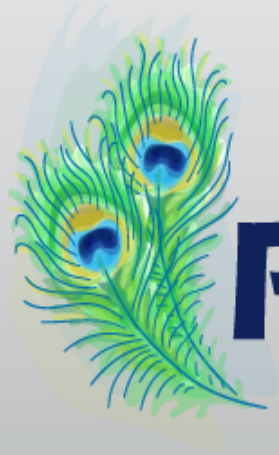


PANOPTESEC

Antonella Del Pozzo
delpozzo@dis.uniroma1.it

Hacker generation

- **First generation** (70's) was inspired by the **need for knowledge**

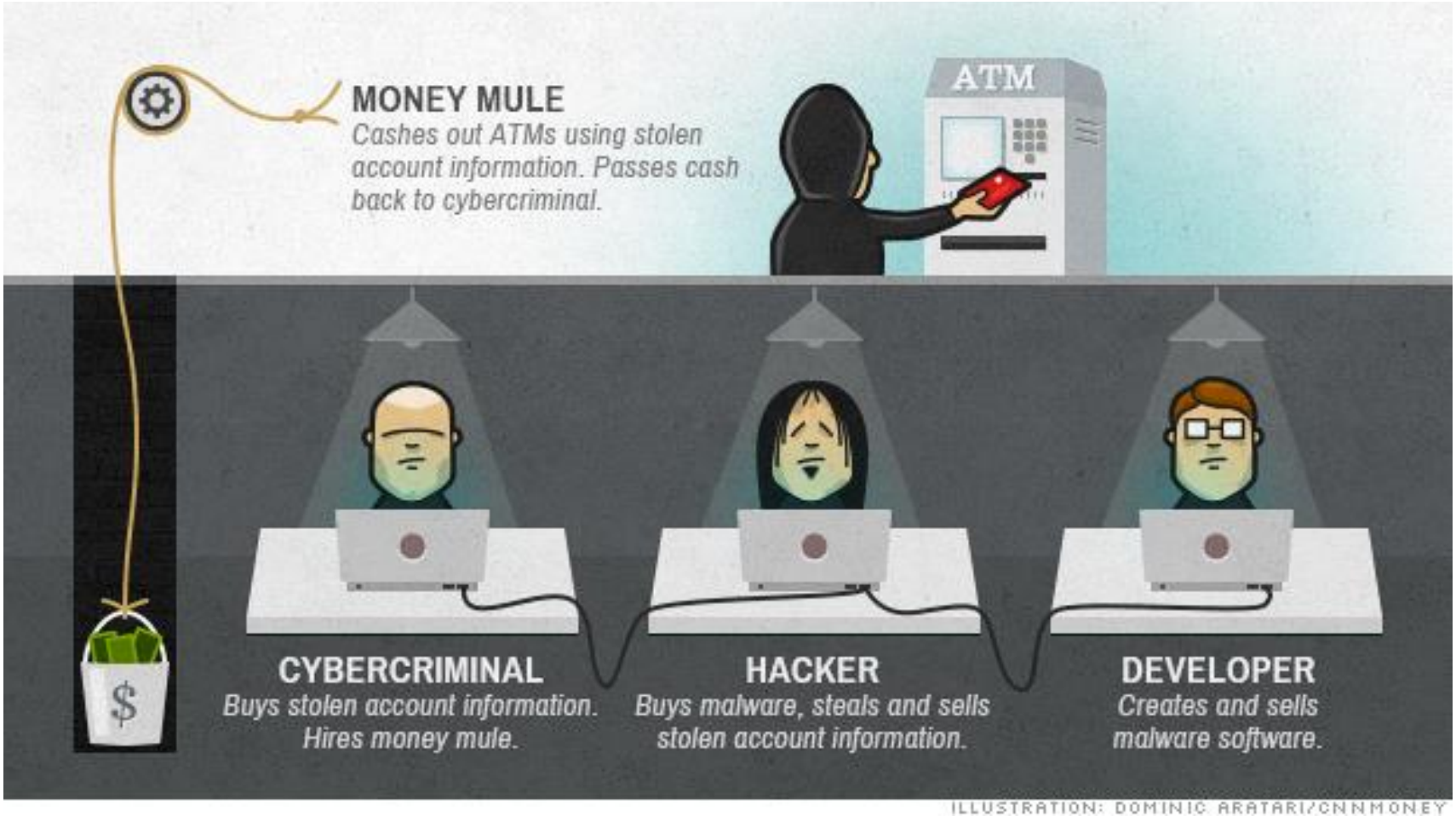
- **Second generation** (1980-1984) was driven by **curiosity** plus the knowledge starving: the only way to learn OSs was to **hack them**; later (1985-1990) hacking becomes a **trend**.

- The **Third one** (90's) was simply pushed by the **anger for hacking**, meaning a mix of **addiction**, **curiosity**, **learning new stuff**, **hacking IT systems and networks**, **exchanging info** with the **underground community**.

- **Fourth generation** (2000-today) is driven by **angerness** and **money**: often we can see subjects with a very low know-how, thinking that it's "cool & bragging" being hackers, while they **are not interested** in hacking & phreaking history, culture and ethics. Here hacking meets with politics (**cyber-hacktivism**) or with the criminal world (**cybercrime**).

PANOPTESEC

Modern bank robbers



PANOPTESEC

Critical Infrastructures

Aurora Generator Test (2007): the experiment involved hacking into a replica of an Idaho power plant's control system. It proved and demonstrated that you can destroy physical equipment with a cyber-attack.

PANOPTESEC

Risk in critical infrastructures

The **IT** systems **which have to manage control automation in industrial environments** are one of the **main sources for risk** when speaking about critical system failure.

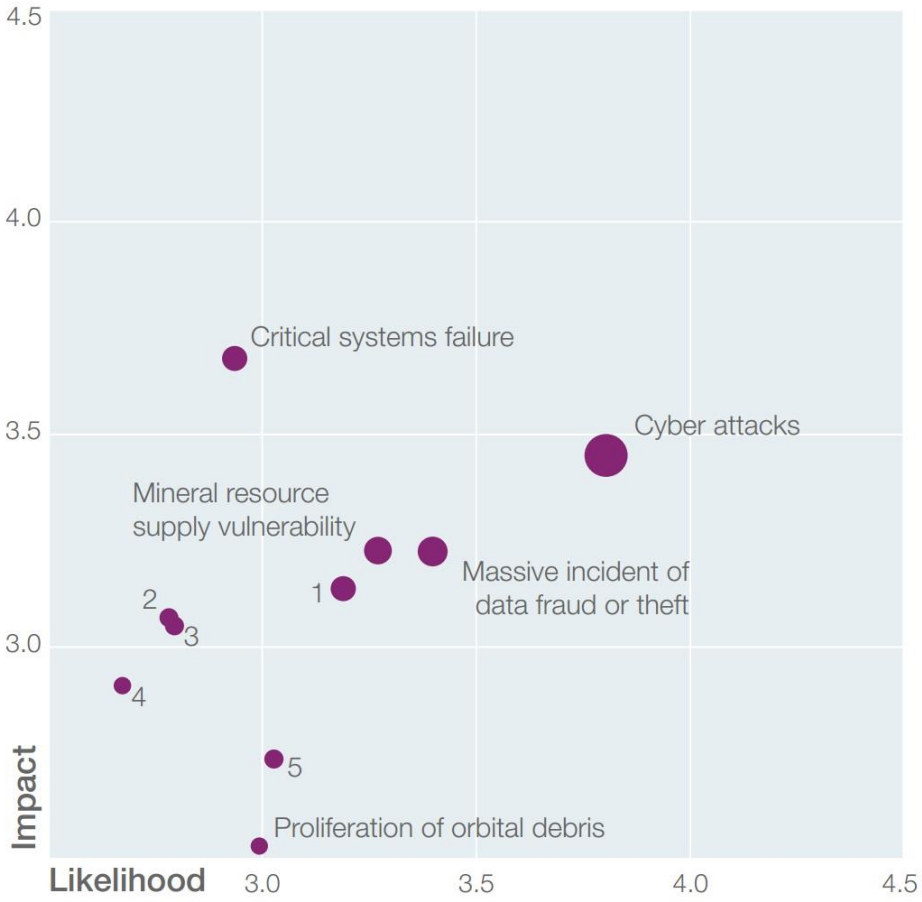
Historically they have not been designed considering IT attacks.

