Reproducible and Adaptable Log Data Generation for Sound Cybersecurity Experiments

Rafael Uetz
Christian Hemminghaus
Louis Hackländer
Philipp Schlipper
Martin Henze

SPONSORED BY THE

Fraunhofer FKIE

RWTH AACHEN UNIVERSITY

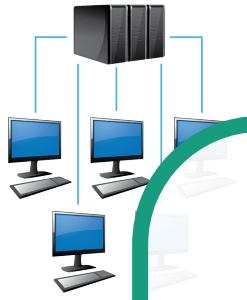
Federal Ministry of Education and Research

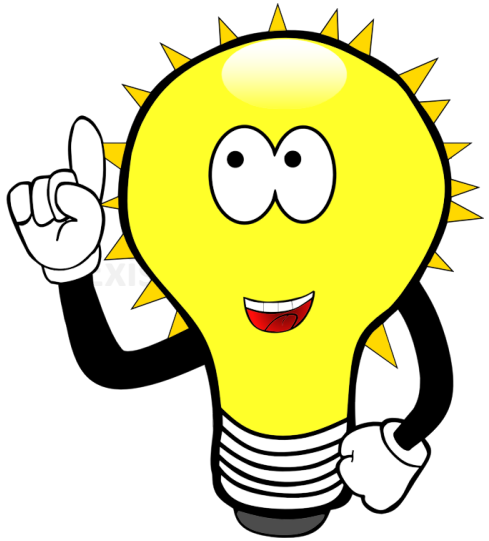
Introduction

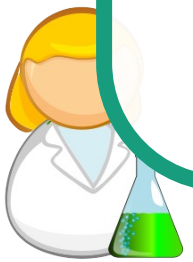
- Numerous organizations are hit by data breaches and ransomware attacks
- Timely detection and analysis are vital to limit the damage
- Log data are an invaluable source for detection and analysis

Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

Motivation: Log Data Sources and Their Challenges







- No access or no permission to perform experiments
- Others cannot adapt the testbed to build on results

