NSF Award 1829698
CyberTraining CIP: Cyberinfrastructure Expertise on High-throughput Networks for Big Science Data Transfers

Hands-on Zeek (Bro) Labs

UTSA®
The University of Texas at San Antonio™
The Cyber Center For Security and Analytics

Elias Bou-Harb, Ph.D., CISSP

Elias Bou-Harb

Associate Director
Associate Professor

Expertise

Operational Cyber Security
Cyber Forensics
Critical Infrastructure Security
Data Science
Digital Investigation
Internet Measurements

Research Grants

NSF OAC Core: Small: Devising Data-driven Methodologies by Employing Large-scale Empirical Data to Fingerprint, Attribute, Remediate and Analyze Internet-scale IoT Maliciousness

NSF CyberTraining CIP: Cyberinfrastructure Expertise on High-throughput Networks for Big Science Data Transfers

Recent Publications

Miranda, C., Kaddoum, G., **Bou-Harb, E.**, Garg, S. and Kaur, K., 2020. A Collaborative Security Framework for Software-Defined Wireless Sensor Networks. IEEE Transactions on Information Forensics and Security.

Pour, M.S., Mangino, A., Friday, K., Rathbun, M., **Bou-Harb, E.**, Iqbal, F., Samtani, S., Crichigno, J. and Ghani, N., 2019. On Data-driven Curation, Learning, and Analysis for Inferring Evolving Internet-of-Things (IoT) Botnets in the Wild. Computers & Security, p.101707.

Kaur, K., Garg, S., Kaddoum, G., **Bou-Harb, E**. and Choo, K.K.R., 2019. A Big Data-Enabled Consolidated Framework for Energy Efficient Software Defined Data Centers in IoT Setups. IEEE Transactions on Industrial Informatics.

Neshenko, N., **Bou-Harb, E.,** Crichigno, J., Kaddoum, G. and Ghani, N., 2019. Demystifying IoT security: an exhaustive survey on IoT vulnerabilities and a first empirical look on internet-scale IoT exploitations. IEEE Communications Surveys & Tutorials, 21(3), pp.2702-2733.

Oliveira, D., Ghani, N., Hayat, M., Crichigno, J. and **Bou-Harb, E.**, 2018. SDN testbed for evaluation of large exo-atmospheric EMP attacks. IEEE Communications Magazine, 57(1), pp.88-97.

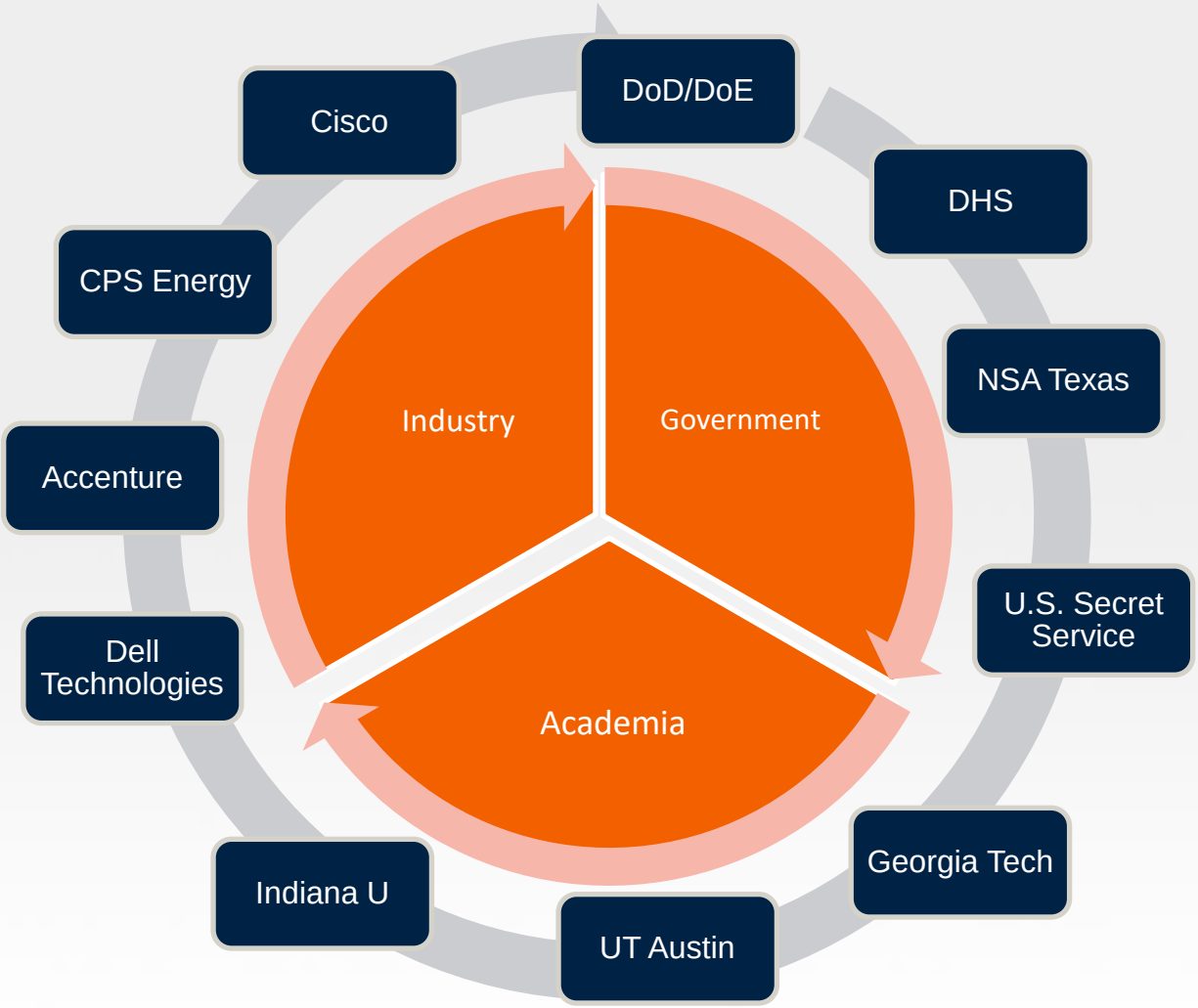
Cyber Threat Intelligence
Digital Forensics

