PoCyMa - Ein Ansatz für rekursive KI-basierte Entscheidungsunterstützung bei IT Vorfällen

Erläuterung der Problemstellung, die mit der Idee gelöst werden soll:

IT-Vorfälle erfordern mehr und mehr das Finden schneller fundierter Entscheidungen, da die Bedeutung einer intakten IT Infrastruktur unerlässlich ist. Dabei sind zumeist mehrere Hierarchiebenen, wie technisches und Managementpersonal, involviert. Dies erfordert unterschiedliche Sichtweisen auf den Vorfall. Technisches Personal benötigt beispielsweise detailiertere Einsichten in den Vorfall, z.B. welche Systeme und Dienste sind in welcher Weise betroffen, wohingegen Managementpersonal in erster Linie eine Lagebildübersicht benötigt. Der Vorfall muss entsprechend auf unterschiedliche Weisen darstellbar sein.

Des Weiteren ist es für fundierte Entscheidungen von Bedeutung, frühere Entscheidungen und deren Folgen mit in Betracht zu ziehen. Auch Handlungsweisen Dritter bei ähnlichen Situationen können entscheidende Informationen liefern.

Diese Schritte sind in der Regel zeitaufwendig, was bei entsprechenden Vorfällen zu erheblichen finanziellen Einbußen führt.

Daraus ergibt sich die Problemstellung, wie kann bei IT Vorfällen möglichst effizient und automatisiert die Entscheidungsfindung über mehrere Hierarchiebenen hinweg dauerhaft unterstützt werden?

Beschreibung der Idee und wie sie das Problem lösen soll:

PoCyMa (https://github.com/localos/PoCyMa) bietet eine Architektur die den Prozess der Entscheidungsfindung bei IT Vorfällen nachhaltig und auf die jeweilige IT Infrastruktur zugeschnitten, unterstützt.
Die Architektur (vgl. Darstellung in der angehängten PDF Datei) besteht in erster Instanz aus drei Planes: Control Plane, Data Plane und Business Intelligence (BI) Plane. Sie ist flexibel erweiterbar (z.B. Forensics Plane).

Die Control Plane ist für die Konfiguration, Administration und Überwachung der Anwendung vorgesehen. Dadurch kann flexibel, an die Bedürfnisse der Infrastruktur angepasst, festgelegt werden, welche Komponenten der Gesamtarchitektur verwendet werden.

Die Data Plane ist für die Sammlung und Aufbereitung der Datenquellen vorgesehen. Als Grundvoraussetzung ist ein Traffic Analysis Tool welches auf A/B Flows arbeitet, wie z.B. Tranalyzer vorgesehen. Als weitere Datenquellen können beispielsweise Intrusion Detection Systeme, wie Bro oder Snort, und Geolocation Datenbanken, wie Maxmind eingebunden werden. Diese werden durch diverse Wrapper- und Parser-Scripte aufbereitet und in ein für die Anwendung definiertes Datenbankmodell gespeichert. In erster Instanz ist vorgesehen, dass diese Scripte auf dem JSON Format oder syslog-standardkonformem Ausgabeformat arbeiten. Ein Webfrontend ist für die Visualisierung der Daten verantwortlich. Da bei derzeitigen IT Infrastrukturen schnell große Datenmengen anfallen, ist für die Abfrage der Daten durch das

Frontend zusätzlich eine Suchindex-Anwendung, wie bspw. Apache Solr, vorgesehen, welche zu Zwecken der Skalierung - ähnlich zum zu verwendenden DBMS - geclustert werden kann. Ein Update-Wrapper-Script sorgt dabei für die notwendige Reindexierung, sobald neue Daten verfügbar sind, um so ein Echtzeitlagebild im Frontend gewährleisten zu können.
Zur Visualisierung des Lagebildes im Frontend ist zum einen die Einbindung einer Karten-API, wie z.B. OpenStreetview, vorgesehen. Dies kann sowohl komplett on-premise mittels eigenem Tile-Server oder auch via Web-API erfolgen. Des Weiteren kann die interne Infrastruktur über eine Network Map visualisiert werden.
Zusätzlich zur Lagebilddarstellung mittels Kartenanzeige, ist eine Datentabelle eingebettet. Diese ermöglicht Business Warehouse Funktionalitäten, wie Drill down, Roll up und Slice, zur Filterung und Anzeige unterschiedlicher Detailierungsgrade der Daten zur Adressierung der unterschiedlichen Hierarchiebenen. Die jeweiligen Anzeigemöglichkeiten können über ein Rollenkonzept mandantenfähig in der Frontendanwendung umgesetzt werden.

