Why?



# Visualizing an Environment (Power BI)

## Network Visualization in Power BI - Why?

- Quickly filter/sort to focus your investigations!

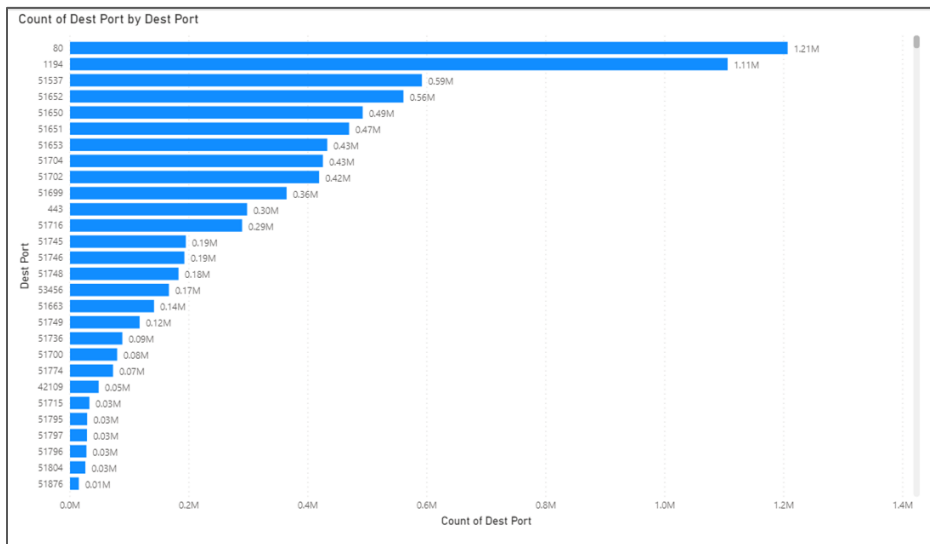


**Not so great to find that one of the kids' gaming consoles talks out mostly on HTTP/80...**

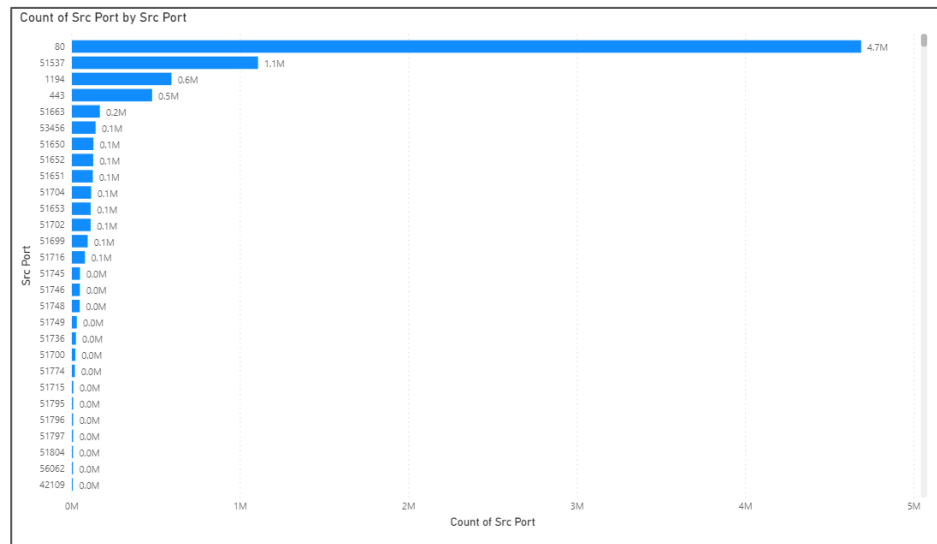
# Visualizing an Environment (Power BI)

## Network Visualization in Power BI - why?

- Quickly conduct long tail analysis!



