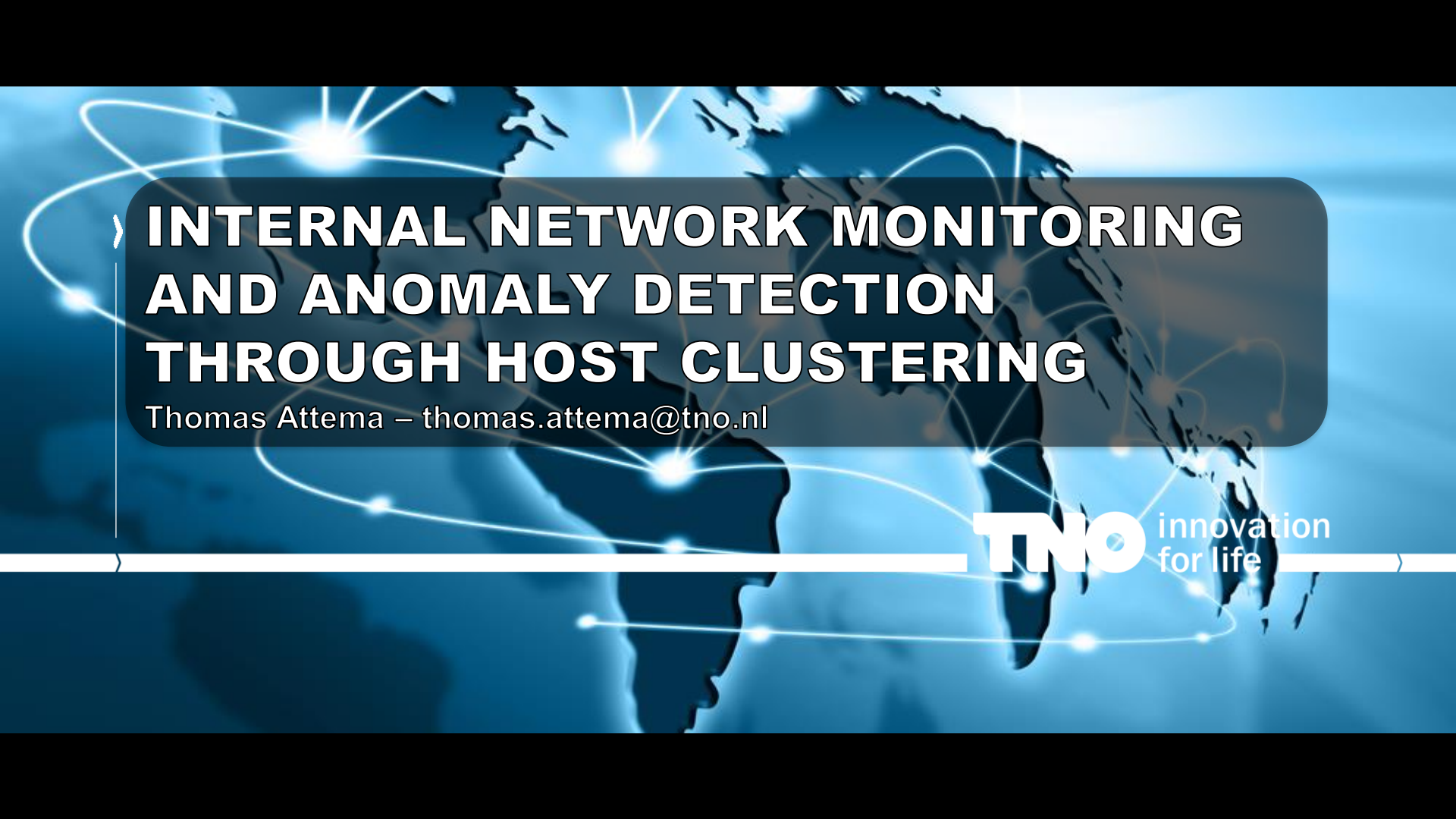

INTERNAL NETWORK MONITORING AND ANOMALY DETECTION THROUGH HOST CLUSTERING
Thomas Attema – thomas.attema@tno.nl
TNO innovation for life

SHARED RESEARCH PROGRAM – CYBER SECURITY
ROADMAP MONITORING AND DETECTION

- **Shared data**
  - Partners provide (anonymized) data to evaluate developed techniques
- **Shared working**
  - Project teams are staffed all partners
- **Shared results**
  - The project results are shared among all partners

ABN·AMRO

ING BANK

Rabobank

achmea

TARGETED ATTACKS

- **Targeted**
  - Small number of victims
  - Making use of very specific knowledge and vulnerabilities of the target
- **Persistent**
  - Patient in gathering information
  - Great effort in staying undetected
- **Highly skilled and well organized attackers**
  - Including state actors