Die BI Plane stellt die Entscheidungsunterstützungsfunktionalität bereit. Gemäß dem ETL-Prozess (Extract, Transfer, Load) werden die Daten aus der Data Plane für die BI Plane extrahiert und aufbereitet. Zusätzlich können Informationen Dritter und Best Practices über diesen Prozess eingebunden werden. Ein neuronales Netz verarbeitet die geladenen Daten und Best Practices und generiert bei stattgefundenen Vorfällen mögliche Handlungsweisen passend zum Vorfall. Durch Feedback der Nutzer über den Erfolg der Vorschläge lernt das System zudem diese neu einzuordnen und zu bewerten. Dies führt zu einer stetigen Anpassung an die vorhandene Infrastruktur und Handlungsweise der Nutzer.

Der beschriebene Ansatz kann dabei in verschiedenen Anwendungsfällen eingesetzt werden. Aus technischer Sicht erleichtert die Anwendung das Troubleshooting durch die visuelle Darstellung der Verbindungen und A/B-Flows im Netz und den Business Warehouse Funktionalitäten.
Dies lässt sich ebenfalls auf das Szenario Incident Response übertragen. Zusätzlich werden hierfür die unterschiedlichen involvierten Hierarchiebenen und Bereiche unterstützt durch die Visualisierung in differenzierten Detailierungsgraden.
Das Szenario Entscheidungsfindung wird durch die BI Plane und deren rekursiven KI-Ansatz behandelt.
Zudem wird eine individuelle operationale Lagebilddarstellung in Echtzeit ermöglicht.
Zur forensischen Aufklärung ist eine Erweiterung der Architekur durch eine Forensik Plane möglich, die die Erfassung der Daten im Sinne der Integrität und Beweisbarkeit absichert. Somit kann auch der Anwendungsfall der forensischen Aufklärung eines Sicherheitsvorfalls adressiert werden.

Zur Umsetzung der vorgestellten Architektur werden lediglich entsprechende, leistungsgerechte Hardware, sowie gegebenenfalls notwendige Lizenzen (bspw. für IDS) benötigt.

In erster Linie sind die notwendigen Akteure im jeweiligen Unternehmens- oder Behördenumfeld zu finden. Zur Verbesserung der Leistung des Entscheidungsunterstützungssystems ist die Kooperation mit weiteren Unternehmen und/oder Behörden von Vorteil.

