Anglia Ruskin University

botprobe

Botprobe - Reducing Network Threat Intelligence Big Data & Pre/Post Forensic Data

adrian.winckles@anglia.ac.uk

project background

PhD: *“a botnet needle in a virtual haystack”*

- a mechanism to **capture** botnet communication traffic in virtualised environments such as ***Cloud Service Providers***.

why?

- cloud providers are building block for IoT
- a great hosting platform for botnets

interesting built environment:

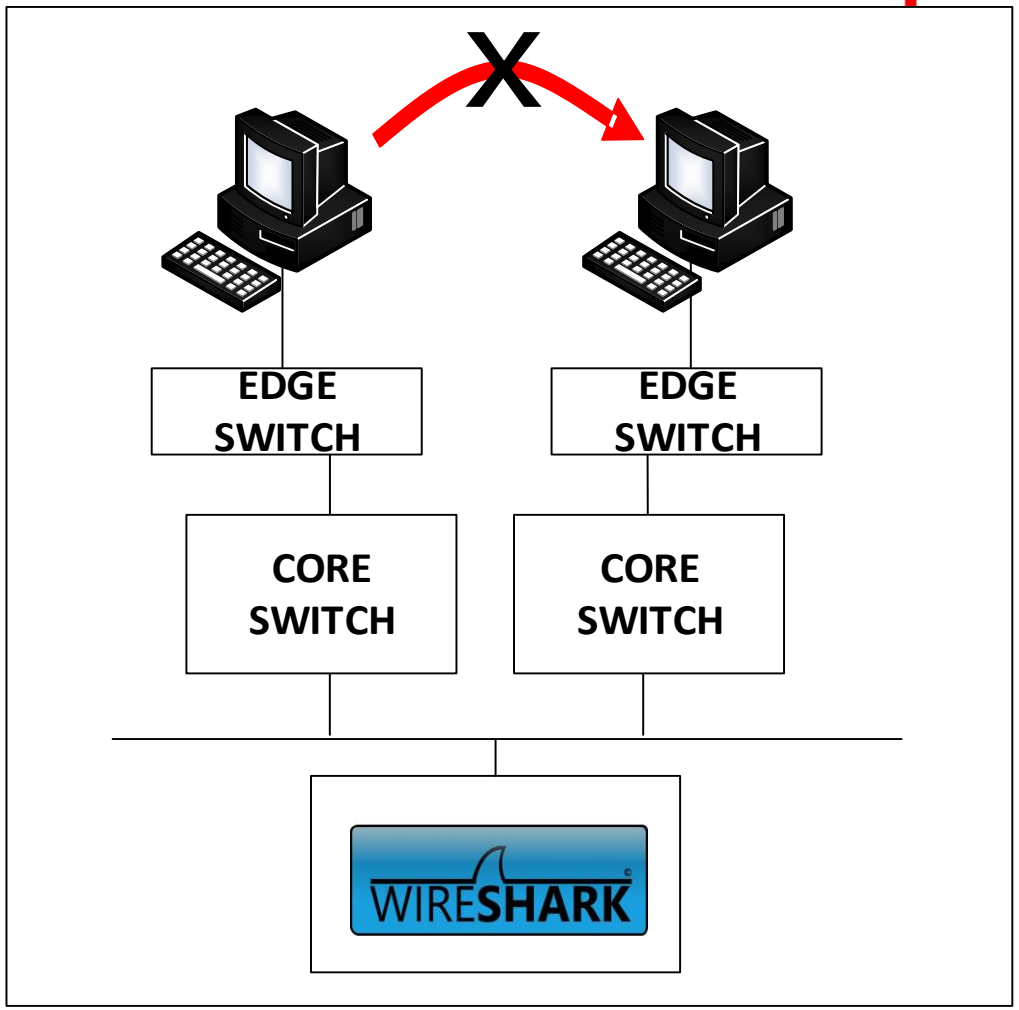
- tenant isolation, data privacy
- internal infrastructure is an attack surface