Word Economic Forum (WEF) http://www3.weforum.org/docs/WEF_GlobalRisks_Report_2013.pdf

PANOPTESEC

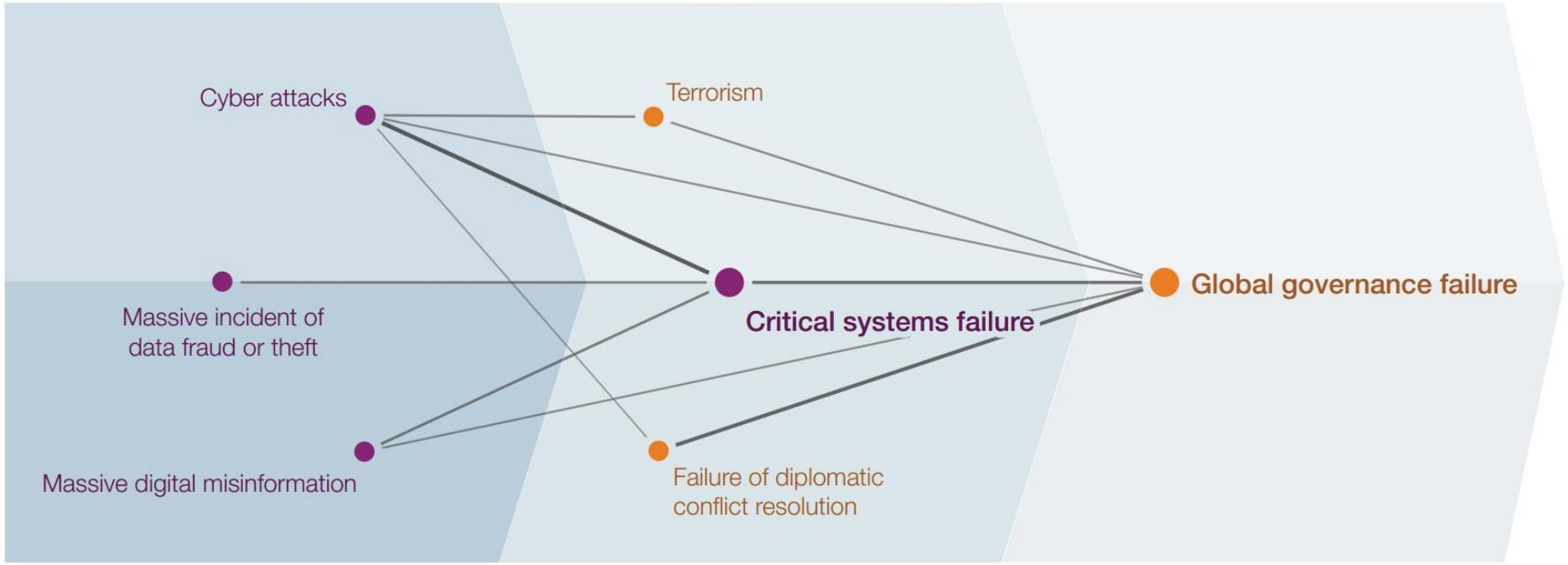
Technological Risks

Figure 38: Technological Risks



PANOPTESEC

Figure 17: The Dark Side of Connectivity Constellation



Origin Risk
Increasing capabilities for cyber crime and attacks.

Pathways
Balance-of-power tips as new actors can wage effective interference and disrupt commerce.

Manifestation
The traditional system of global governance is undermined.

Source: World Economic Forum

PANOPTESEC

Let’s start easy

How to protect a laptop?

- Proactive way:

Find its vulnerabilities and patch them to prevent their exploitation.

- Reactive way:

Detect current attacks and execute countermeasures (What antivirus does).

PANOPTESEC

A possible vulnerability definition

A weakness in system security procedures, hardware design, internal controls, etc. , which could be exploited to gain unauthorized access to […] sensitive information (or privileges).

PANOPTESEC

Patch

A Patch basically is a fix for security vulnerabilities and bugs.

Once I know that there is a vulnerability on my laptop I can patch it.

Attack detection

Intrusion Detection System (IDS)

A tool that monitors network or system activities for malicious activities or policy violations. When something wrong is detected an IDS alert is raised.

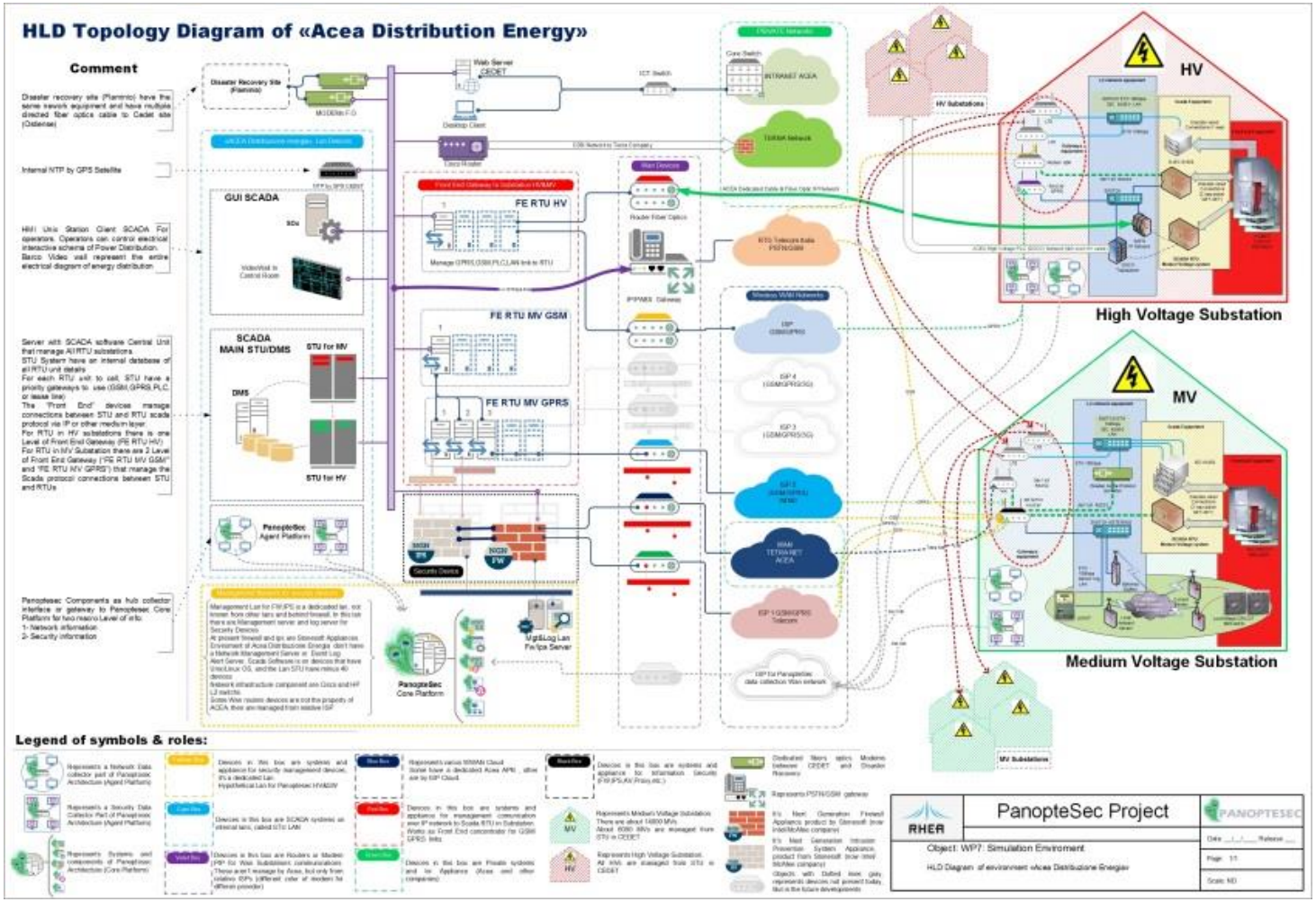
Countermeasures

A countermeasure is an action that reduces a threat, or an attack by minimizing the harm it can cause.

e.g., close ports, shut-down the server.