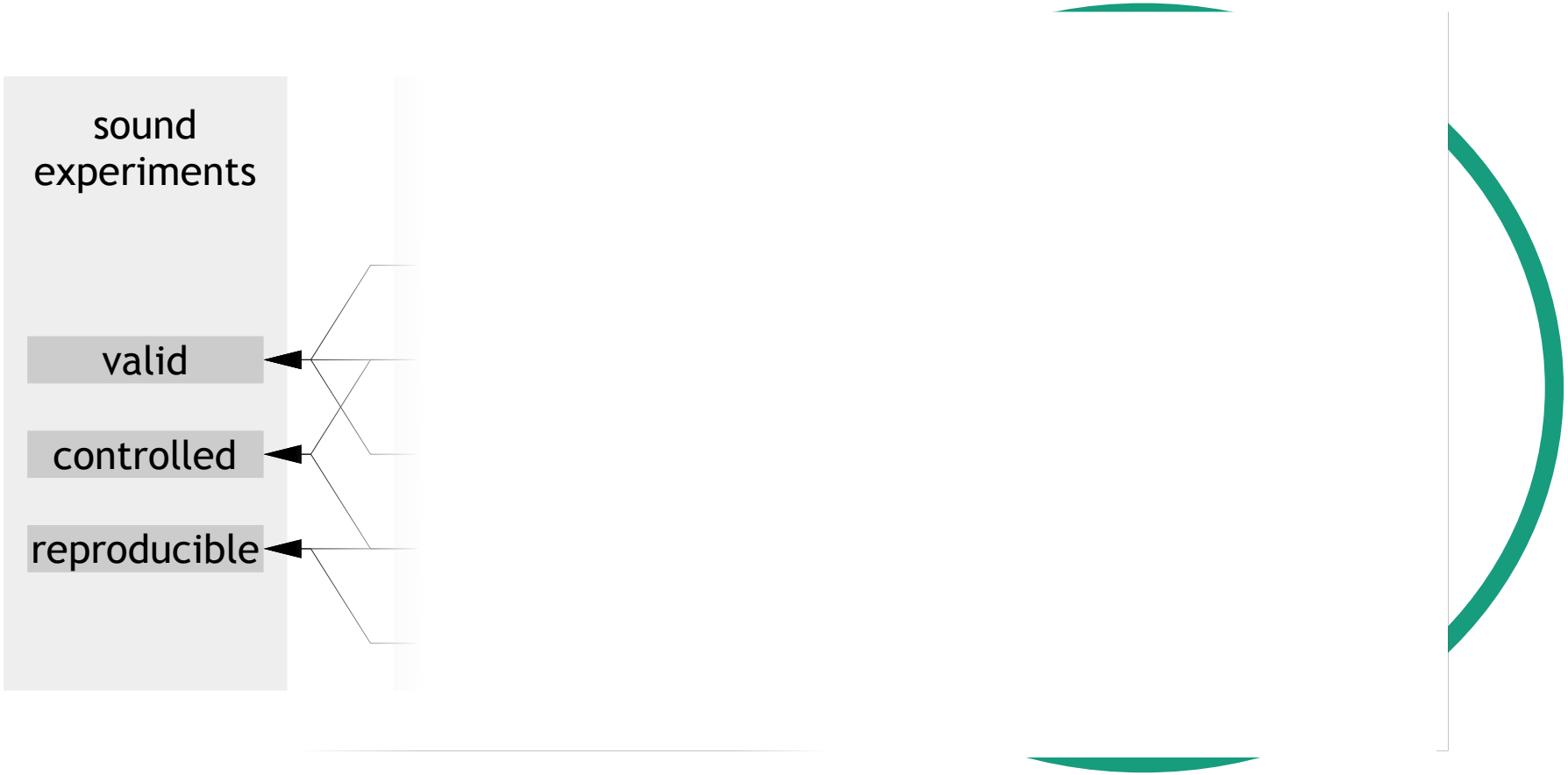
The scientific community requires

reproducible and adaptable testbeds

for log data generation

Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

Testbed Requirements



Rafael Ue @ rafael.uetz@fkie.fraunhofer.de 37z u37

Proof-Of-Concept Testbed: SOCBED

Self-Contained Open-Source Cyberattack Experimentation Testbed

Main features

- Simulation of a small company's network using virtual machines
- Simulation of user activity (web surfing, emailing, file editing)
- Simulation of common multi-step cyberattacks
- Best-practice log data collection
- Optional collection of network traffic