Infrastructure Security

Data Science

Software/ICT Security

Industry-Public Sector-Academia



Vision

R&D Projects

- Blockchain-Based Cyber Security
- Early Detection of Stealthy User Behavior
- Analytics-Driven Profiling of Grid Security
- IoT Threat Analytics

Studies

- IoT Malware
- Cyberterrorism Communities
- Offensive Language Classifiers
- Modeling public response to data breaches

Training

- NSF/DoD Cybertraining
- NSA/CAE Training
- U.S. Army Reserve
- CTF/Hackathons

Impact

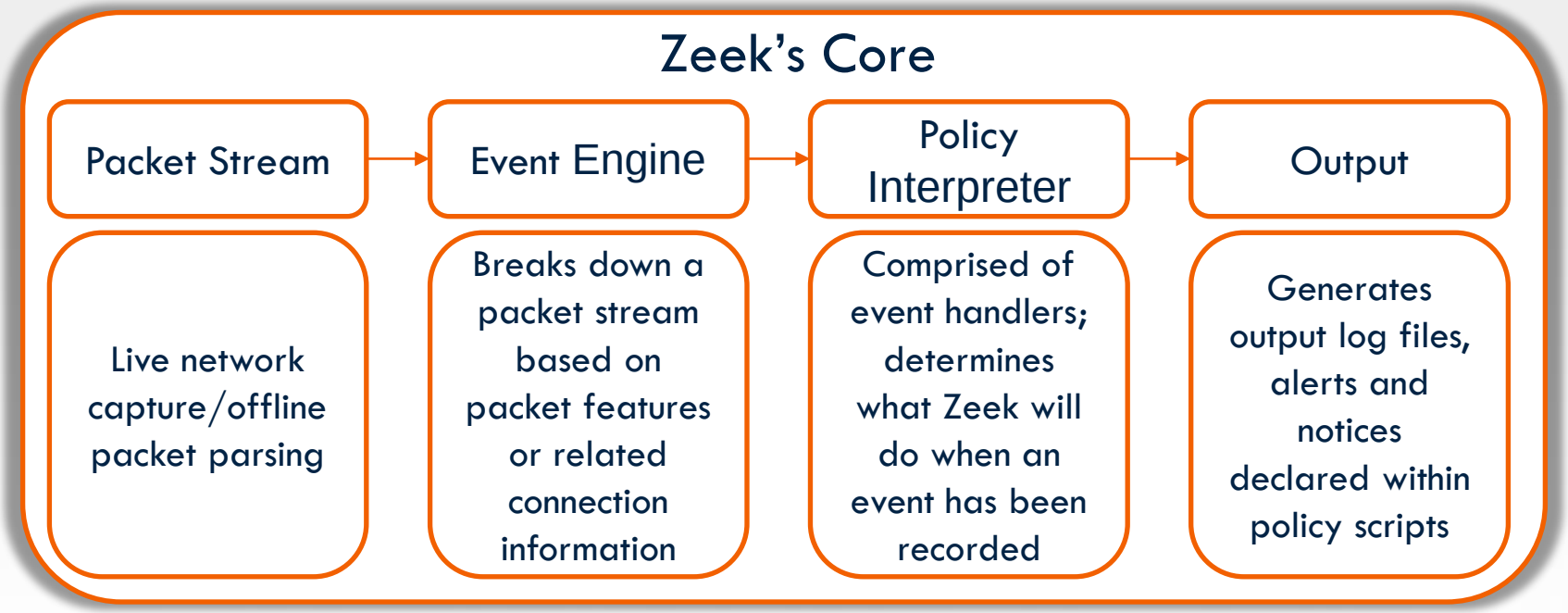
- R, D & T For K12/Minorities
- IoT Exploitations' Database
- Ransomware Prevention
- TTP/Patents