Once an attack has been detected then countermeasures can be instantiated.

What about an Enterprise Network?



The concept of multi-step attack arises.

At the proactive side

- Proactive side issues:
  - Lots of vulnerabilities to address, to which I should start?
    - Not all device may be physically reachable
    - It may cost in terms of time and money
    - Not all patches may be worthy to be applied
    - Vulnerabilities are connected between them, how to reason on that?

At the reactive side

- Reactive side issues:
  - What can I understand with IDS alerts on different devices?
  - What kind of countermeasures should I take (e.g., where should I cut my network )?
  - How to take into account enterprise assets?

PANOPTESEC

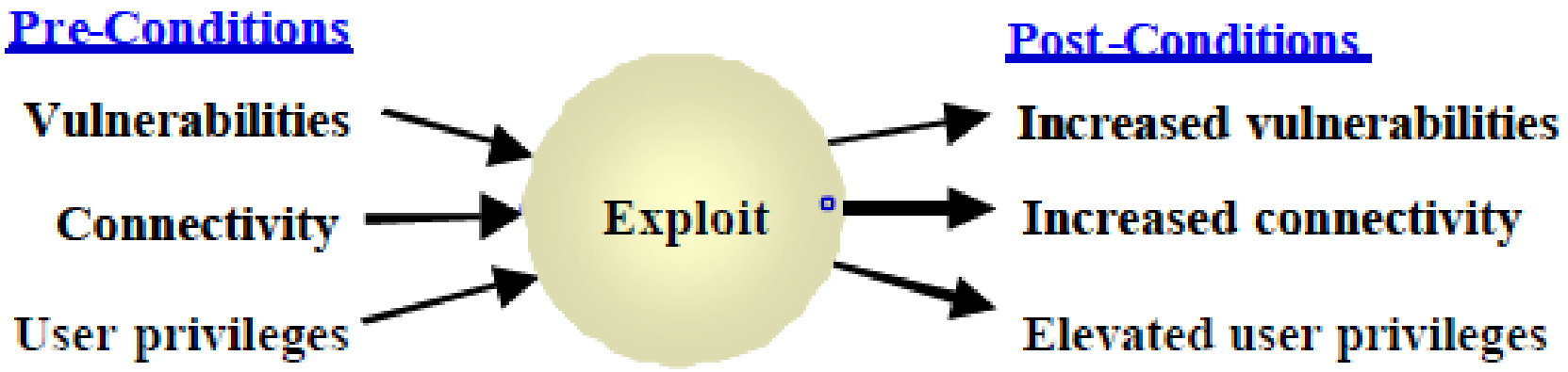
How to reason on all these aspects

A formalization is required to connect and reason on all these information:

- Vulnerabilities
- Network connectivity
- IDS alerts

PANOPTESEC

Pre-condition and Post-condition



[1]

Exploit Post-Conditions are Pre-Conditions of another one, and so on.

Multi-step attack

[1] Ritchey, Ronald, Brian O'Berry, and Steven Noel. "Representing TCP/IP connectivity for topological analysis of network security." *Computer Security Applications Conference, 2002. Proceedings. 18th Annual*. IEEE, 2002.

PANOPTESEC

Vulnerabilities Inventory

A structured representation of the vulnerabilities that could be exploited.

Vulnerabilities characteristic should be expressed in a public scoring system (e.g., Common Vulnerability Scoring System (CVSS). It is an open and standardized method for rating IT vulnerabilities).

Network Connectivity

Reachability Matrix:

- on any column and row a device IP address.
- each cell carry the information of the IP ports and protocols authorized between the couple of IP addresses, in one direction and the other.

Summing up

- We have an inventory with all the known vulnerabilities in our network
- We know that if an attacker exploits a specific vulnerability on node A can reach node B.
- We have the Reachability Matrix

How to mix all together up?

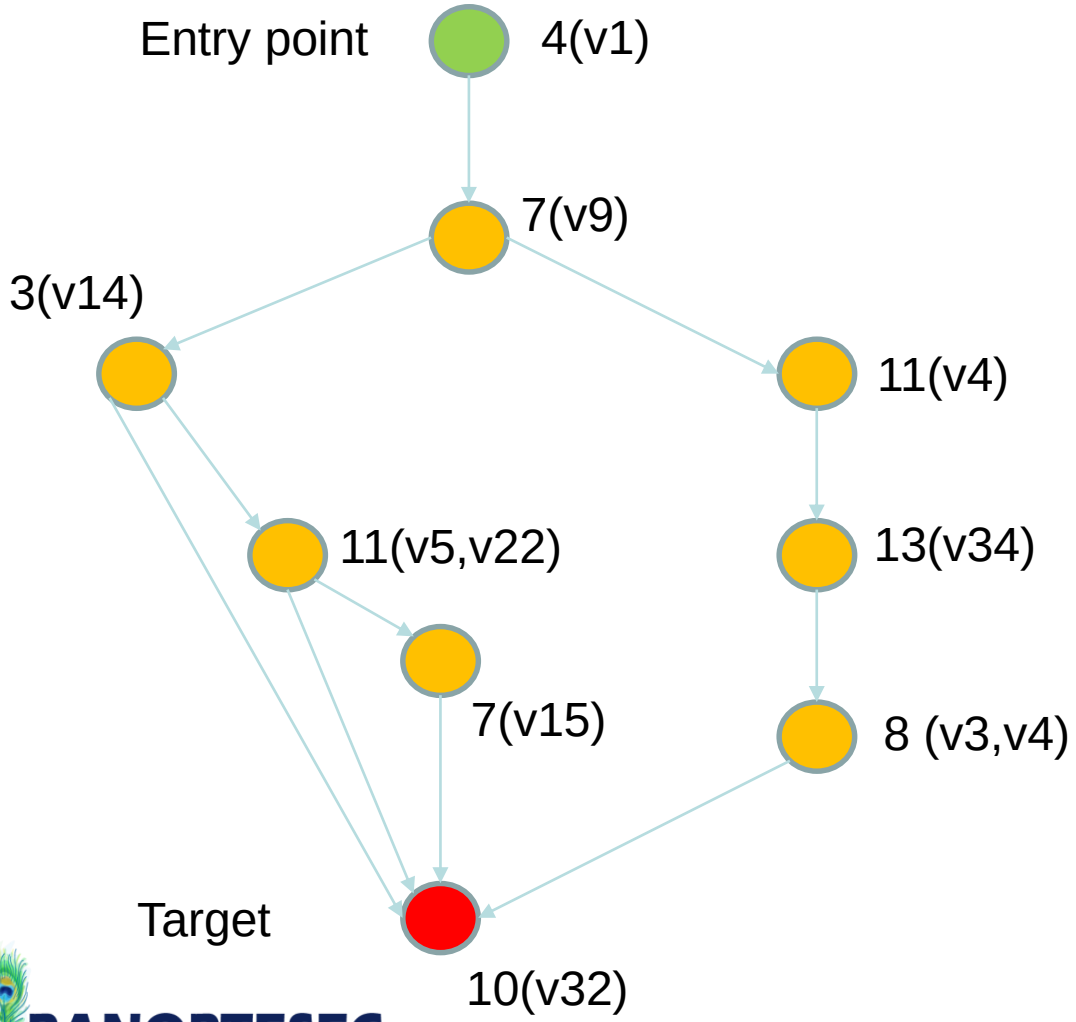
Attack Graph (AG)

There is not an unique syntax.

A possible AG representation is:

- nodes: devices and vulnerabilities
- edges: relationships as pre-condition and post-condition.