Destination Port

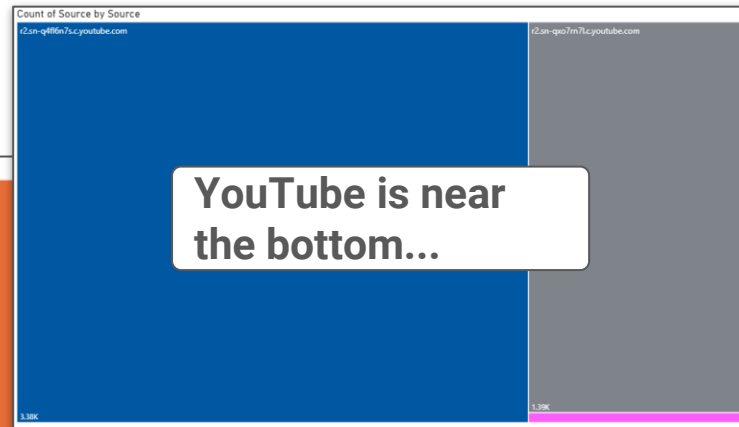
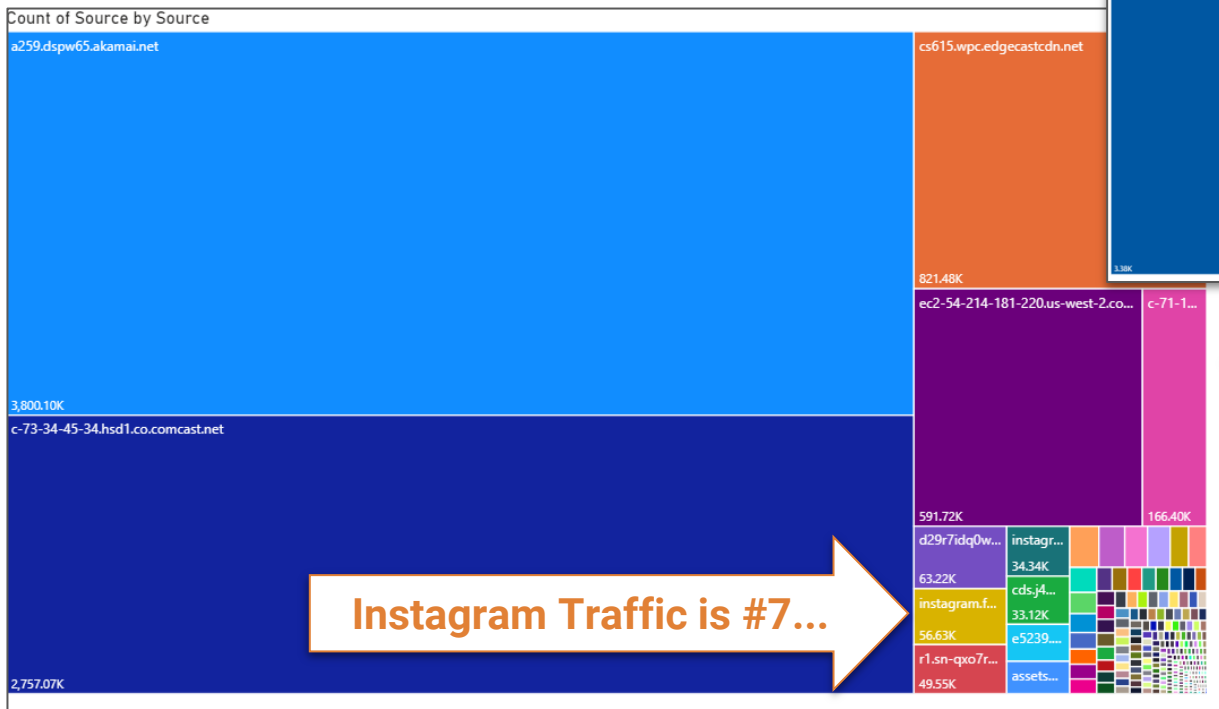


Source Port

# Visualizing an Environment (Power BI)

## Network Visualization in Power BI - Why?

- Quickly discover bulk flows!