Intro to Zeek (Bro)

Network Intrusion Detection Systems

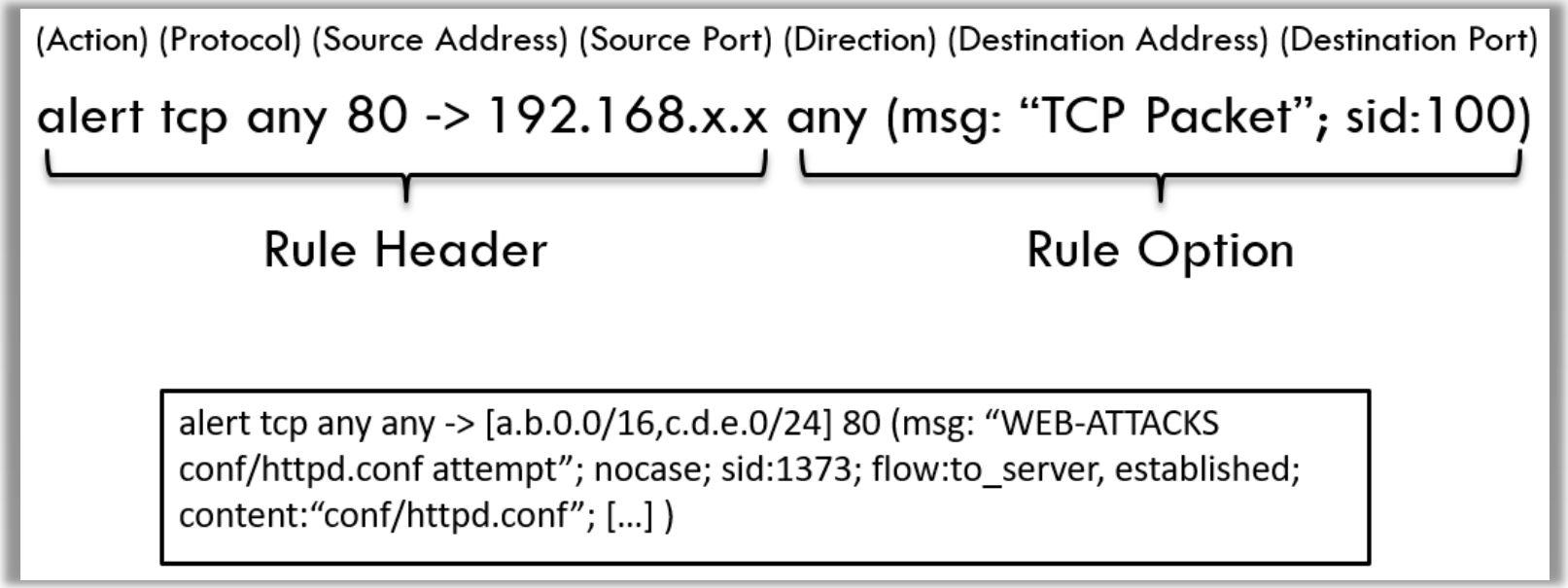
- Software/hardware systems that actively monitor live networks for malicious traffic, policy violations and unidentified anomalies
- Deployed to protect operational networks without disturbing normal/benign packet traffic flows
- In contrast to firewalls, NIDS are most often passive, looking for signatures or anomalies, although they can operate as NIPS as well