botprobe

packet capture

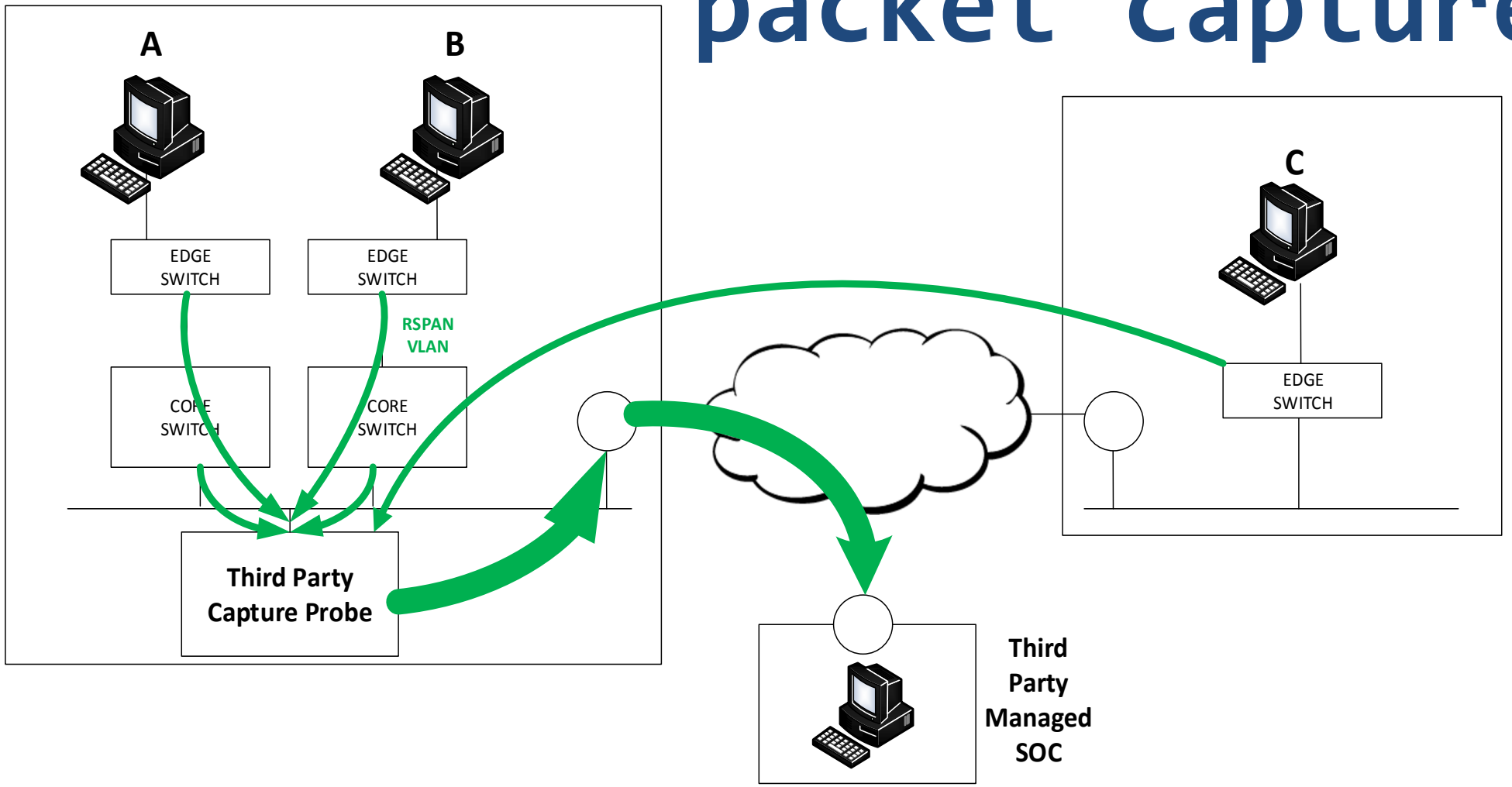
if you want to capture network traffic for threat detection:

use wireshark/tcpdump et al



packet capture

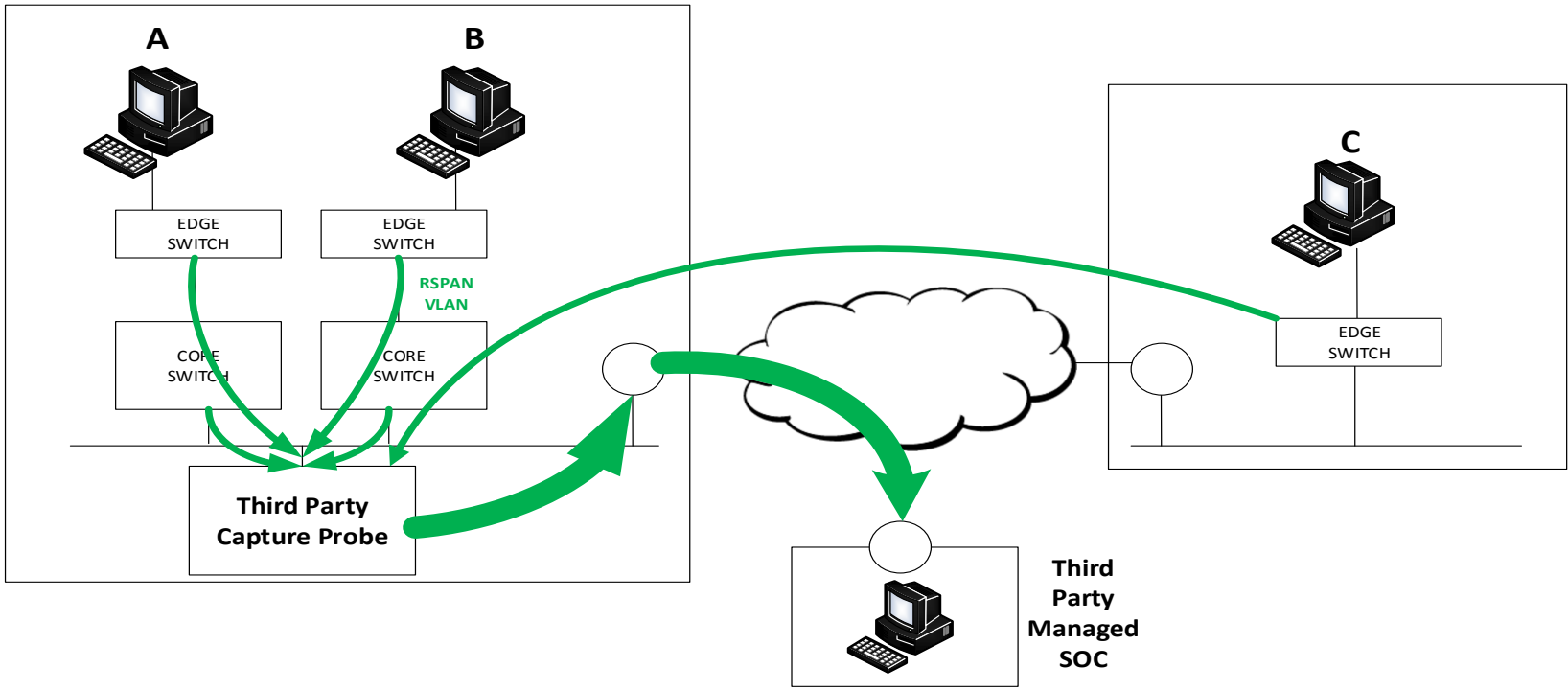
botprobe



packet capture

Three drawbacks in this scenario:

1) port mirroring **doubles** network bandwidth volumes
2) assumes the monitored devices support **mirroring**
3) **big fat pipe** to send traffic to a 3rd party SOC



history lesson

1980’s

Simple Network Management Protocol (SNMP)

- MIB information is limited, so use syslog
- Syslog is unstructured

1990’s

- 1991 – IETF proposed packet aggregation into flows
- 1993 – Disbanded due to lack of interest
- 1996 – Cisco patented NetFlow