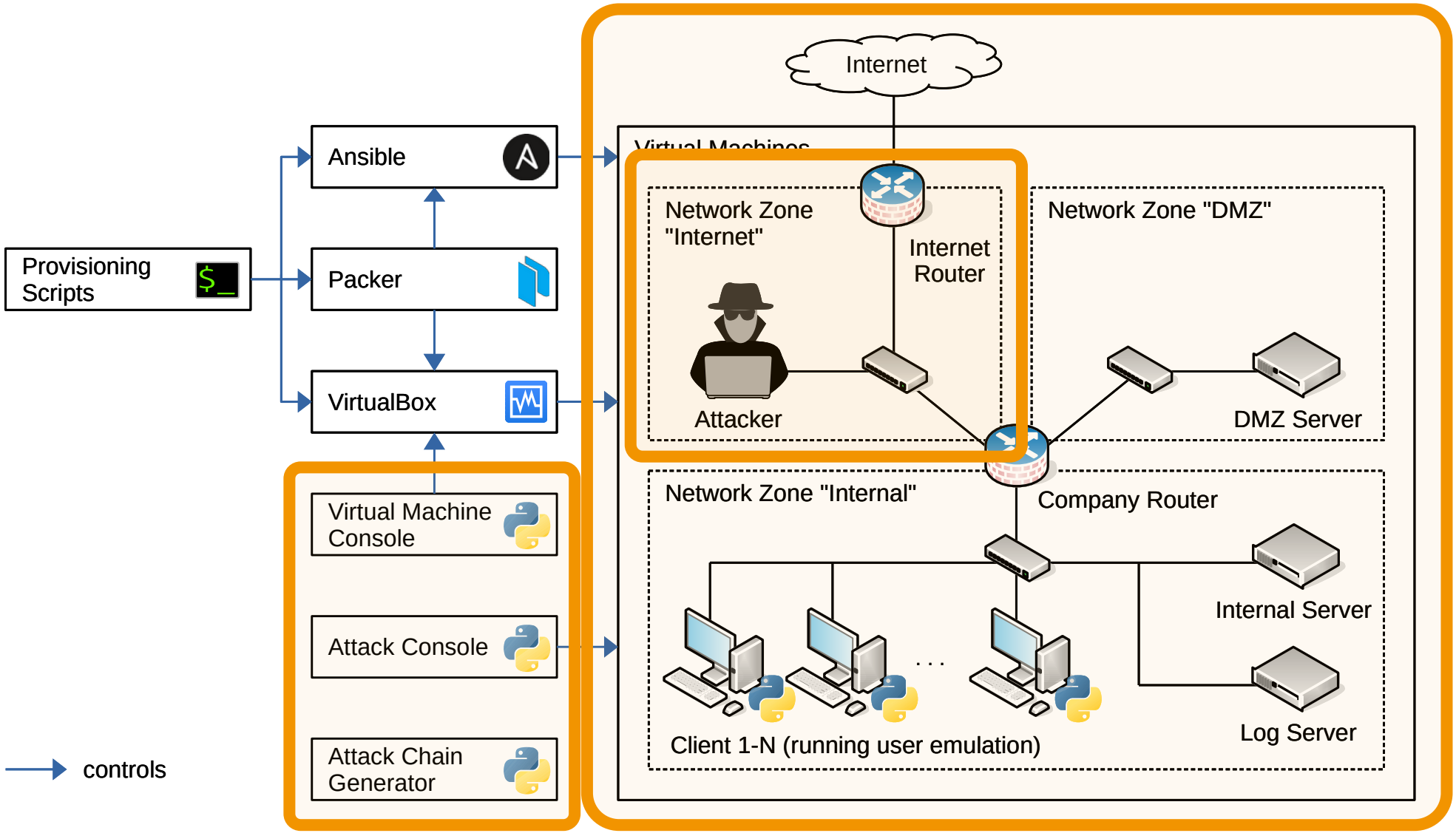
fkie-cad/socbed

Distinctive design principles

- Runs on commodity computers, open source, full infrastructure as code
- Self-contained, numerous self-tests (unit and system)
- Deterministic user activity and fully scriptable attacks

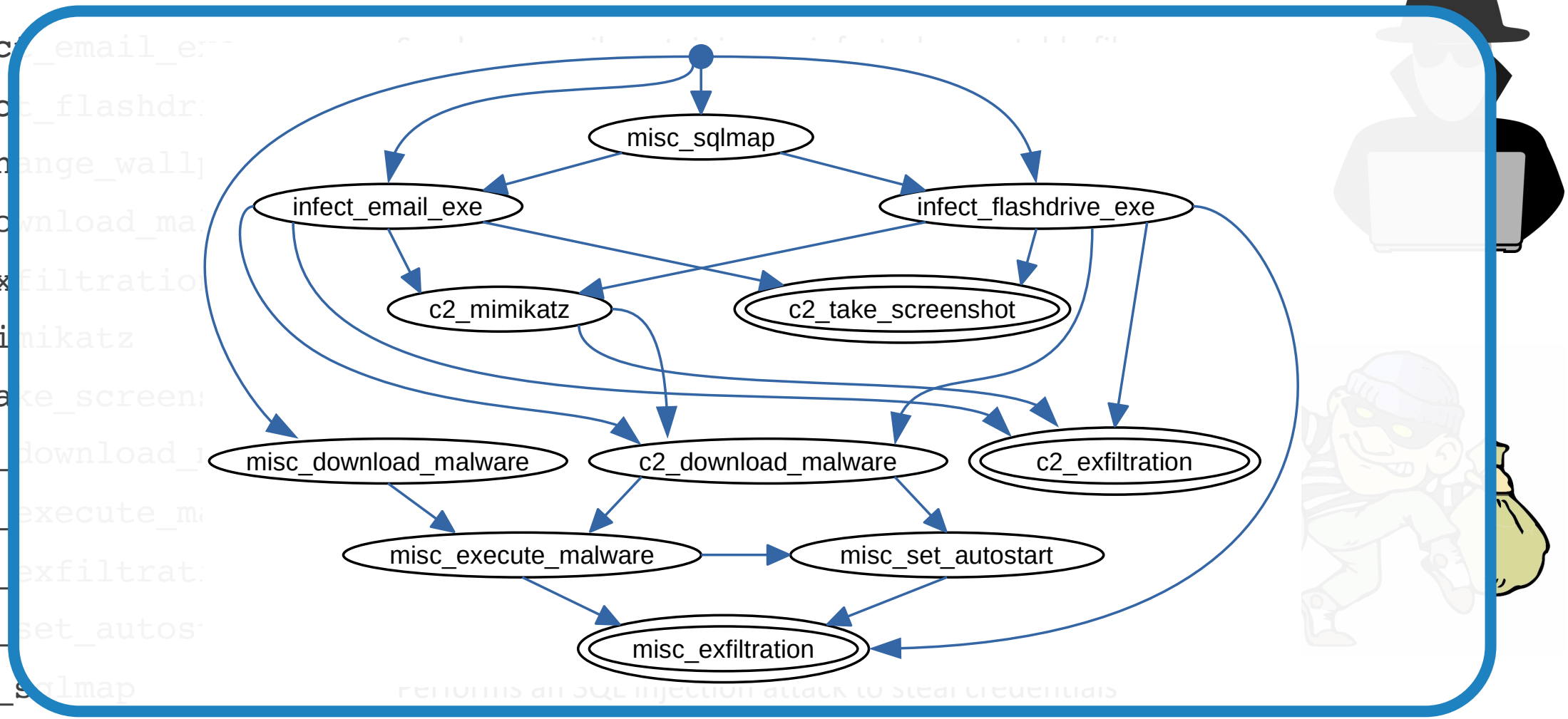
Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