## Network Visualization in Power BI - Why?

- [illegible]



# Let's see it live...

---

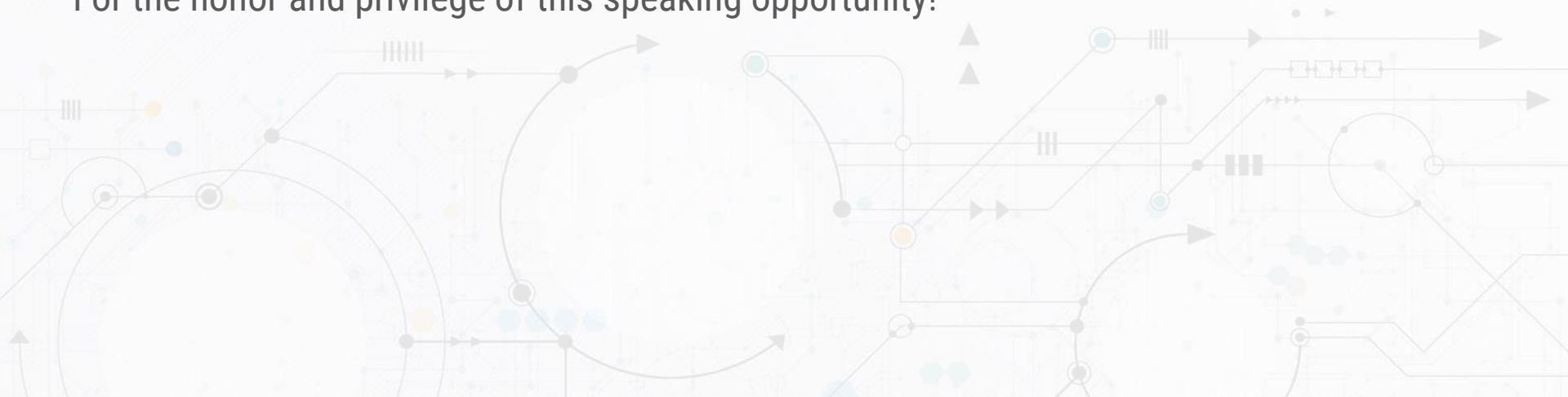
*A “PCAP” of badness (what do we find) using Power BI & Tableau Public*

# Resources for You!

- Power BI - <https://powerbi.microsoft.com/en-us/desktop/>
- Tableau Public - <https://public.tableau.com/en-us/s/>
- Pandas - <https://pandas.pydata.org/>
- Wireshark - <https://www.wireshark.org/>
- Elastic - <https://www.elastic.co/>
- Security Onion - <https://securityonion.net/>

# Thank you, Peak Cyber Symposium!

For the honor and privilege of this speaking opportunity!





# Questions?



**Presentation on  
GitHub:**  
[https://github.com/  
cyberguy514/pres  
entations](https://github.com/cyberguy514/presentations)

## Contact Details:

Civilian: [brhodes@zvelo.com](mailto:brhodes@zvelo.com)

Military: [brad.e.rhodes.mil@mail.mil](mailto:brad.e.rhodes.mil@mail.mil)

MCPA: [brad.rhodes@milcyber.org](mailto:brad.rhodes@milcyber.org)