- Zeek's Development began in 1995 by Vern Paxon (as Bro)
- Zeek's scripting language creates a versatile environment for fine-grained anomaly-related detection and processing
- Versatile formatting of output data for preprocessing and advanced analytics



Network Traffic Signatures: A *Snort* Signature

Follows a rule-based format



Network Traffic Signatures: A *Zeek* Signature

Follows a variable/data object-based format

Variables support strings, integers and floats

```
signature sid-1371 {
        ip-proto == tcp
        dst-ip == a.b.0.0/16,c.d.e.0/24
        dst-port == 80
payload /.*conf/\httpd\.conf/
tcp-state established, originator
event "WEB-ATTACKS conf/http.conf attempt"
}
```

Zeek Log Files

- After processing network traffic, Zeek will output statistical log files
- By default, log files will be separated by the transport protocol and related characteristics
- At a basic level, these log files can be used to determine the presence of an anomaly
- Zeek log files can be formatted and exported to external processing software

Connection	Protocol-Specific	Detection	Observations
conn.log	http.log	notice.log	known_certs.log
files.log	ftp.log	signatures.log	known_services.log
x509.log	dns.log	traceroute.log	weird.log

Zeek Policy Scripts and Filters

- The Zeek scripting language is used to develop and implement filters and policies for the event-based engine
- Scripts can be implemented to permanently update Zeek's event handling or used as a non-permanent filter
- Script events include (but are not limited to):
  - Protocol-specific events
  - Application-level headers
  - Unknown/broken connection handling
- Packet data is accessible within the filters to be used for calculations or to be exported into separate log files

Zeek Policy Scripts and Filters

```
event udp_request(u:connection){
        print fmt("A UDP Request was found!");
        print fmt("Source Address: %s Destination Port: %s",
                u$id$orig_h, u$id$resp_p);
}
event udp_reply(u: connection){
        print fmt("A UDP Reply was found!");
        print fmt("Source Address: %s Destination Address: %s",
                u$id$orig_h, u$id$resp_h);
}
```

Protocol-oriented Zeek Filter

Custom-based detectors

```
export {
const addr_scan_interval = 5min &redef;
const addr_scan_threshold = 20 &redef;
}
function horizontal_scanning(c: connection):bool {
        if (num_requests(c$id$orig_h) > addr_scan_threshold &&
                time_alive(c$connection) < addr_scan_interval) {
                print fmt("Horizontal Scanner Detected!");
                return c$id$orig_h;
        }
}//end function
```