Example of attack graph



Nodes represent attack actions (identified by device identificator and vulnerabilities)

Edges represent that a dependency link or a causality relationship exists between attack actions.

PANOPTESEC

How to use AG

An attack graph gives us the possible paths (as far as we know) that can be used to reach a target device.

In a proactive way it is useful to understand where to apply patches.

In a reactive way it is useful to try to understand, given IDS alerts, where an attacker is going to.

Missions of the Organization

Any countermeasure, both in the proactive and reactive way, must take into account the effectiveness and business assets (missions).

e.g., the shut-down of a highly vulnerable server might reduce the risk of it being exploited, however as the mission depends on this server, it highly affects the mission.

PANOPTESEC

Panoptesec

The PANOPTESEC consortium will deliver a beyond-state-of-the-art prototype of a cyber defence decision support system, demonstrating a risk based approach to automated cyber defence that accounts for the dynamic nature of information and communications technologies (ICT) and the constantly evolving capabilities of cyber attackers.

PANOPTESEC

PANOPTESEC Project

SME - Experienced technology company
- Corporate experience in both FP programmes and cyber defence

Large telecom equipment provider
- Security research dept. Bell Labs France focus on cyber security and cloud security

SME – Knowledge based systems for industrial applications
- Specialized in description logics and semantic web

Sapienza Research Centre on Cyber Intelligence and Information Security
- Research focus on distributed systems and information visualization

Institute for Software Systems and Institute of Security in Distributed Systems
- Research focus on knowledge based systems and description logics

Industrial development of ICT through the science and technology
- Distributed Services, Architectures, Modelling, Validation and Network Administration

Important Italian public utility which focuses on energy and water
- Operational environment for experimentation and demonstration

The “CIDre” team (*Confidentialité, Integrité, Disponibilité, et repartition*)
- Research focus on intrusion detection, correlation and visualization