LinkedIn: <https://www.linkedin.com/in/brad-rhodes-1951ba7/>

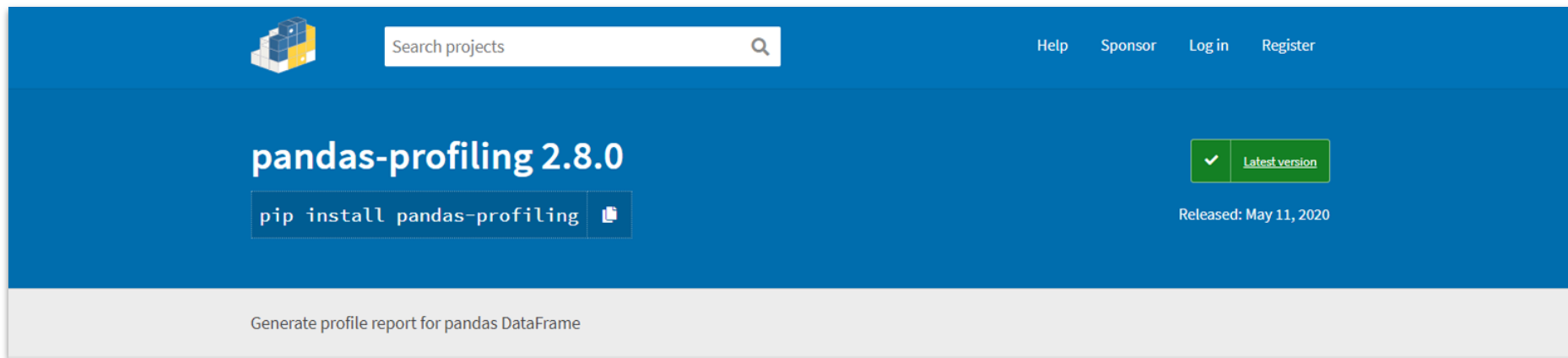
Twitter: [@cyber514](https://twitter.com/cyber514)

# Bonus Materials

The background of the slide features a complex, abstract graphic. It consists of a network of thin, light gray lines that resemble circuit traces or data paths. These lines are interconnected with various nodes, some of which are represented by small circles in shades of blue, yellow, and gray. There are also larger, faint circular outlines and some rectangular blocks that look like components on a circuit board. The overall aesthetic is technical and digital, fitting the theme of the presentation.

# Introducing Pandas Profiling

- Automated & quick data analysis using Pandas!
- <https://github.com/pandas-profiling/pandas-profiling>
- <https://pypi.org/project/pandas-profiling/>



The screenshot shows the PyPI project page for pandas-profiling 2.8.0. The header is blue with a search bar and links for Help, Sponsor, Log in, and Register. The main content area is also blue and features the project name 'pandas-profiling 2.8.0' in large white text. Below the name is a dark blue button with the text 'pip install pandas-profiling' and a copy icon. To the right, there is a green badge with a checkmark and the text 'Latest version', and below it, the release date 'Released: May 11, 2020'. At the bottom of the page, there is a light gray bar with the text 'Generate profile report for pandas DataFrame'.

Search projects

Help Sponsor Log in Register

**pandas-profiling 2.8.0**

`pip install pandas-profiling`

✓ Latest version

Released: May 11, 2020

Generate profile report for pandas DataFrame

# Windows Logs Analysis with Power BI (for IR)

- Process...
  - Ingest the Windows Log
  - Export CSV
  - Ingest into Power BI
  - Given a LOG with evil — Where is it?
  - But first...