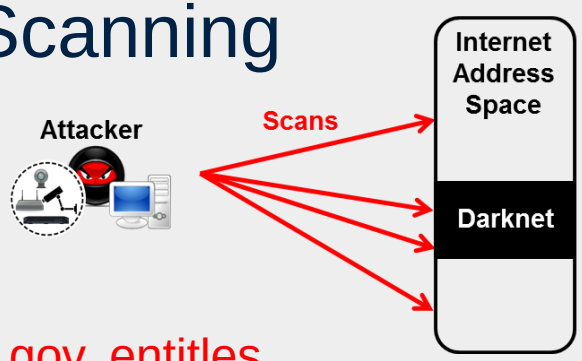
Zeek Inferring IoT-generating Scanning

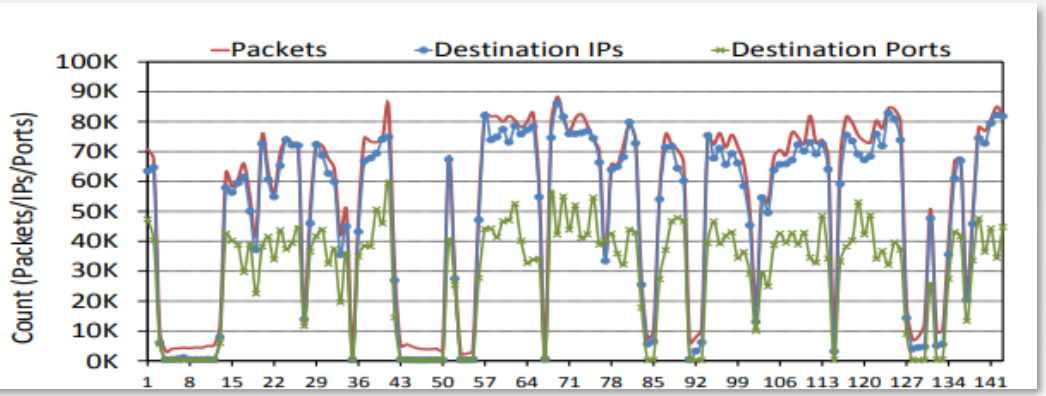
The Insecurity of the IoT Paradigm

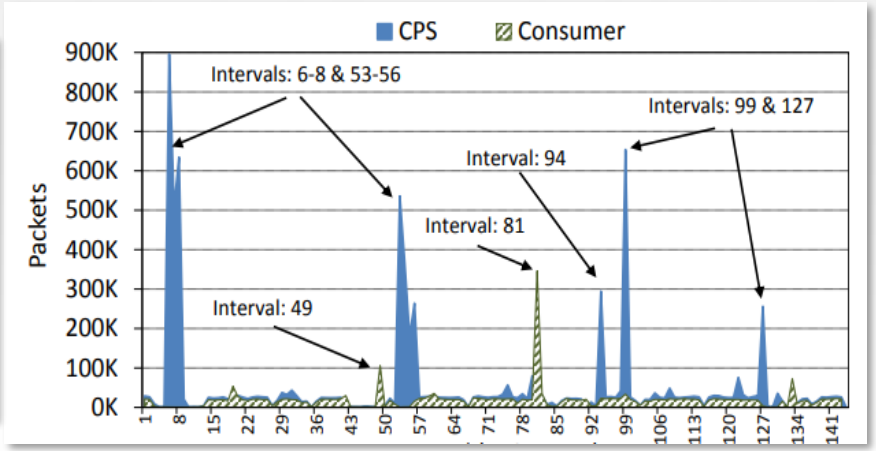
Zeek Inferring IoT-generating Scanning

Malicious scans from compromised IoT devices

- 2 TB of Darknet Data (Daily)
- 840K global IoT exploitations (25K in the US)
- Exploitations in health services, manufacturing plants, gov. entitles



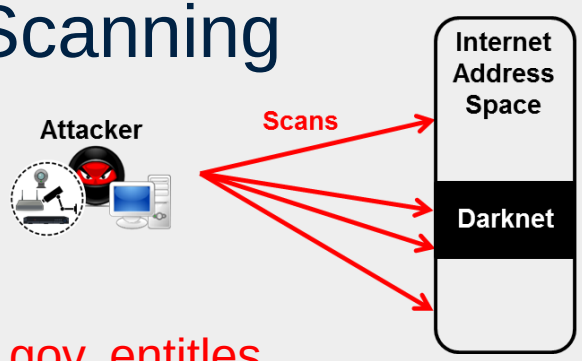




Zeek Inferring IoT-generating Scanning

Malicious scans from compromised IoT devices

- 2 TB of Darknet Data (Daily)
- 840K global IoT exploitations (25K in the US)