PANOPTESEC

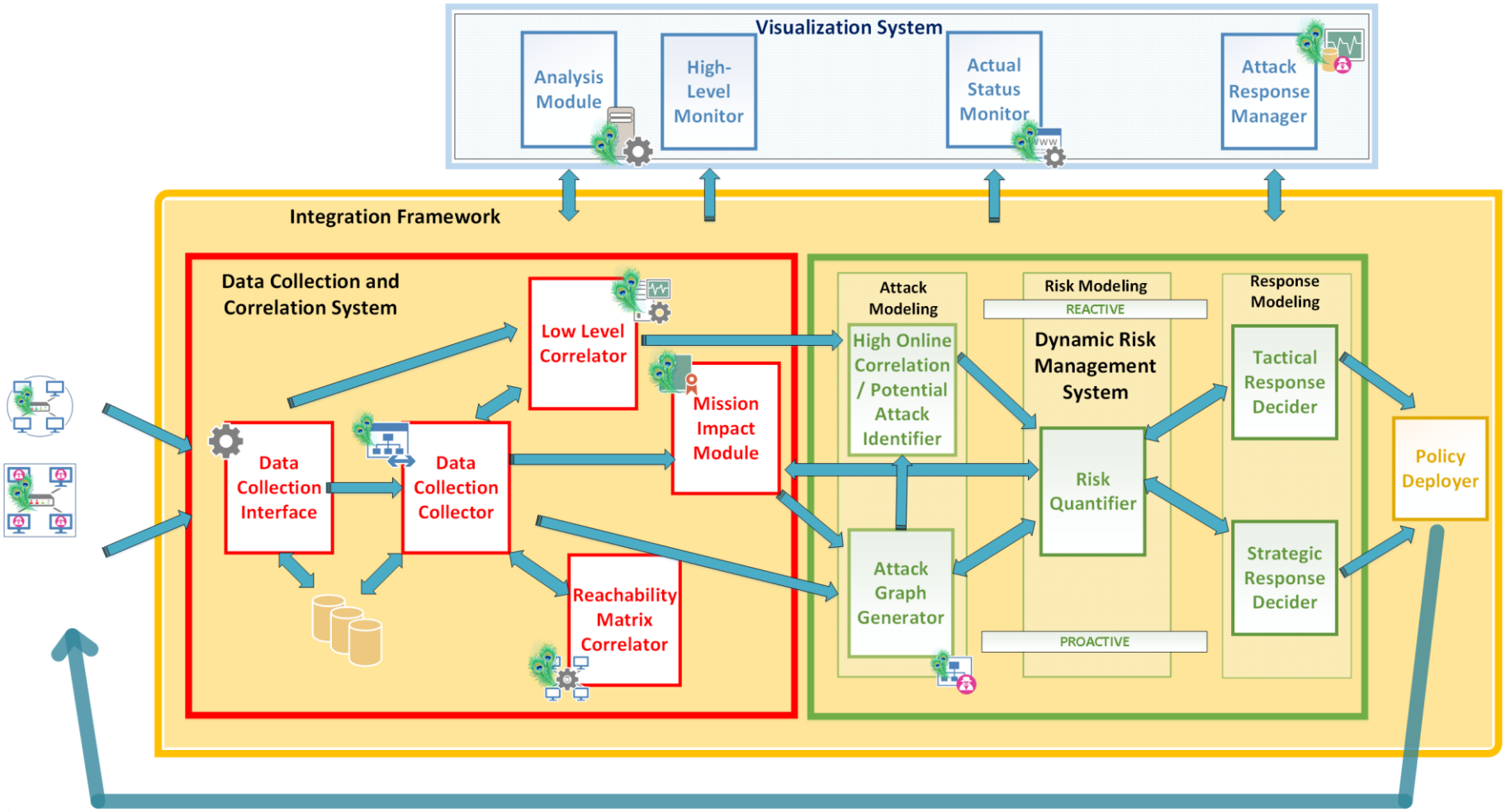
PANOPTESEC Project

SME - Experienced technology company

- Corporate experience in both FP programmes and cyber defence

Special guest from Rhea: Ing. Matteo Merialdo

Panoptesec Architecture

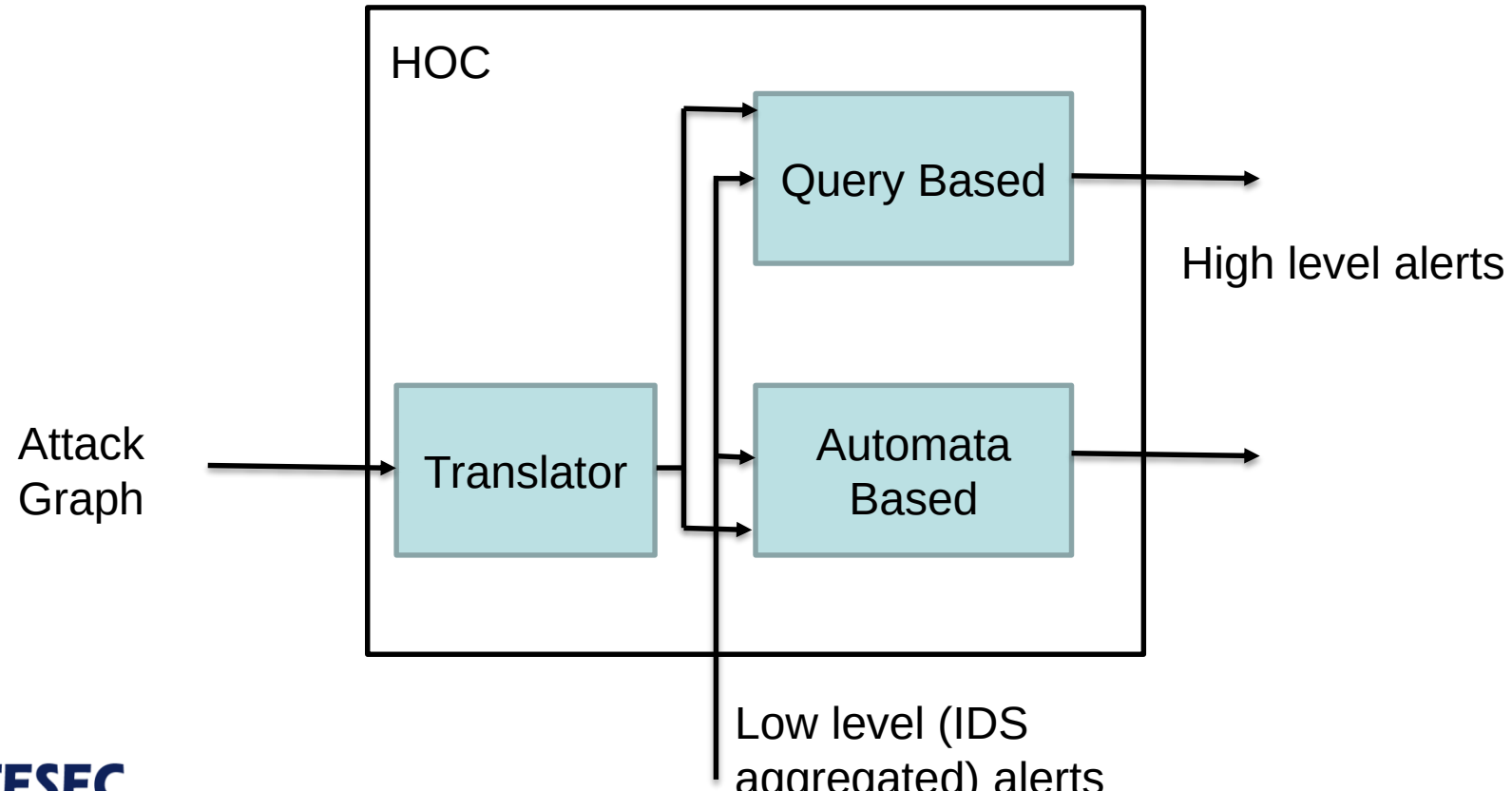


PANOPTESEC

High Level On-Line Correlation (HOC)

Two orthogonal approaches:

- Automata Based (AB)
- Query Based (QB)



PANOPTESEC

High level view

The main purposes of the module are:

1. Understand on whose attack paths the received low level alert are located.
2. Identify the attack path currently under attack
3. adapt itself automatically to any new attack graph.

PANOPTESEC

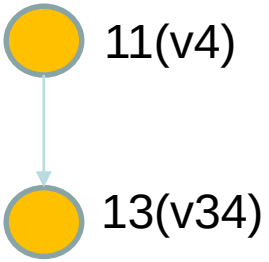
Matching low-level alerts on attack paths

Any low level alert carries the following information:

- Source node
- destination node
- exploited vulnerability

It matches, if any, an edge of the attack path.

From 11, to 13, exploiting v34.



PANOPTESEC

What if there is no match?

What does it means that a low level alert does not match any attack path?

- The alert could be an IDS false positive

- The exploited vulnerability does not belong to our knowledge base. It is a *zero day*!

Stream of low level alerts

Low level alerts are provided to the module as a stream of event. These have to be analysed in real time. Given their number, they can not be stored and then analysed.

A Complex Event Processing (CEP) engine is employed. Esper in that case.

http://www.espertech.com/esper/index.php

PANOPTESEC

ESPER

Open Source CEP, available both for Java and .NET.

Esper enables rapid development of applications that process large volumes of incoming messages or events.

Esper works as real time engine that triggers actions when event conditions occur among event streams.

Event Processing Language

Esper comes with an “SQL-Like” language, called Event Processing Language (EPL). Very similar to SQL: select, where, having, group by and so on…

EPL is used to specify event conditions that have to occur in the event stream in order to trigger a specific action.

These event conditions are nothing more than queries used to filter our event stream.

PANOPTESEC

Example

Very basic example:

select * from pattern [every (A or B).win:time(100 sec)]

Means: I want to be alerted when an A or a B event arrives in the last 100 seconds

Attack Path as queries

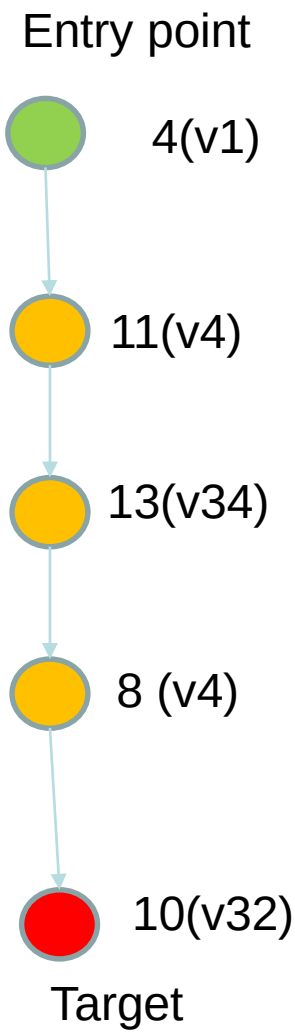
In our case we do not need for complex queries. Given a stream of alert we want to know to whose attack paths do they match.

As we told, an alert match, if any, to an edge in the attack path.

- We need for a query for any attack path.
- Any query is an OR of all the edges compounding the attack path.

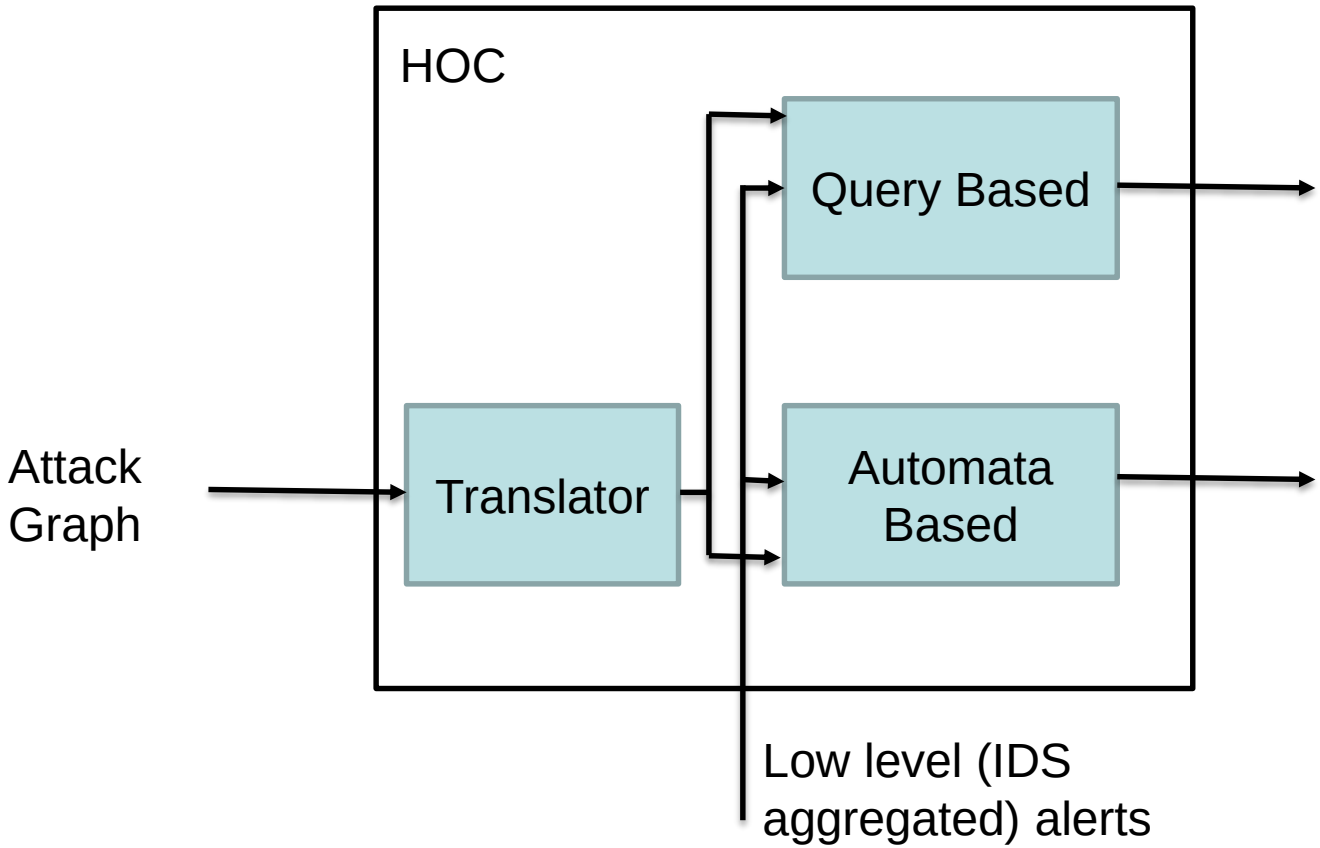
Basic example:

```
select *
from EventStream
where [
        (s=4 AND d=11 AND v=v4) OR
        (s=11 AND d=13 AND v=v34) OR
        (s=13 AND d=8 AND v=v4) OR
        (s=8 AND d=10 AND v=v32)
]
```