SOCBED Architecture



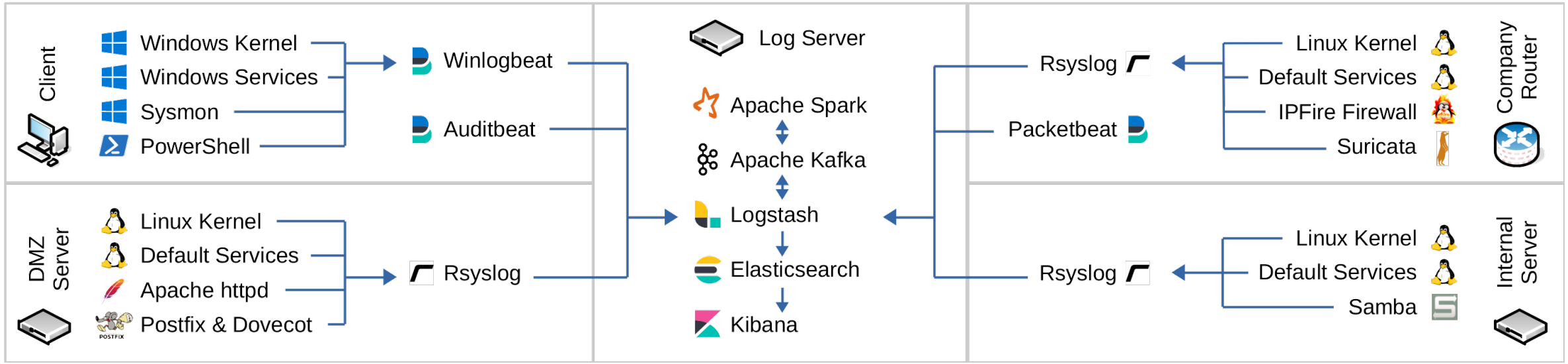
Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

SOCBED Cyberattacks



Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

SOCBED Logging Facilities



- Log data collection from numerous sources
- Storage, search and visualization on a dedicated log server
- Easy download to create log datasets

Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

Evaluation

Goal	Show that SOCBED facilitates sound experiments
Approach	Show that exemplary experiment is valid, controlled, and reproducible
Experiment	Compare logging configs for detection of multi-step attack
Iterations	Repeated ten times on two independent hosts
Results	Consistent attack detection under reproduction and adaptation
Conclusion	**SOCBED facilitates sound experiments**

Attack step
(1) Scan and exploit web server
(2) Send email with malware
(3) Open malicious attachment
(4) Capture screen
(5) Collect cached credentials
(6) Search network & download files
(7) Download custom backdoor
(8) Set autostart for backdoor
(9) Execute backdoor
Number of detected attack steps

Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de

Summary

- Log data are crucial for cybersecurity research
- Testbeds are an important means for generating log data
- Existing testbeds lack reproducibility and adaptability
- SOCBED is a proof-of-concept testbed to facilitate sound experiments
- SOCBED and our evaluation dataset are freely available on GitHub

`fkie-cad/socbed`

`fkie-cad/socbed-eval-acsac-2021`

Please see our paper for more details. Thank you very much for your attention!

Rafael Uetz @ rafael.uetz@fkie.fraunhofer.de