Dell SecureWorks attack lifecycle

TARGETED ATTACKS

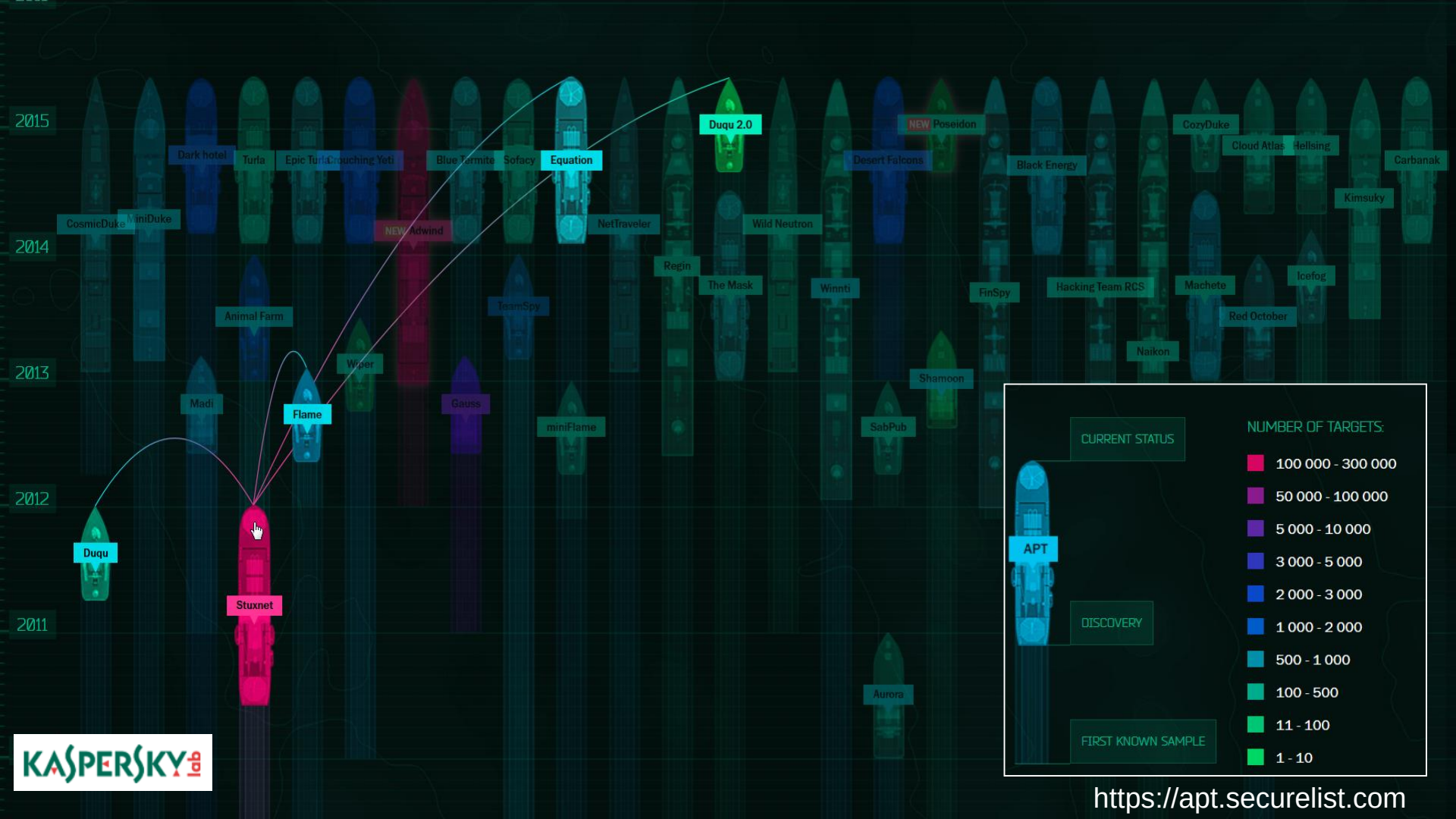
Political motives
(e.g. Stuxnet, 2010)

Financial motives
(e.g. Carbanak, 2015)

Intelligence gathering
(e.g. Duqu, 2011)

Shaming
(e.g. Sony Hack, 2014)

Terrorism
(next?)



https://apt.securelist.com

EXAMPLE: COMMAND AND CONTROL TRENDS (1/4)

A) FOCUS ON RESILIENCE - TRENDS

- Huge networks of p0wned systems
  - IOT will extend this
  - For CnC host, proxies, staging, etc.
  - Semi-volatile
- Two-stage / multi-stage CnC channels
- Attacks more targeted on *specific* organisations
  - Specific objective → Low noise
  - Organisation-specific IPs and domains → IOCs become useless
  - Spread over long periods

EXAMPLE: COMMAND AND CONTROL TRENDS (2/4)

B) HIDING THE CNC INFRA - TRENDS

- Use of steganography

- Apply secret handshake before connections
  - Investigate IP whitelists of targeted institutions

- Using common services → Legitimate sites
  - URL shorteners
  - Cloud services/social networks
  - Cloudfront
  - CDNs

- New obfuscation techniques to hide domains and IP addresses

EXAMPLE: COMMAND AND CONTROL TRENDS (3/4)

C) HIDING THE CHANNEL - TRENDS

- Encryption
  - Everything communication, storage, malware
  - Multiple layers of encryption
- Use of decoys
- Hiding in common usage patterns
  - Low frequency interaction
  - Lined up with system/network usage patterns

EXAMPLE: COMMAND AND CONTROL TRENDS (4/4)

D) HIDING THE ATTACKERS - TRENDS

- More encryption
- Attacks leave smaller forensic footprints (everything in memory)
- Detection sandboxes, virtualisation, debugging
- Falsified footprint information
  - Copying from other attacks
  - Different languages
- Multiple attacker organisations cooperate

GENERAL TARGETED ATTACK TRENDS (1/2)

- Attackers are more professional
  - Better prepared, better tools, better organised
  - More state actors
  - C2aaS, BaaS
- Lots of attacks are seemingly connected
  → Possible organised groups carrying out multiple attacks
- Increasing use of standardised toolkits for every stage
  - Tailored for each attack