- continued…

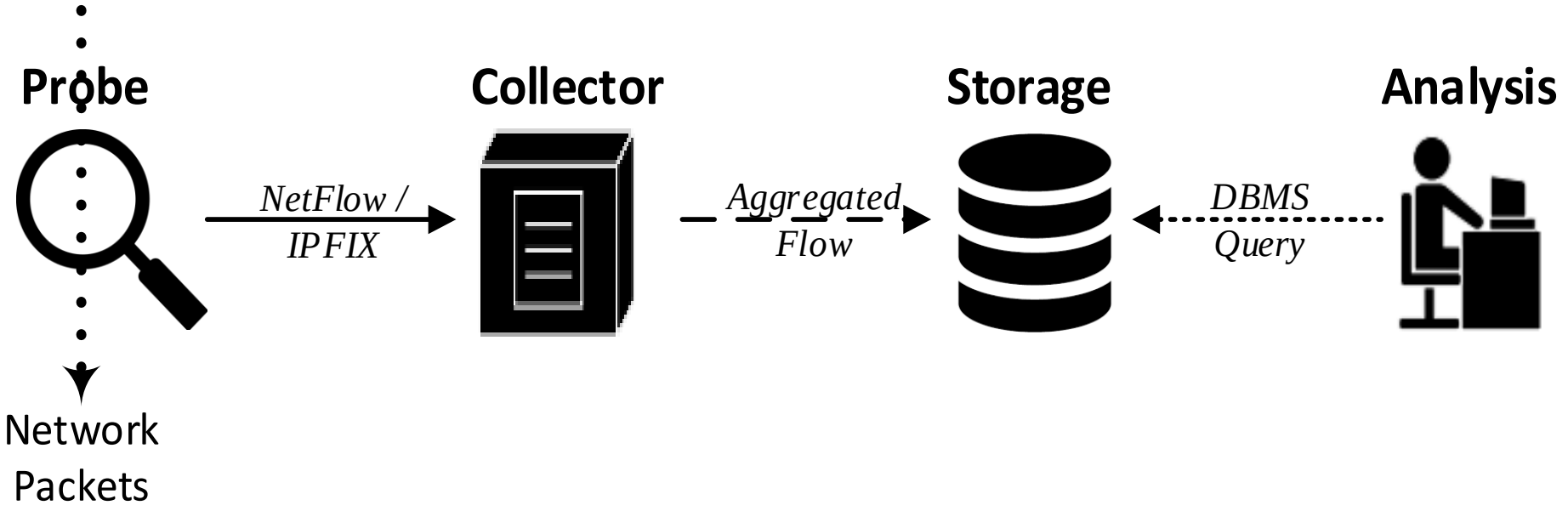
history lesson

1996 – Cisco patented NetFlow

flow aggregates similar traffic based on an attributes tuple:
e.g. 5 field flow tuple: { ***sIP***, ***dIP***, ***sPort***, ***dPort***, ***protocol*** }

- *PCAP is a phone call*
- *flow is the phone bill (who, when, how long)*

flow export architecture



history lesson

1996 – Cisco patented NetFlow
2002 – NetFlow v5
2004 – NetFlow v9

NetFlow was designed for application to **network management**,
but has limitations when applied to **threat detection**:

- NFv5 has 18 **fixed** fields (only **10** useful!)
- **header** information ONLY
- transport layer is **UDP** only
- **no support for**: MPLS, IPv6, VLANs, MAC addresses
- [typically 1:50 sampling rates]

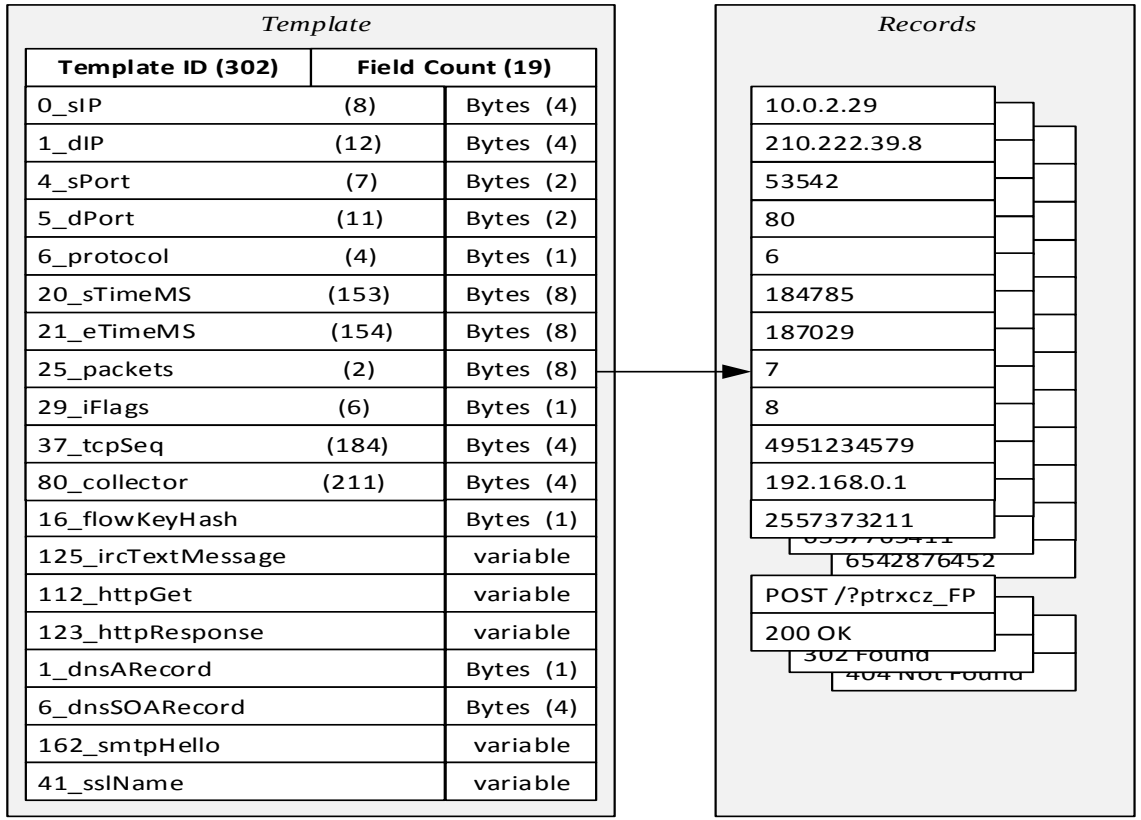
Cisco NFv9 supports (most) of these, but is proprietary.