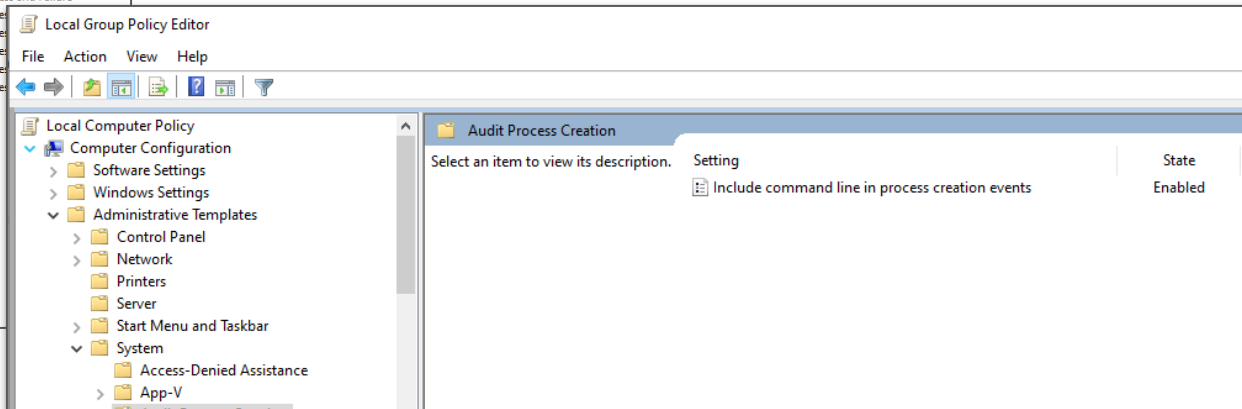
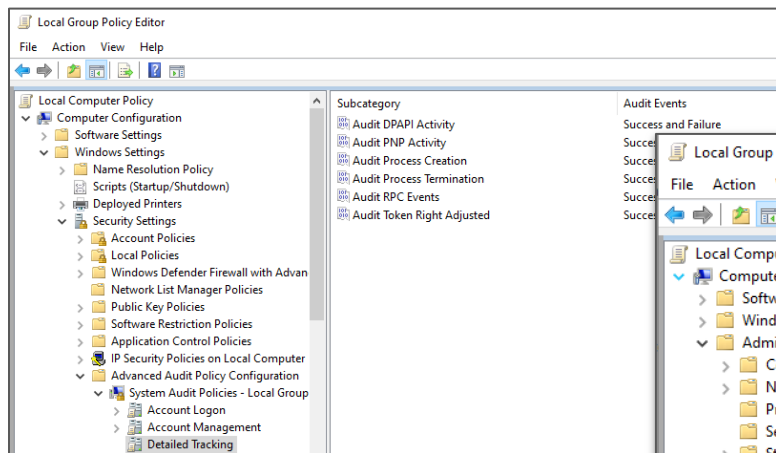
<https://www.nist.gov/cyberframework>

# Windows Logs Analysis with Power BI (for IR)

- Hopefully, you're doing good Windows Logging!?
- At a minimum, turn on detailed tracking...
- And command line logging via GPO!

```
C:\Windows\system32>AuditPol /get /category:*
```