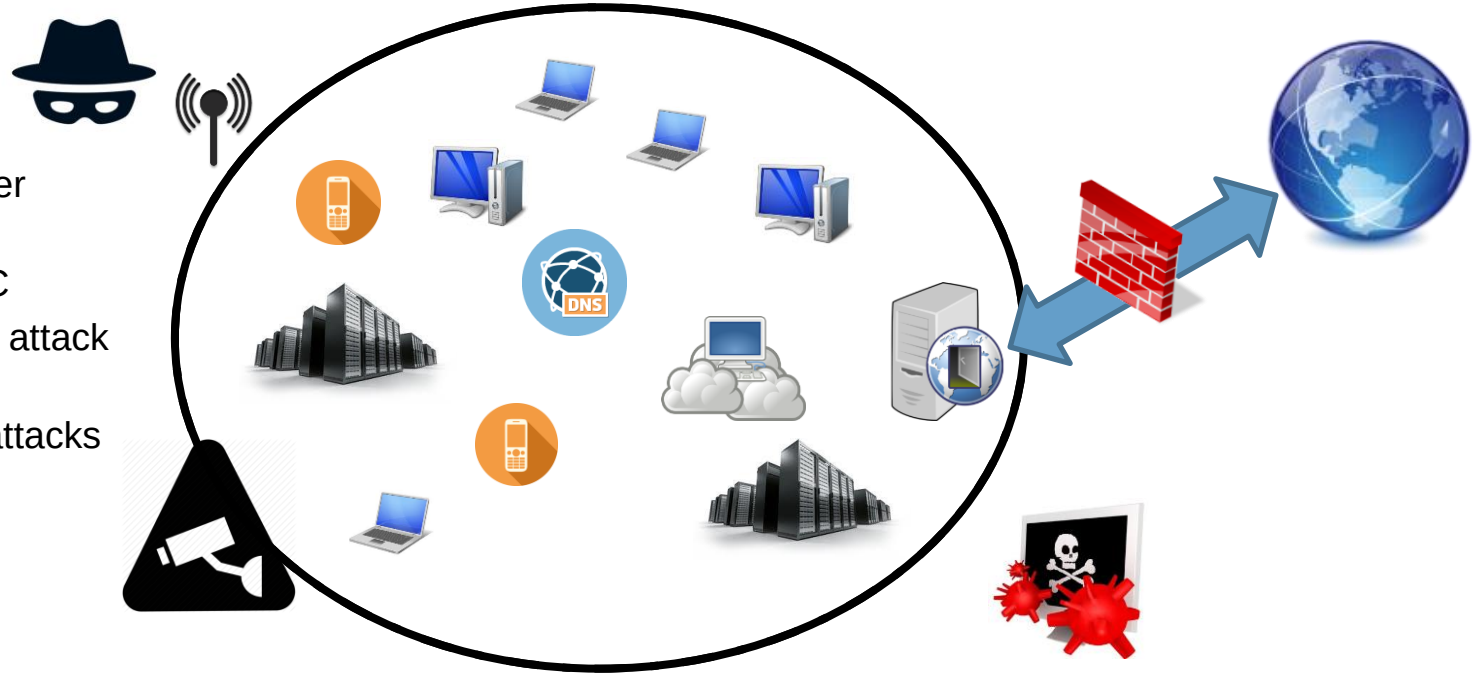
GENERAL TARGETED ATTACK TRENDS (2/2)

- Development pace increases
- Attacks, infrastructure, techniques diversify
- Attacks persist for long times → evolve gradually
- Supply chains are attacked as well
  - Pre-installed rootkits
  - Malicious hardware components

HARD OUTER LAYER & SOFT INNER LAYER

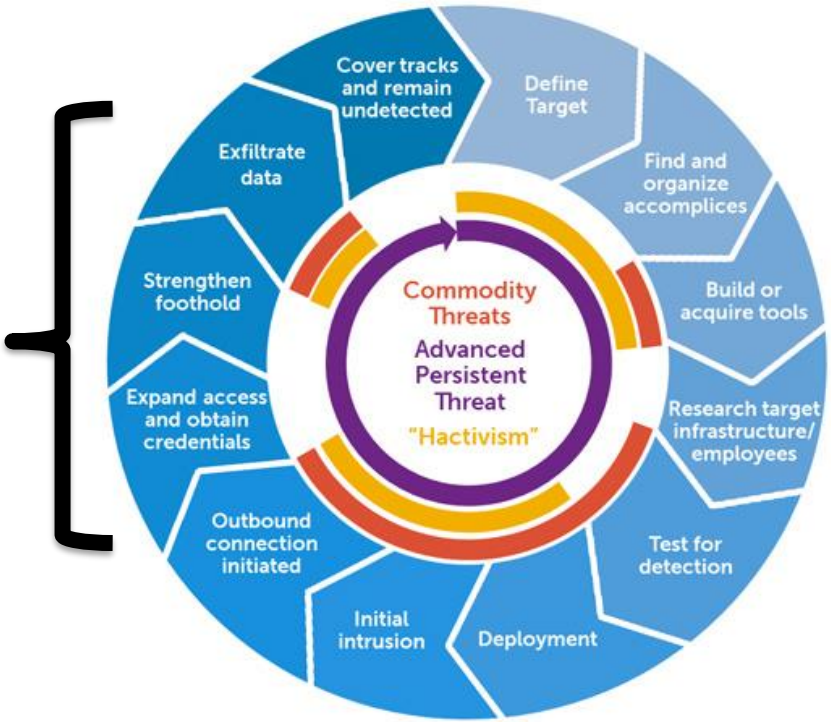
NEED FOR INTERNAL NETWORK MONITORING AND DETECTION

- Advanced cyber attacks
- Advanced CnC
- IoT: increasing attack surface
- Supply chain attacks
- ...