IP Flow Information eXport

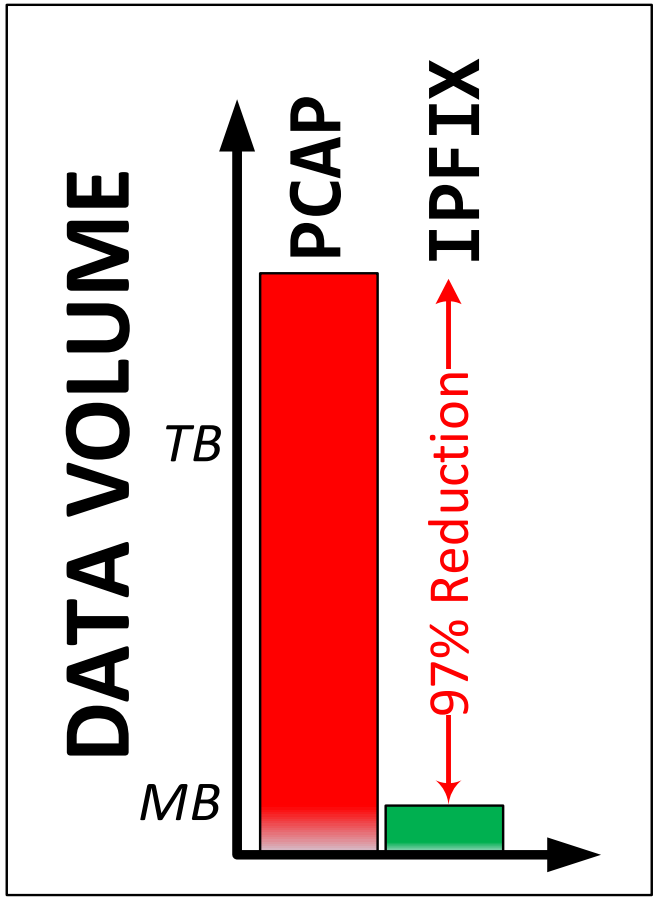
2013 – IPFIX the flow export standard (**RFC7011 - RFC7015**)
IPFIX IS A FLOW EXPORT PROTOCOL IN ITS OWN RIGHT *(not NFv10)*

- Standards-based: **vendor neutrality**
- Extensible: NFv5 – fixed template: 18 fields
  NFv9 - 79 fields (104 if Cisco)
  IPFIX - 433 **Information Elements** (IANA)
- EEs: create your own bespoke **Enterprise Elements**
- Security: security by design
- Future-proof: supports IPv6, MPLS and multi-cast

botprobe template



botprobe performance



	pcap	botprobe
data volumes:	99.4 MB	3.0 MB
load/analysis:	172.5 sec	0.2 sec

botnet detection

repeated 30 botnet experiments:

- **97% less** capture data volume
- **faster** capture
- **no change** to algorithm feeds

		Gates, et al.	BLINC Karagiannis, et al.	Karasaridis, et al.	RISHI Goebel & Holtz	BOTHUNTER Gu, et al.	BOTSNIFFER Gu, et al.	BOTMINER Gu, et al.	Strayer, et al.	BOTLAB John, et al.	Wurzinger, et al.
	Publication Year	2004	2005	2007	2007	2007	2008	2008	2008	2009	2009
	IRC, HTTP, P2P	-	P	I	I	I	I,H	I,H,P	I	H	I,H,P
	P = Packet Capture 5 = NFv5, 9 = NFv9	5	5	5,P	P	P	P	5,P	5	-	P
NFv5 Attributes	srcIPv4	✓	✓	✓	✓	✓	✓	✓		✓	✓
	dstIPv4	✓	✓	✓	✓	✓	✓	✓		✓	✓
	srcPort	✓		✓	✓	✓	✓	✓		✓	✓
	dstPort	✓		✓	✓	✓	✓	✓		✓	✓
	proto	✓	✓			✓	✓		✓	✓	✓
	packetTotal	✓	✓	✓				✓	✓		✓
	byteTotal	✓	✓	✓				✓	✓		✓
	TCPFlag	✓		✓		✓	✓		✓		
	timeStamp	✓	✓	✓	✓	✓	✓	✓	✓		✓
Non-NFv5 Attributes	flowDirection		✓					✓			
	flowsTotal			✓							
	1stPacketSize										
	payloadSize										✓
	irc_Header			✓	✓	✓	✓				
	http_URL				✓		✓			✓	✓
	http_UserAgent										✓
	http_Server										
	http_Response										
	smtp							✓		✓	✓
	dns							✓		✓	

ipfix v pcap

botprobe

PCAP	IPFIX	so what?
SPAN mirroring	Inline TAP	mirroring doubles network bandwidth, TAP is more efficient
dedicated infrastructure	s/w probe on any device	more control over data capture, lower data volumes
plain Text	encryption, replay protection	security by design, can be sent over internet
unstructured data	structured data	easier search
TB data volumes	MB data volumes	97% reduction data volumes
full packet: payload	L3/L7 templates-capture	privacy, lawful inspection.