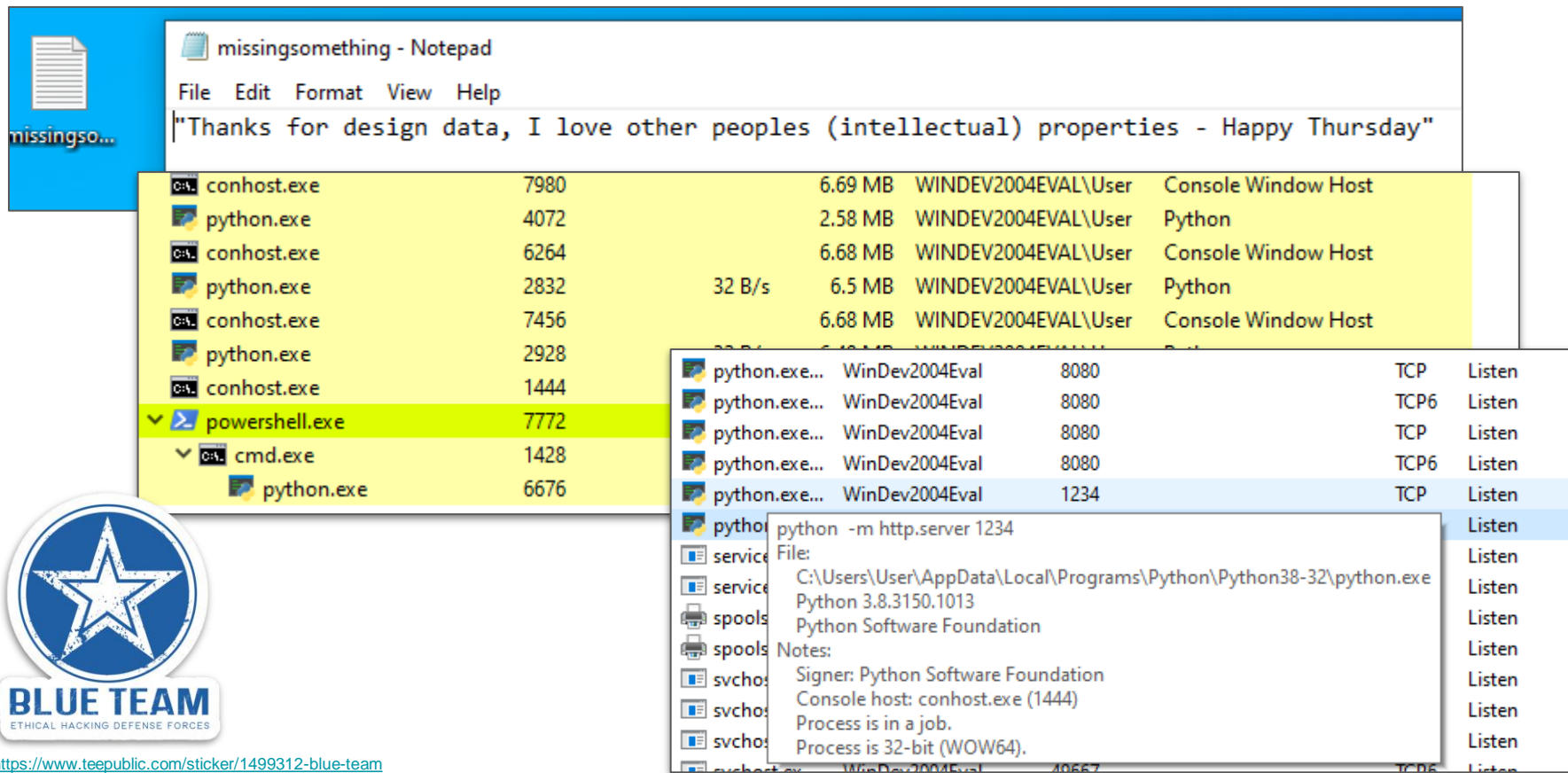
```
Detailed Tracking
Process Creation      No Auditing
Process Termination  No Auditing
DPAPI Activity        No Auditing
RPC Events            No Auditing
Plug and Play Events  No Auditing
Token Right Adjusted  No Auditing
```



<https://www.itprotoday.com/strategy/understanding-and-enabling-command-line-auditing>

<https://www.malwarearchaeology.com/cheat-sheets>

# Windows Logs Analysis with Power BI (for IR)



# Windows Logs Analysis with Power BI (for IR)



<https://www.teepublic.com/sticker/1499286-red-team>

```
root@kali-zv:~# nmap 192.168.56.103
Starting Nmap 7.80 ( https://nmap.org ) at 2020-06-11 08:28 MDT
Nmap scan report for 192.168.56.103
Host is up (0.00076s latency).
Not shown: 998 filtered ports
PORT      STATE SERVICE
1234/tcp  open  hotline
8080/tcp  open  http-proxy
MAC Address: 08:00:27:6C:B7:81 (Oracle VirtualBox virtual NIC)
Nmap done: 1 IP address (1 host up) scanned in 17.83 seconds
```