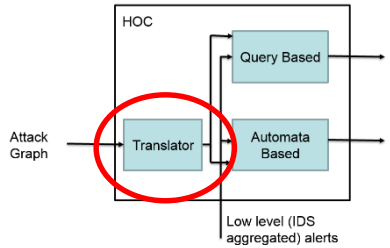
PANOPTESEC

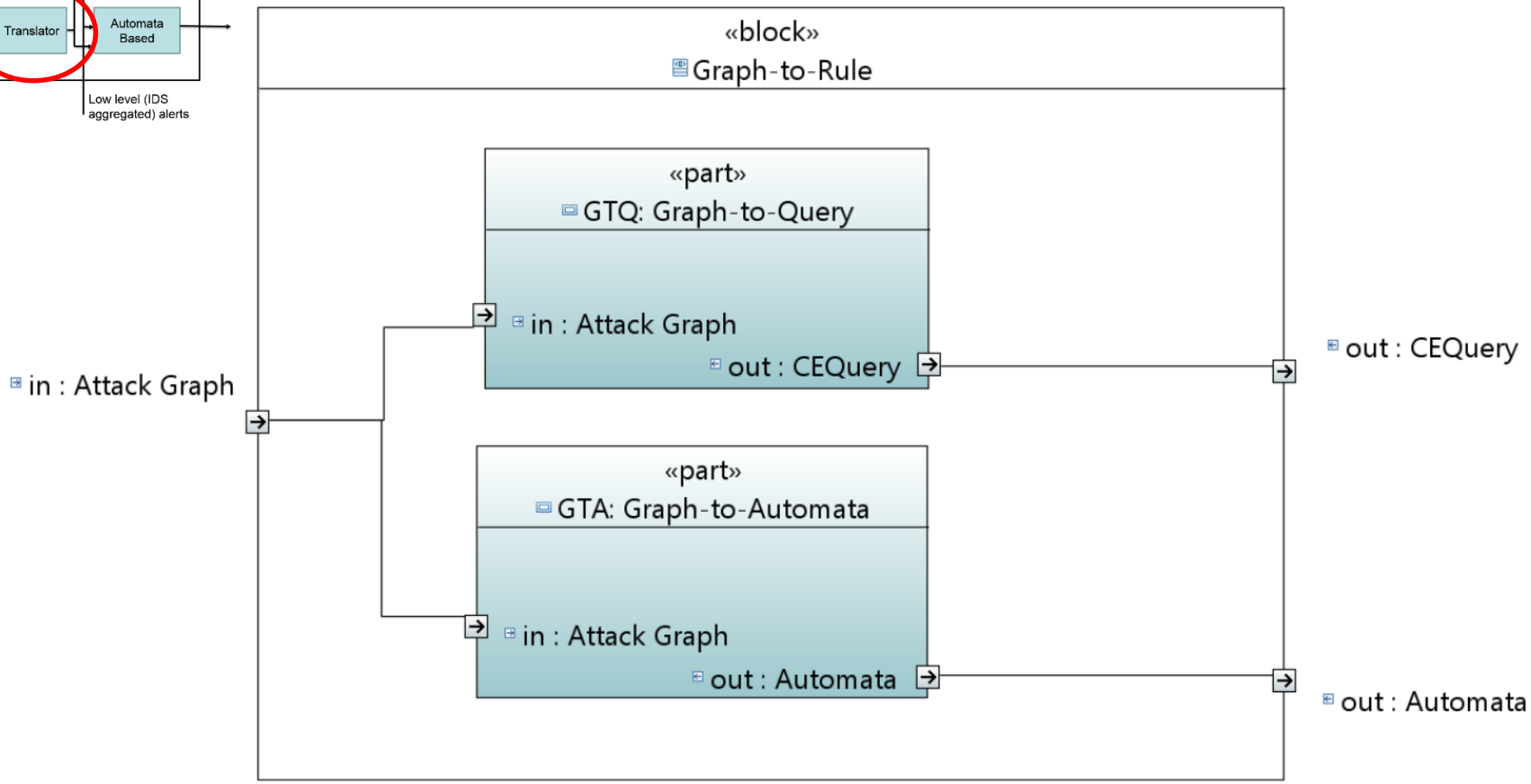
Back to HOC



PANOPTESEC

Translator


HOC
Query Based
Attack Graph
Translator
Automata Based
Low level (IDS aggregated) alerts


«block»
Graph-to-Rule
«part»
GTQ: Graph-to-Query
in : Attack Graph
out : CEQuery
«part»
GTA: Graph-to-Automata
in : Attack Graph
out : Automata
in : Attack Graph
out : CEQuery
out : Automata

Graph to Rule (GTR)

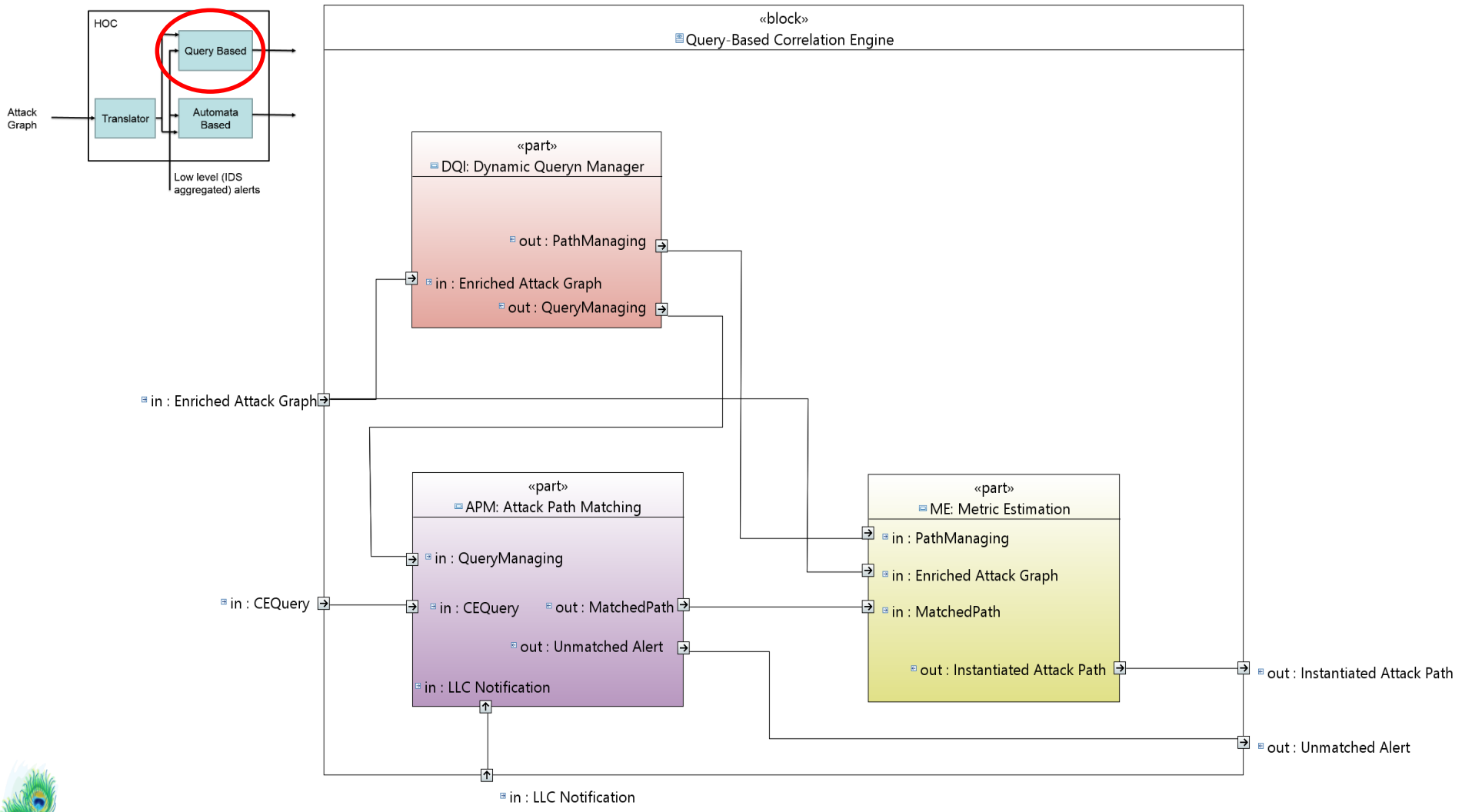
GTR:

- takes the Attack Graph as input
- subdivides it in attack paths (any for each couple of entry point and target)
- translate attack paths in formats suitable for the two subsequent modules (queries for QB module).

Informally speaking it is a parser.

PANOPTESEC

Query Based



PANOPTESEC

Dynamic Query Manager

It is in charge to update all the module data structures when a new attack graph is provided.

This can be done following two approaches:

- epoch based: the new attack graph substitute the previous one.

- incremental: data structures are updated with the differences between the previous and the new one.

PANOPTESEC

Attack Path Matching

Takes in input queries representing attack path and the low level alerts stream.

Checks, using ESPER, if such alert matches some attack path. If yes it returns, if any, both alert and the attack paths matching it.

Metric Estimator

Takes as input attack paths along with the alerts matching them.

Computes the similarity between any attack path (considering its exploited edges) and the correspondent whole exploited attack path.

If the computed metric overcomes a threshold then it rises an high level alert.

Metrics

Different metrics are employed.

There is not a right one. The idea is to test them and to find which is the more suitable in which situations.

Jaccard Coefficient

Used to compare the similarity between finite sample sets.

Defined as the size of the intersection over the size of the union of the sample sets.

$$J(A,B) = \frac{|A \cap B|}{|A \cup B|}$$

A: the set of exploited edges
B: the set of the whole edges.

PANOPTESEC

Cosine Similarity

It is a kind of metric used to compute similarity between texts. Any text is represented as vector in the geometric space

$$\cos(\theta) = \frac{A \cdot B}{\left||A|\right| \cdot \left||B|\right|}$$

A: the vector representing the exploited edges
B: the vector representing the whole edges.

PANOPTESEC

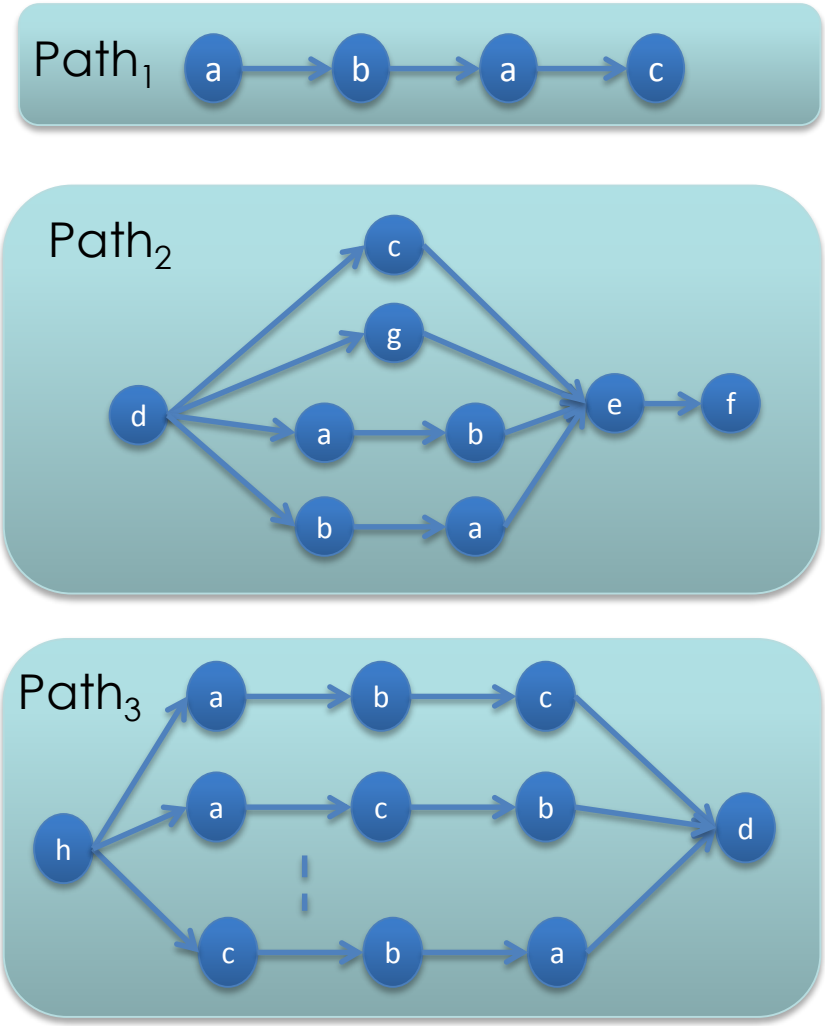
Attack Path Matching

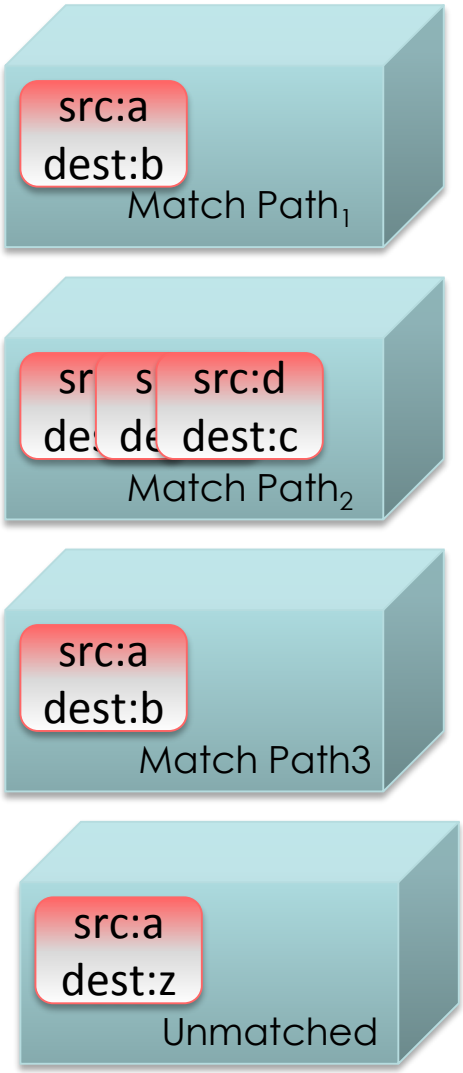
src:d
dest:c
...

src:a
dest:b
...

src:a
dest:z
...