botprobe

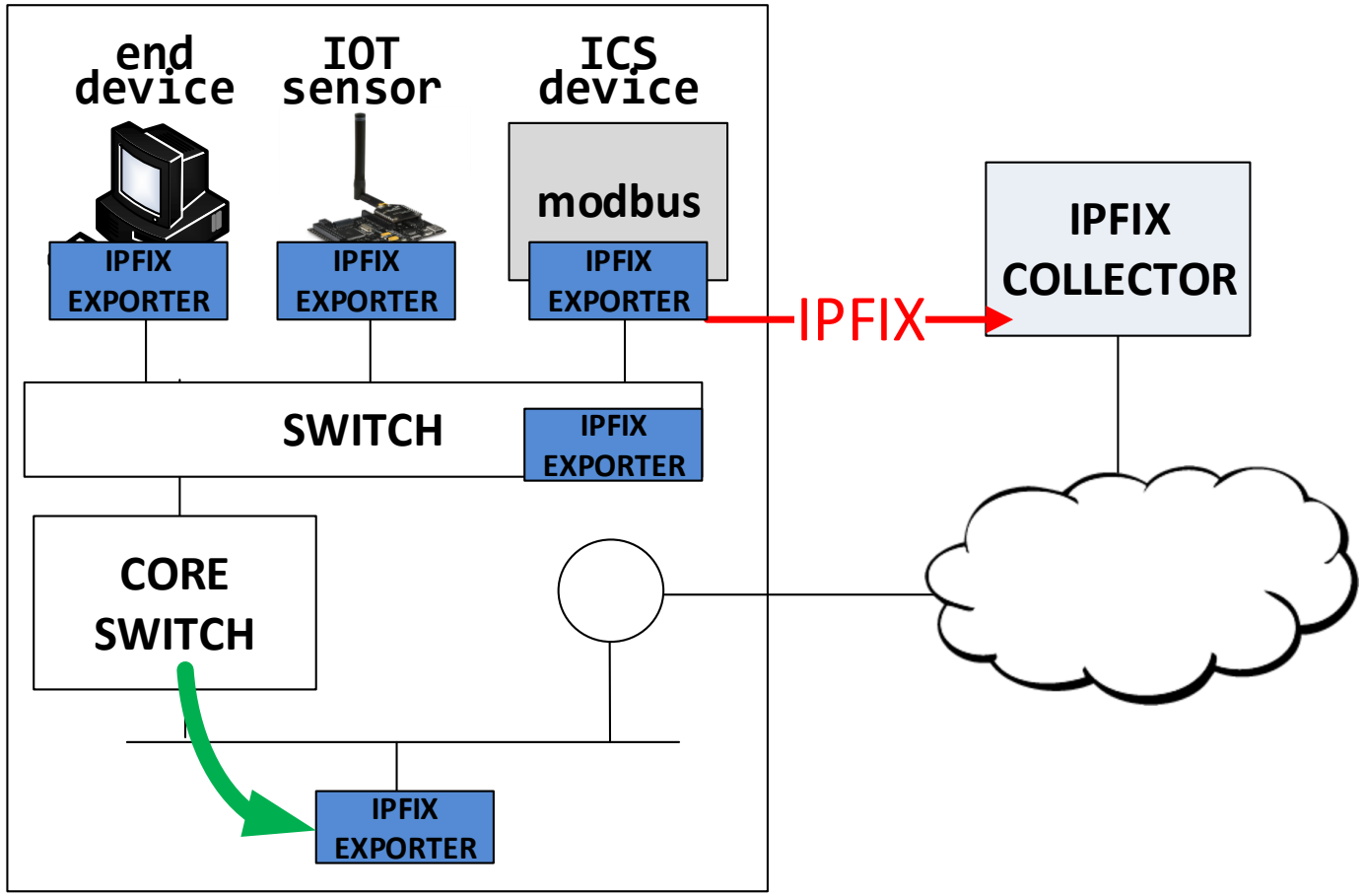
case studies

further IPFIX templates:

- **botprobe** : botnets
- **smtpprobe** : spam traffic
- **httpprobe** : malicious http streams
- **iotprobe** : malicious IoT traffic
- **icsprobe** : malicious Industrial Control Systems traffic

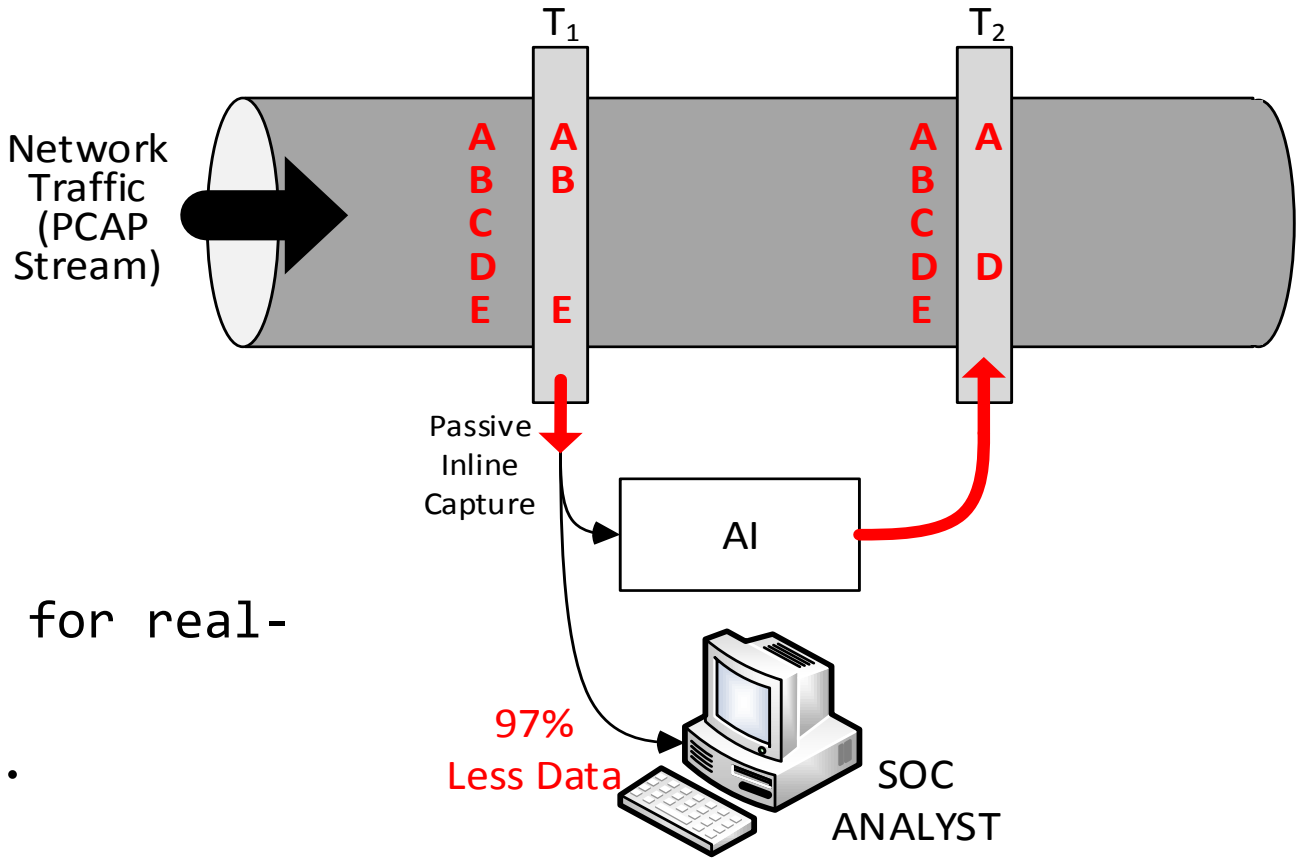
if an attribute is present in a packet [header or payload], we can capture it.