PoCyMa

KI based traffic analysis and decision support tool

Team localos

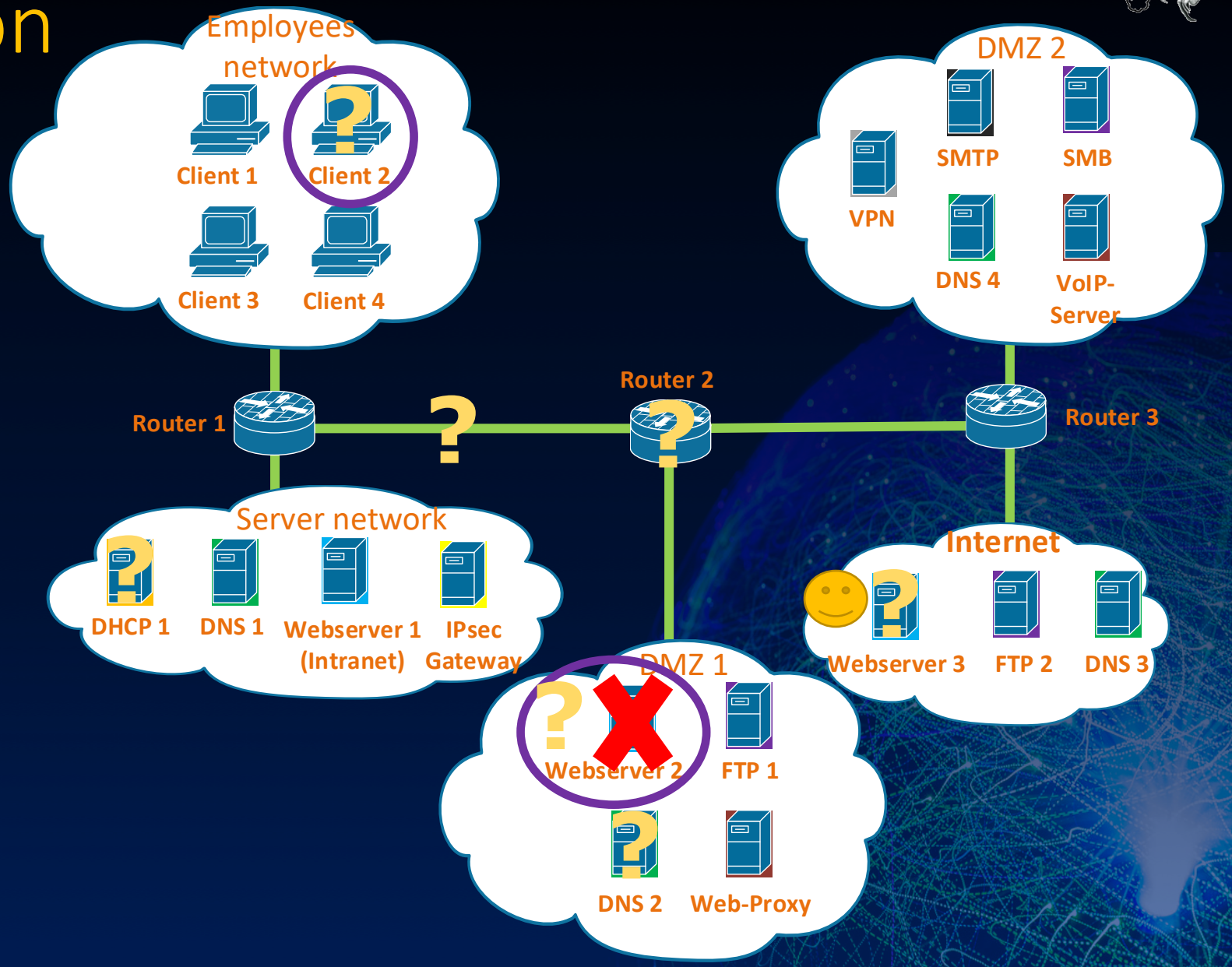
Presenter: Dr. Peter Hillmann

pocyma@mailbox.org

https://github.com/localos/PoCyMa

10.07.2019

Motivation


Employees network
Client 1
Client 2
Client 3
Client 4
DMZ 2
VPN
SMTP
SMB
DNS 4
VoIP-Server
Router 1
Router 2
Router 3
Server network
DHCP 1
DNS 1
Webserver 1 (Intranet)
IPsec Gateway
Internet
Webserver 3
FTP 2
DNS 3
DMZ 1
Webserver 2
FTP 1
DNS 2
Web-Proxy