- Exploitations in health services, manufacturing plants, gov. entitles



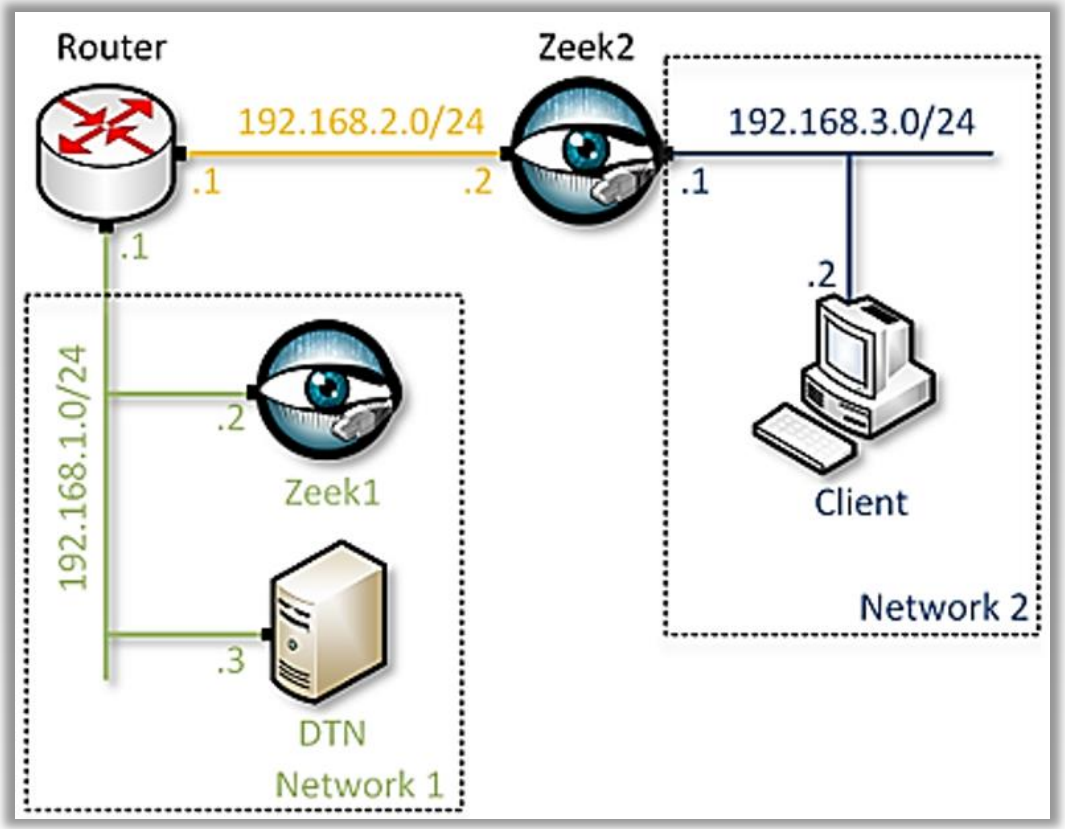
Date	2019-10-08	2019-11-25	2020-03-11	2020-04-26
# all infected hosts	754,169	836,255	806,326	839,082 (752,348)
# Compromised IoT	274,699	229,488	224,964	480,049 (405,184)
# all infected hosts (in USA)	16,614	15,957	23,779	25,468 (16,981)
# Compromised IoT (in USA)	6,569	5,489	8,541	12,909 (4,920)
# infected hosts **in Medical**	131	6	160	**323 (311)**
# Compromised IoT **in Medical**	17	0	10	**76 (58)**
# infected hosts **in Medical (US)**	26	2	22	**58 (54)**
# Compromised IoT **in Medical (US)**	3	0	2	**11 (10)**