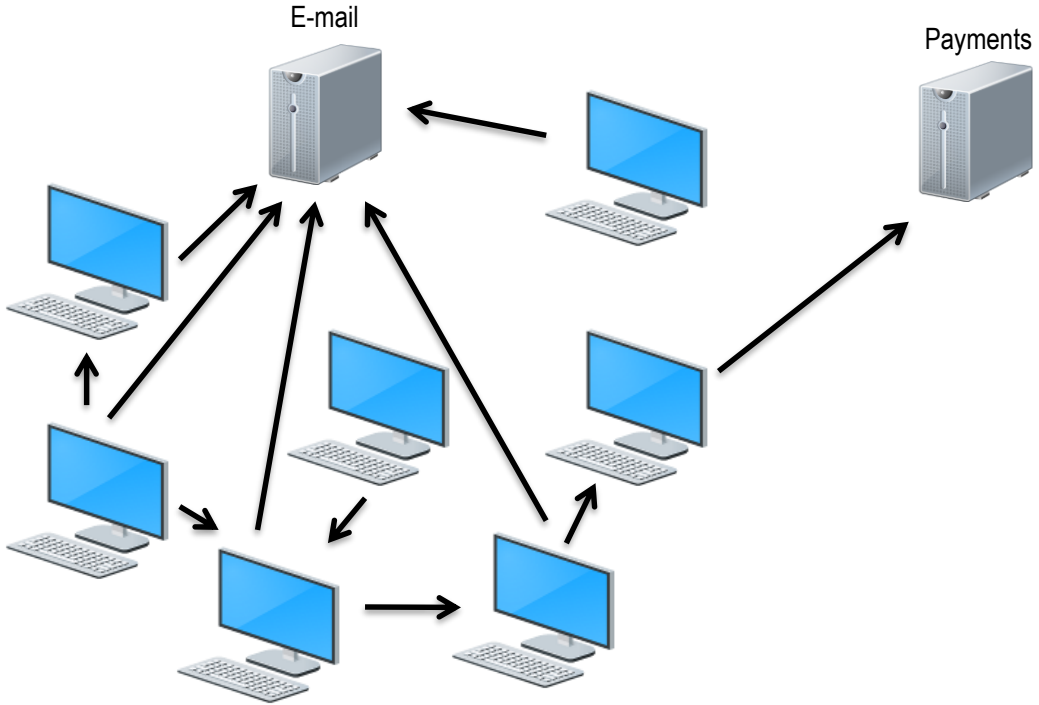


INTERNAL NETWORK TRAFFIC IS A VALUABLE SOURCE OF INFORMATION

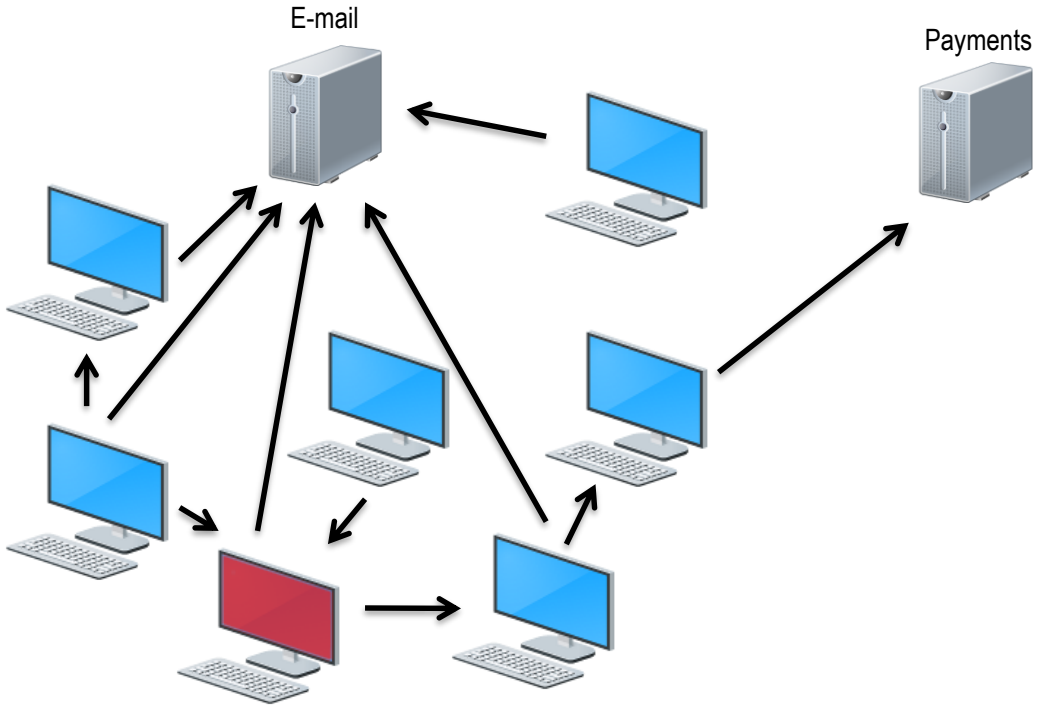
- Undervalued source of information
  - Many detection mechanism focus on the border of the network
- Targeted attack behaviour is visible in the internal network traffic
  - Probing hosts for vulnerabilities
  - P2P C&C channel
  - Local proxies
  - Staging servers
  - ...