Problem / Gap analysis

- Top-Down: *Different views and information at different hierarchy levels*

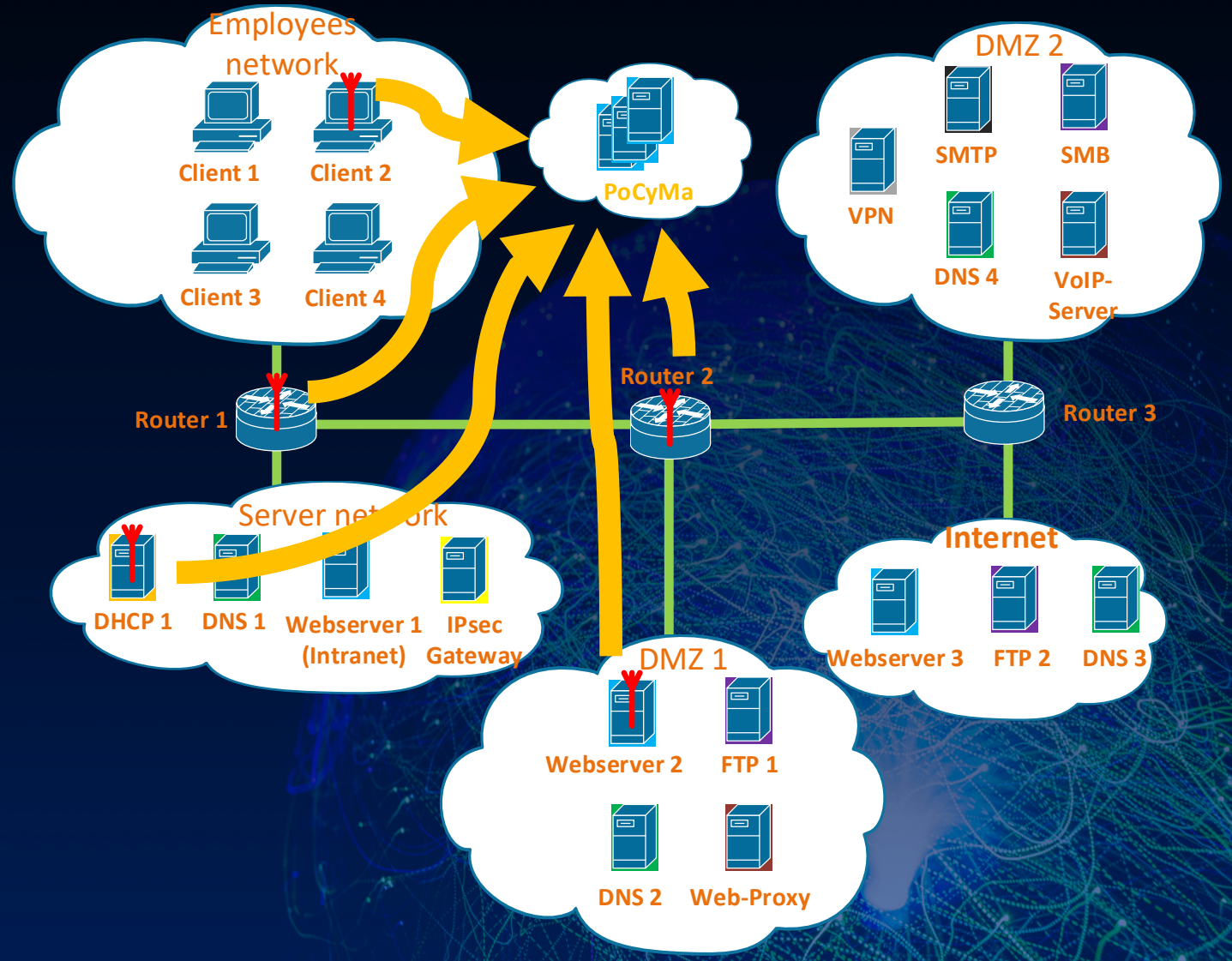
Human mind / Knowledge / Experience / Collaboration
→ Missing IT Alignment

- Bottom-Up: *Technical solution activities*

→ Improve the recovery and collaboration with tailored solutions

Solution Approach - PoCyMa

- Decision support for multiple hierarchy levels
- Holistic and integrated approach
- Preparation
  - Deploy multiple sensor
  - Usage of data from other monitoring systems



Solution Architecture

Control Plane

Configuration/ Administration

Controller

Information Overview

Data Plane

Data Analyzer (Monitoring)

Network

Host

PCAP pro Flow

Syslog

Tranalyzer T2

IDS/IPS/ SIEM

Meta Data Host/Service/ Documentation

DB1

DB2

Big Data Correlation *Apache Solr*

GeoDB

VulDB

IP Geolocation

Vulnerability Repository

OpenStreetview Server

Website

Network Infrastructure

JSON

Anonymizer

Collaboration

Business Intelligence Plane

Extractor

Aggregator

Normalizer

DB 3

Training & Test

Solution Repository

Artificial Intelligence

Vision / Added value

- Decision support
  - Incident Response
  - Prioritized kind of incident
  - Best Practices / step-by-step guides
- Feedback loop
- Improved problem recovery
  - → Resilience
  - → Decision level collaboration
  - → Automated Healing