Hands-on Zeek (Bro) Labs

Zeek Hands-on Labs

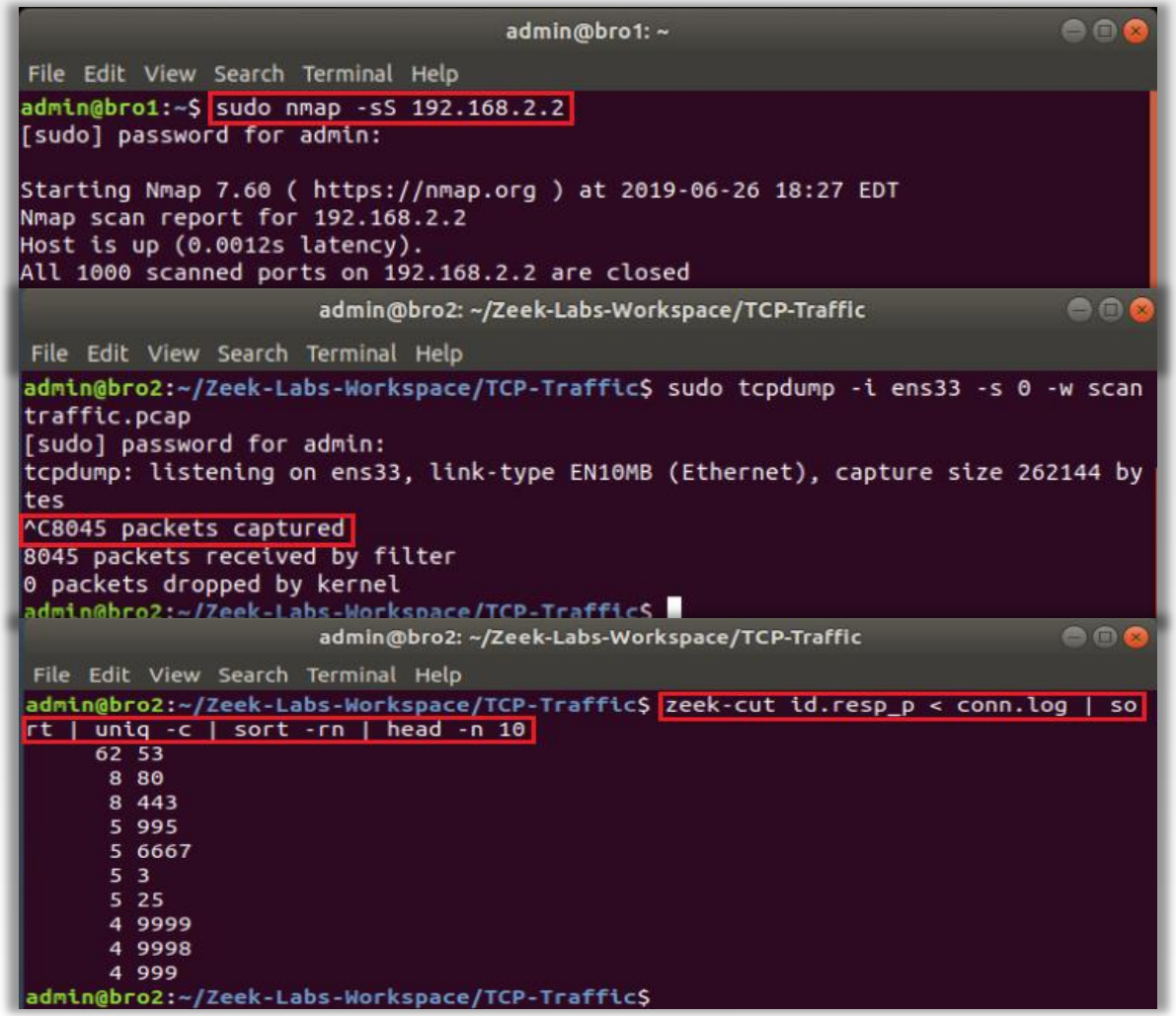
- Lab 1- Introduction to the Capabilities of Zeek
- Lab 2 - An Overview of Zeek Logs
- Lab 3 - Parsing, Reading and Organizing Zeek Files
- Lab 4 - Generating, Capturing and Analyzing Network Scanner Traffic
- Lab 5 - Generating, Capturing and Analyzing DoS and DDoS-centric Network Traffic
- Lab 6 - Introduction to Zeek Scripting
- Lab 7 - Advanced Zeek Scripting for Anomaly and Malicious Event Detection
- Lab 8 - Preprocessing of Zeek Output Logs for Machine Learning
- Lab 9 - Developing Machine Learning Classifiers for Anomaly Inference and Classification
- Lab 10 - Profiling and Performance Metrics of Zeek