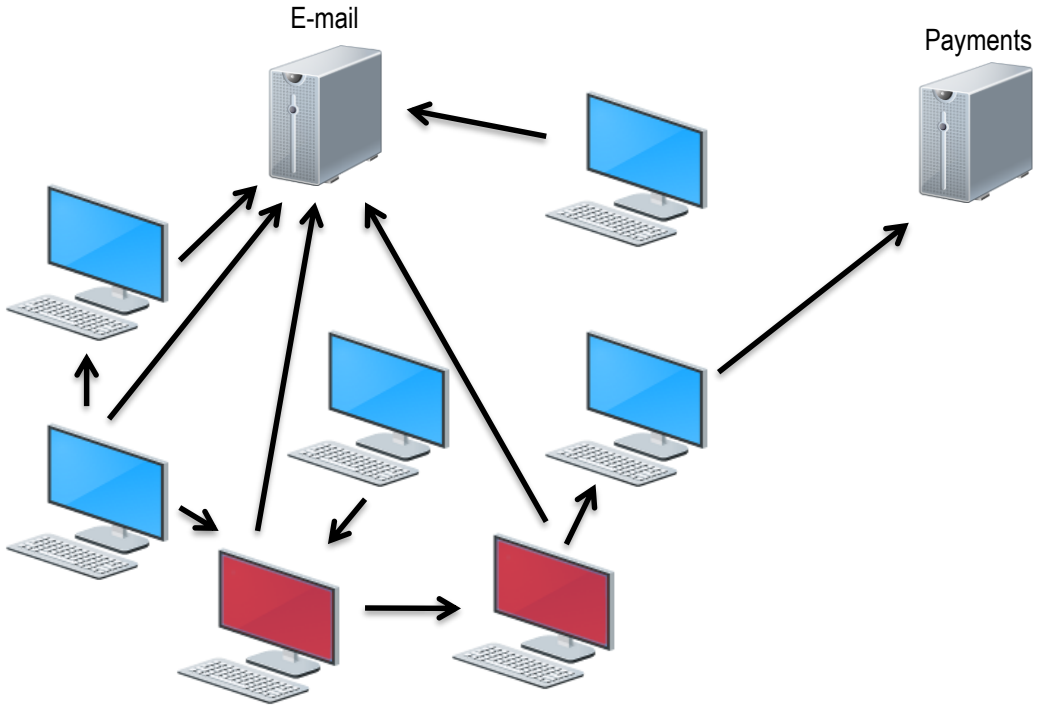
INTERNAL NETWORK ACTIVITY IN CARBANAK (1/4)



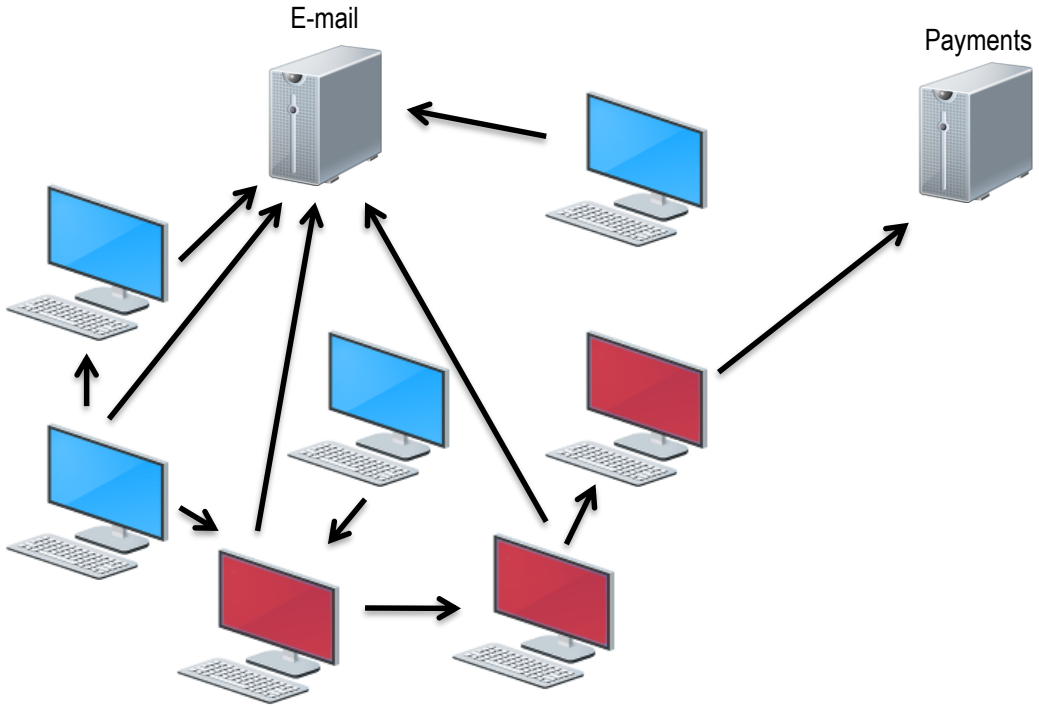
INTERNAL NETWORK ACTIVITY IN CARBANAK (2/4)



INTERNAL NETWORK ACTIVITY IN CARBANAK (3/4)



INTERNAL NETWORK ACTIVITY IN CARBANAK (4/4)



Online-banking
Money was transferred to fraudsters' accounts

E-payment systems
Money was transferred to banks in China and the US

Inflating account balances
The extra funds were pocketed via a fraudulent transaction

Controlling ATMs
Orders to dispense cash at a pre-determined time

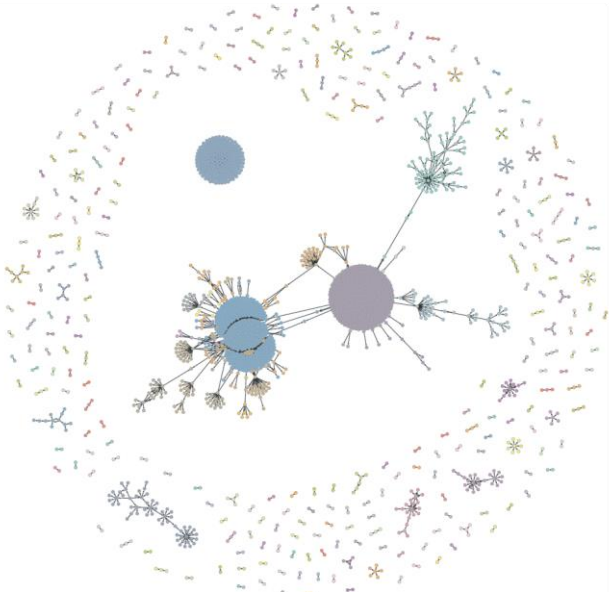
ANOMALY DETECTION APPROACH TO BE ABLE TO DETECT NEW THREATS

- **Misuse-based**
  - Uses signatures
  - Low false positive rate
  - Crucial because of the modular approach in which parts of attacks are re-used
  - Can only discover known threats

- **Anomaly-based**
  - Uses heuristics
  - Often high false-positive rate
  - Capable of discovering new and organization specific threats

INFORMATION CONCENTRATORS ARE USED TO EFFICIENTLY EXTRACT RELEVANT NETWORK DATA

- Full PCAP
  - The amount of data is huge
  - Difficult to monitor the entire network
- Alternatively, we aim for information concentrators within the network
  - DNS servers
  - IAM servers
  - Internal NetFlow
  - ...