Directory listing for /

192.168.56.103:1234

Kali Linux Kali Training Kali Tools Kali Docs Kali Forum

## Directory listing for /

- [bash.txt](#)
- [designdata.txt](#)
- [desktop.ini](#)
- [EULA.pdf](#)
- [No AV/](#)
- [Process Hacker 2](#)
- [Visual Studio 201](#)
- [Visual Studio Cod](#)

```
meterpreter > shell
Process 7896 created.
Channel 5 created.
Microsoft Windows [Version 10.0.18363.836]
(c) 2019 Microsoft Corporation. All rights reserved.

C:\Users\User\Desktop>cd ..
cd ..

C:\Users\User\Desktop>powershell
powershell
Windows PowerShell
Copyright (C) Microsoft Corporation. All rights reserved.

Try the new cross-platform PowerShell https://aka.ms/pscore6

PS C:\Users\User\Desktop>

PS C:\Users\User\Desktop>cmd
cmd
Microsoft Windows [Version 10.0.18363.836]
(c) 2019 Microsoft Corporation. All rights reserved.
MAC Address: 08:00:27:6C:B7:81 (Oracle VirtualBox virtual NIC)
C:\Users\User\Desktop>python -m http.server 1234
python -m http.server 1234
```

```
C:\Users\User\Desktop>type missingsomething.txt
type missingsomething.txt
"Thanks for design data, I love other peoples (intellectual) properties - Happy Thursday"
```

# Let's see it live...

---

*Quick Windows Log Analysis in Power BI*



# Windows Logs Analysis with Power BI (for IR)

- Windows Security Events you should **always** care about:
  - 4688 - Process Creation
  - 4689 - Process Exit
- Know your networks, systems, data, users!

```
C:\Windows\system32>AuditPol /get /category:*
System audit policy
Category/Subcategory      Setting
System
  Security System Extension No Auditing
  System Integrity         No Auditing
  IPsec Driver              No Auditing
  Other System Events       No Auditing
  Security State Change     No Auditing
Logon/Logoff
  Logon                     No Auditing
  Logoff                    No Auditing
  Account Lockout           No Auditing
  IPsec Main Mode           No Auditing
  IPsec Quick Mode          No Auditing
  IPsec Extended Mode       No Auditing
  Special Logon              No Auditing
  Other Logon/Logoff Events No Auditing
  Network Policy Server     No Auditing
  User / Device Claims       No Auditing
  Group Membership           No Auditing
Object Access
  File System                No Auditing
  Registry                   No Auditing
  Kernel Object              No Auditing
  SAM                        No Auditing
  Certification Services     No Auditing
  Application Generated      No Auditing
  Handle Manipulation        No Auditing
  File Share                  No Auditing
  Filtering Platform Packet Drop No Auditing
  Filtering Platform Connection No Auditing
  Other Object Access Events No Auditing
  Detailed File Share        No Auditing
  Removable Storage          No Auditing
  Central Policy Staging     No Auditing
Privilege Use
  Non Sensitive Privilege Use No Auditing
  Other Privilege Use Events  No Auditing
  Sensitive Privilege Use     No Auditing
```

```
Detailed Tracking
Process Creation           Success and Failure
Process Termination        Success and Failure
DRAPAPI Activity           Success and Failure
RPC Events                  Success and Failure
Plug and Play Events       Success and Failure
Token Right Adjusted Events Success and Failure
Policy Change
  Audit Policy Change       No Auditing
  Authentication Policy Change No Auditing
  Authorization Policy Change No Auditing
  MPSSVC Rule-Level Policy Change No Auditing
  Filtering Platform Policy Change No Auditing
  Other Policy Change Events No Auditing
Account Management
  Computer Account Management No Auditing
  Security Group Management   No Auditing
  Distribution Group Management No Auditing
  Application Group Management No Auditing
  Other Account Management Events No Auditing
  User Account Management     No Auditing
DS Access
  Directory Service Access    No Auditing
  Directory Service Changes   No Auditing
  Directory Service Replication No Auditing
  Detailed Directory Service Replication No Auditing
Account Logon
  Kerberos Service Ticket Operations No Auditing
  Other Account Logon Events No Auditing
  Kerberos Authentication Service No Auditing
  Credential Validation        No Auditing
```

Turn on more logging...or you'll miss stuff!

<https://www.ultimatewindowssecurity.com/securitylog/encyclopedia/>