Lab 4 - Generating, Capturing and Analyzing Network Scanner Traffic

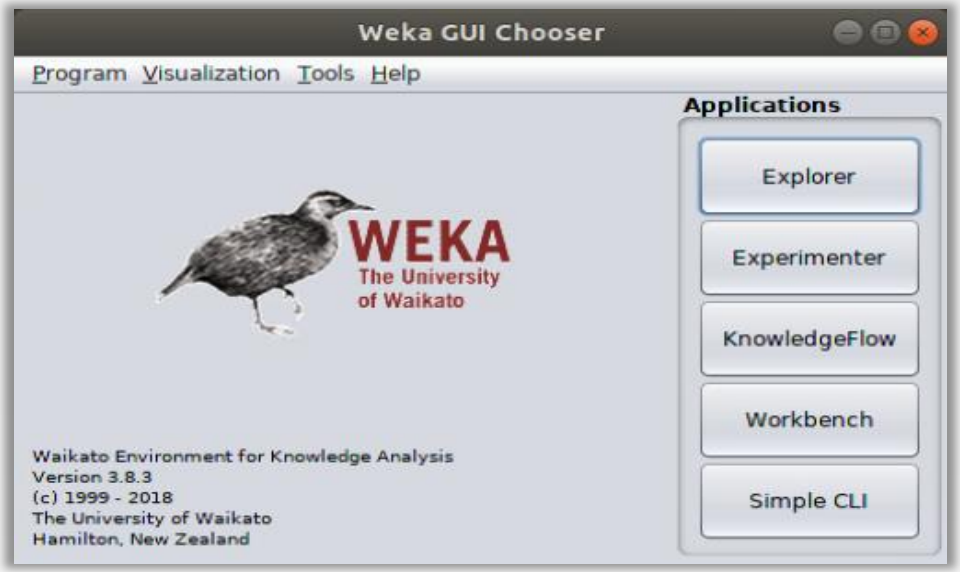


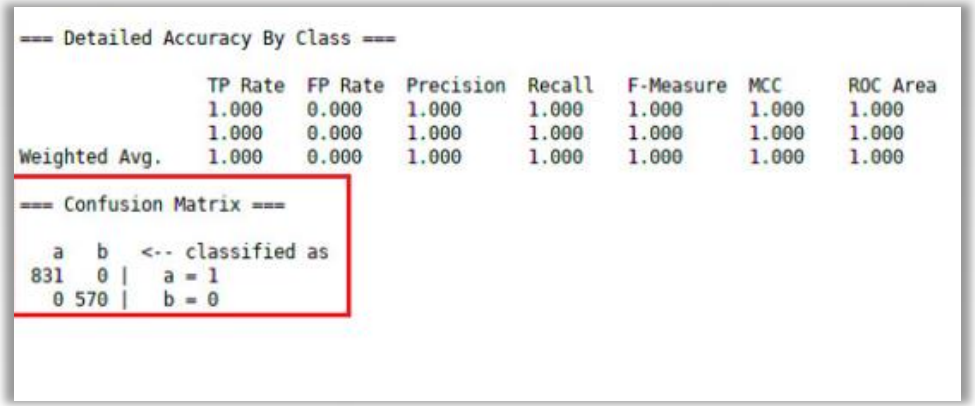
Virtual Machine	IP Address	Account	Password
Zeek1	192.168.1.2	admin	password
DTN	192.168.1.3	root	password
Client	192.168.3.2	root	@dmin123
Zeek2	192.168.2.2 192.168.3.1	admin	password
Router	192.168.1.1 192.168.2.1 203.0.113.2	root	password

Lab 4 - Generating, Capturing and Analyzing Network Scanner Traffic



Lab 9 - Developing Machine Learning Classifiers for Anomaly Inference and Classification



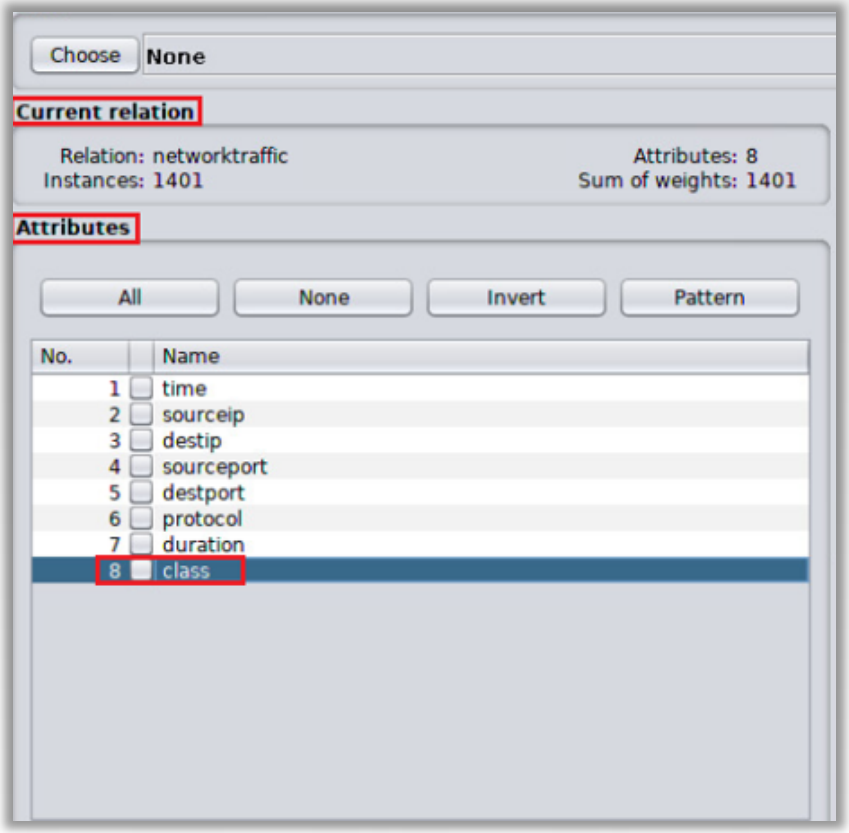


```
=== Detailed Accuracy By Class ===

                 TP Rate  FP Rate  Precision  Recall  F-Measure  MCC      ROC Area
                 1.000    0.000    1.000      1.000   1.000      1.000    1.000
                 1.000    0.000    1.000      1.000   1.000      1.000    1.000
Weighted Avg.    1.000    0.000    1.000      1.000   1.000      1.000    1.000

=== Confusion Matrix ===

   a   b   <-- classified as
 831   0 |   a = 1
   0 570 |   b = 0
```



NSF Award 1829698
CyberTraining CIP: Cyberinfrastructure Expertise on High-throughput Networks for Big Science Data Transfers

Hands-on Zeek (Bro) Labs

Enjoy the Labs ☺
Thanks for the support!

Elias Bou-Harb, Ph.D., CISSP
elias.bouharb@utsa.edu