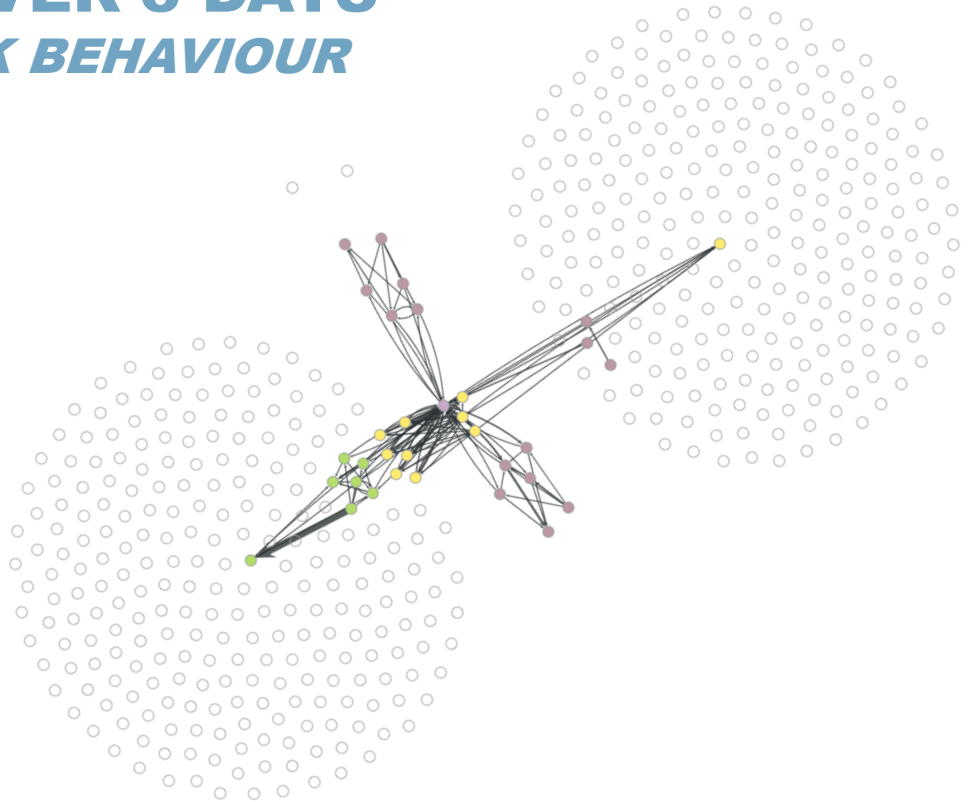


Bank network
Internal DNS data

ANOMALY DETECTION OVER 3 DAYS

BASED ON INTERNAL NETWORK BEHAVIOUR

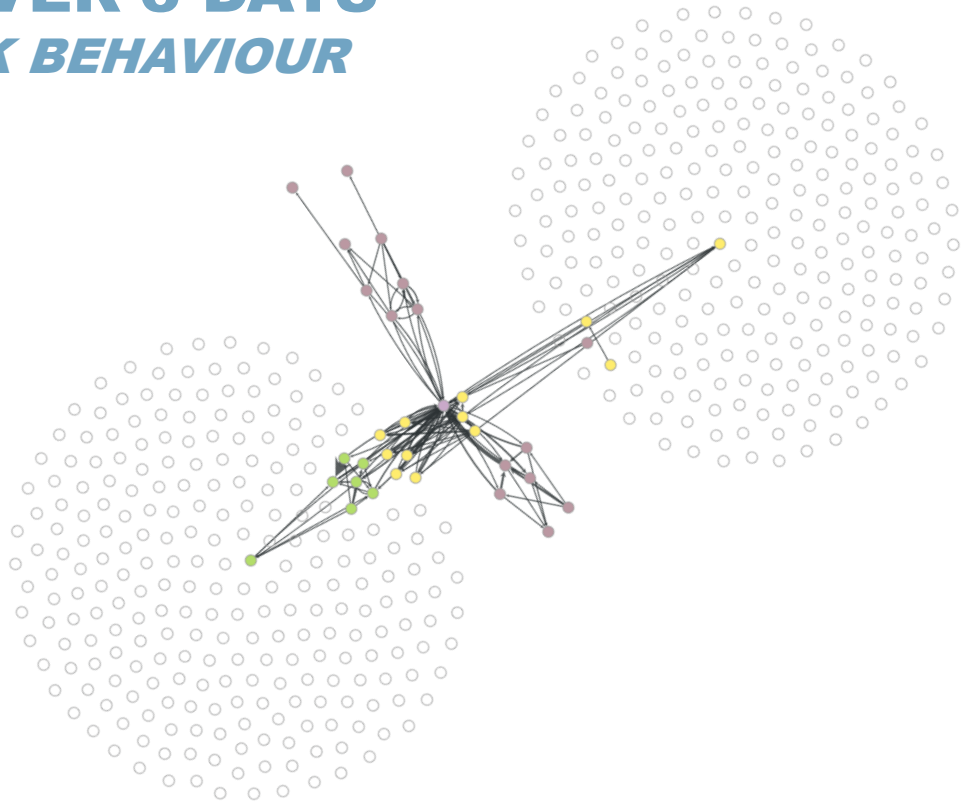
- Friday



ANOMALY DETECTION OVER 3 DAYS

BASED ON INTERNAL NETWORK BEHAVIOUR

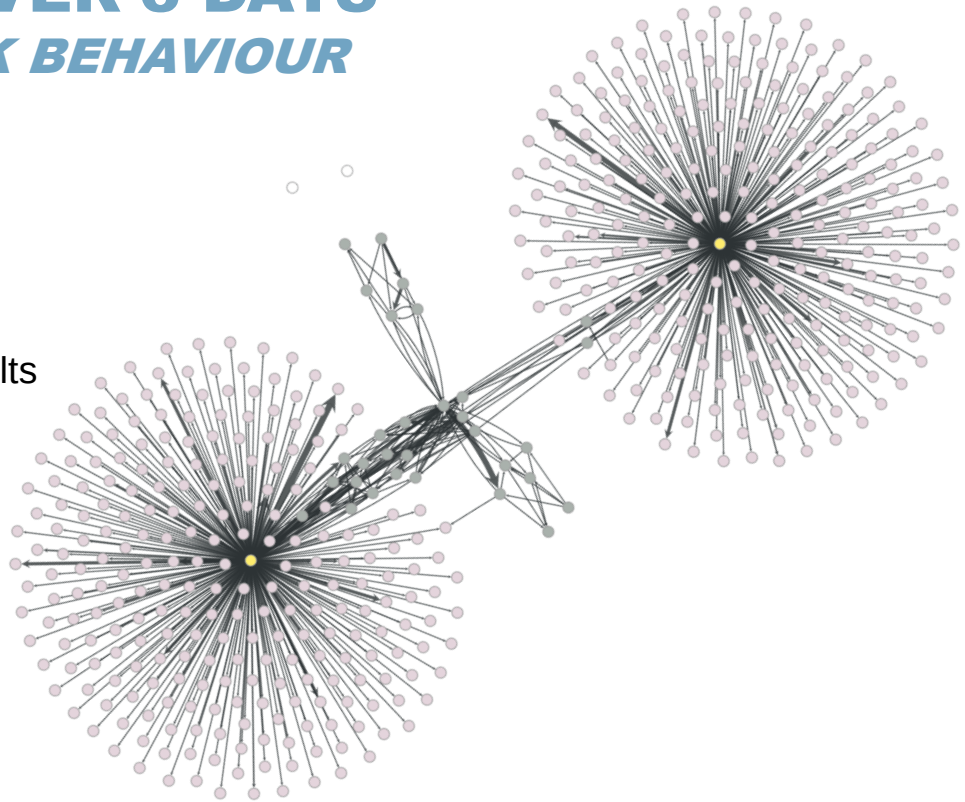
- Friday
- Saturday



ANOMALY DETECTION OVER 3 DAYS

BASED ON INTERNAL NETWORK BEHAVIOUR

- Friday
- Saturday
- Sunday - Port scan influences the clustering results