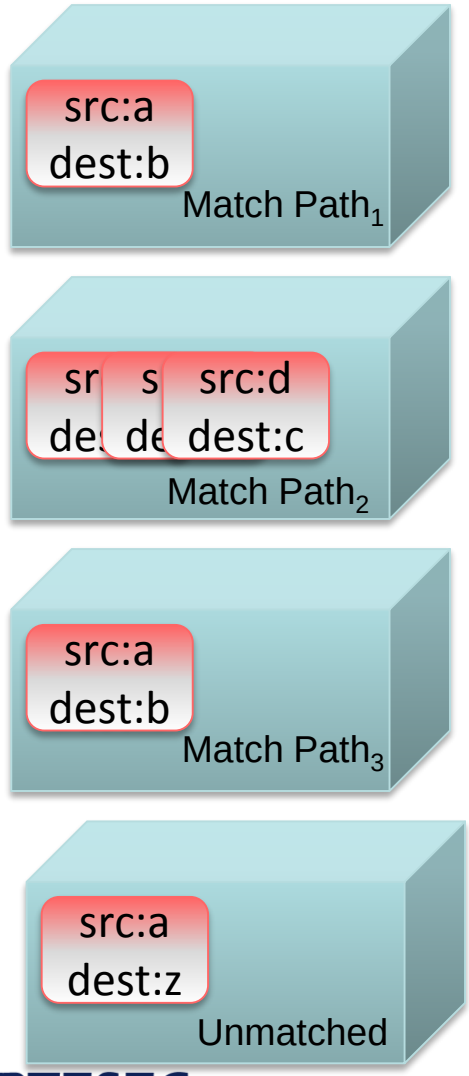
src:c
dest:e
...





PANOPTESEC

Metrics Estimation



Path id	T	Metric Value
1	0,5	0,25
2	0,6	0,66
3	0,8	0,25

PANOPTESEC

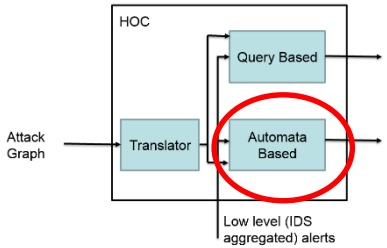
Next Step

- Definition of a new metric considering
  - Distance to the target
  - Length of the path
  - Density of alerts

Possible topics for master thesis!

PANOPTESEC

Automata Based – overview

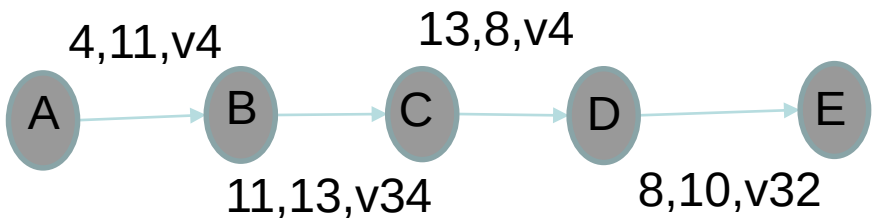


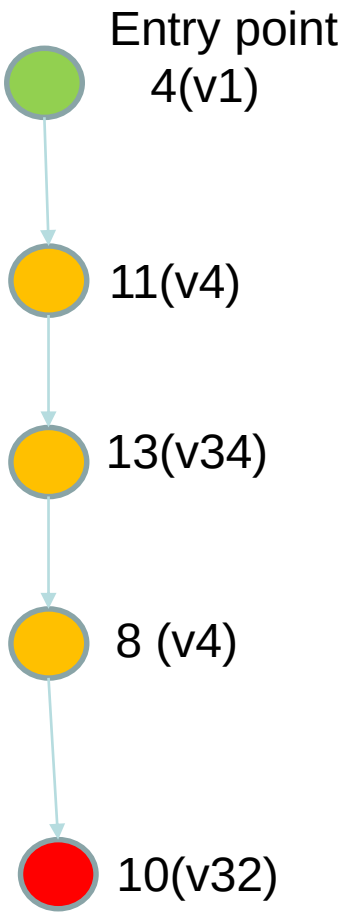
Each attack path is translated in a finite state automata.

The automata state changes any time the expected low level alert is detected.

Example

Finite State Automata





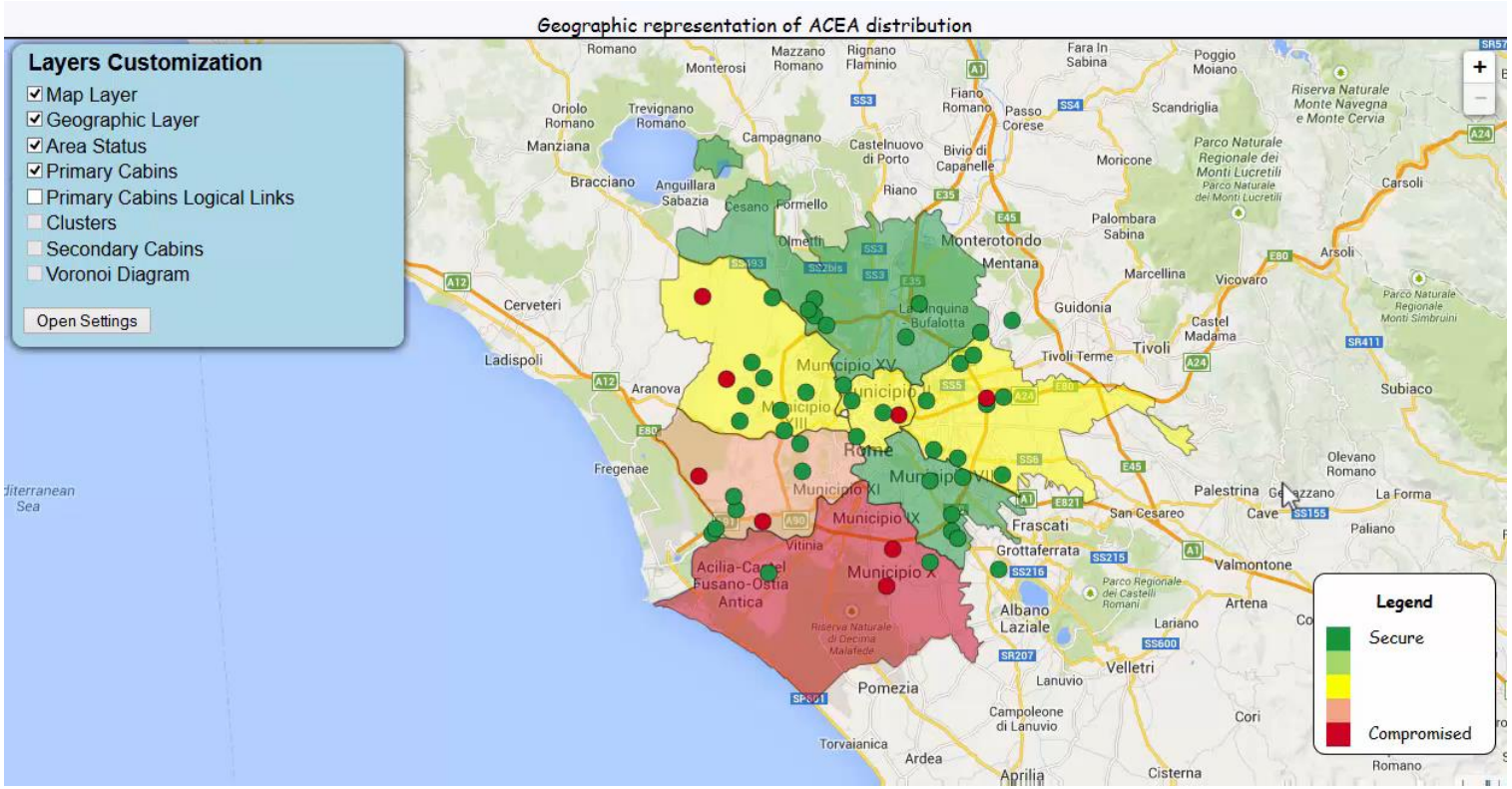
PANOPTESEC

Differences between the two approaches

- What is the big difference between the two approaches?
  - What if two subsequent alerts are delivered in incorrect order?
  - What if there are IDS false negative?
- Which one could produce more false positive?
  - Too early high level alert.
- Which one should be faster?

PANOPTESEC

Visualization



Kindly provided by Dott. Marco Angelini, visualization group at DIAG.

Q&A