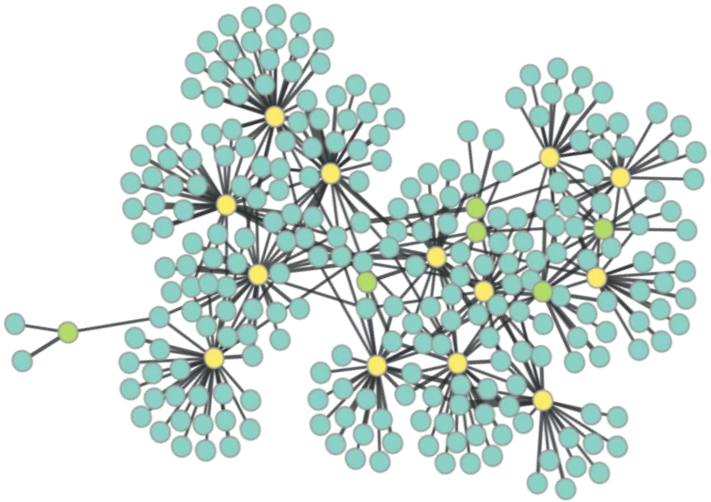


MODEL INTERNAL NETWORK BEHAVIOUR TO DETECT ANOMALOUS HOSTS

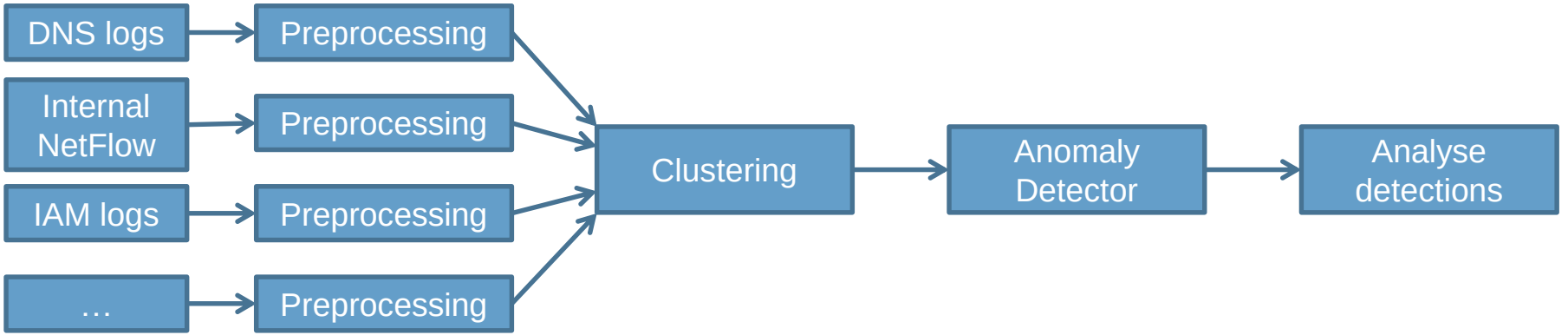
- Modelling individual hosts:
  - Many false-positives
  - Minor day to day differences in network behaviour will be considered anomalous
- Modelling all hosts together:
  - Many false-negatives
  - E.g. a workstation displaying server like network behaviour will not be detected

CLUSTERING HOSTS WITH SIMILAR CONNECTION BEHAVIOUR

- Groups of hosts often behave similarly
  - E.g. mail servers, printers, workstations, mobile devices
- Determine ‘normal’ behaviour **per cluster** and observe when individual host deviates

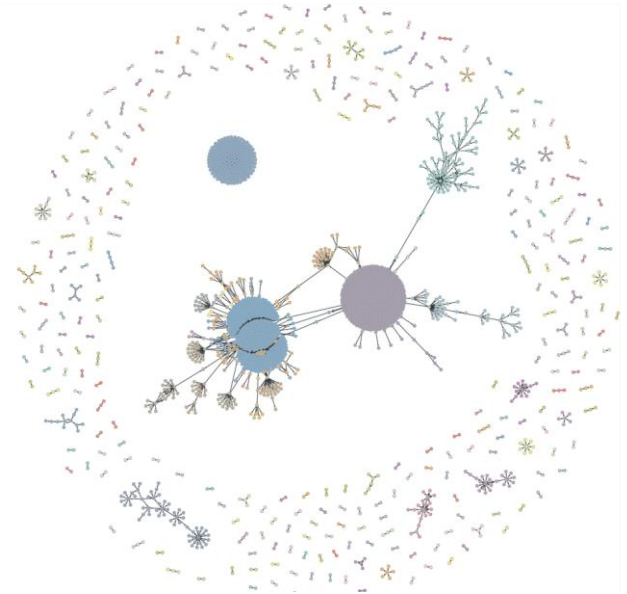


MODULAR MONITORING AND DETECTION FRAMEWORK FOR INTERNAL NETWORK SOURCES



CLUSTERING – LOUVAIN COMMUNITY DETECTION

- Find communities within a network
  - Many connections within the community
  - Little connections across different communities
- Efficient algorithm
- Anomaly detection:
  - *Rare to see connections between different communities*

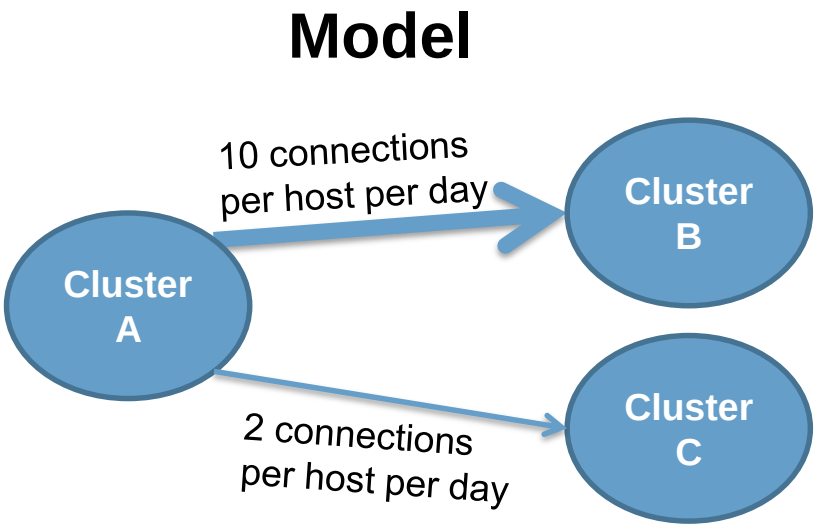


Bank network

Colours indicate the different clusters

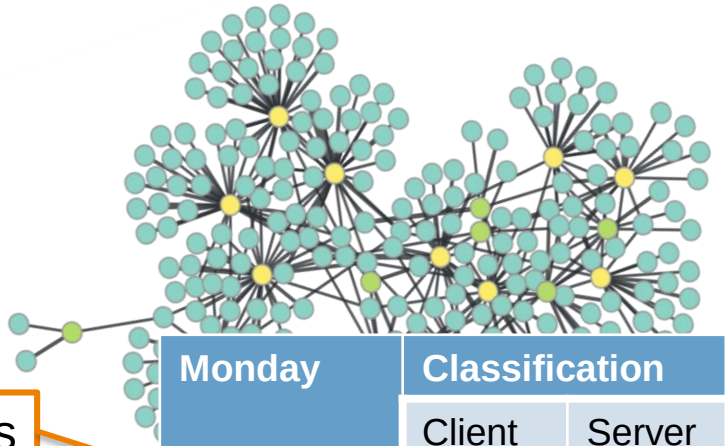
MODELLING THE CONNECTION BEHAVIOUR OF THE CLUSTERED HOSTS

- For each cluster we create a different model
  - This models displays the average number of connections to each of the other clusters
- For all hosts we will determine the number of queries to the each of the different clusters on a new day
- These numbers are then compared to the averages of the model
  - Number of standard deviations from the mean



CLIENT SERVER CLASSIFICATION ALGORITHM BASED ON INTERNAL NETWORK BEHAVIOUR

- Support vector machine based on DNS data
  - Number of times a host is queried
  - Query types
  - Centrality of host within the network
  - …



Friday		Classification	
		Client	Server
Type	Client	99,7 %	0,35 %
	Server	3,2 %	96,7 %

Anomalies

Monday		Classification	
		Client	Server
Type	Client	99,7 %	3,4 %
	Server	8,4 %	91,6 %

CONCLUSIONS AND TAKEAWAYS

- **Internal network traffic:**
  - Valuable source of information for targeted attack detection
- **Behavioural based clustering techniques:**
  - Improve anomaly detection algorithms


THANK YOU FOR YOUR ATTENTION
TNO
innovation for life