ipfix capture



- **software** probes
- **end-point** protection
- **increased** visibility for fewer probes
- **lower** capture volumes

adaptive capture



machine learning **genetic algorithms** for real-time template adaption as traffic profiles change.

threat detection

three key phases of a cyber attack:

- infection
- detection
- response

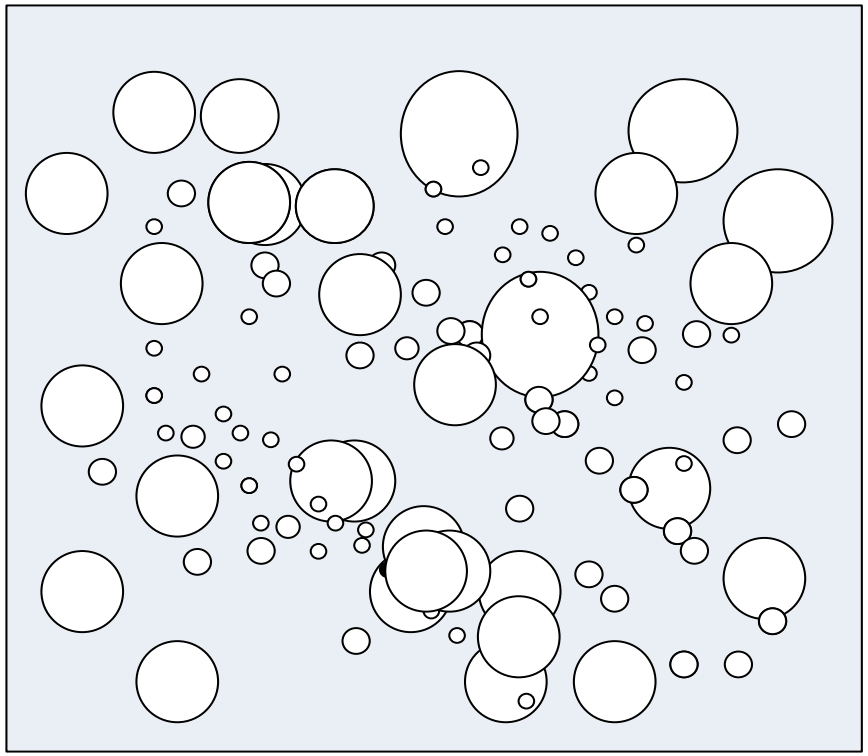
average time to detect a cyberattack is **205 days** (Gartner, 2016)

the cost of a cyber attack is **reputational**, not just financial.

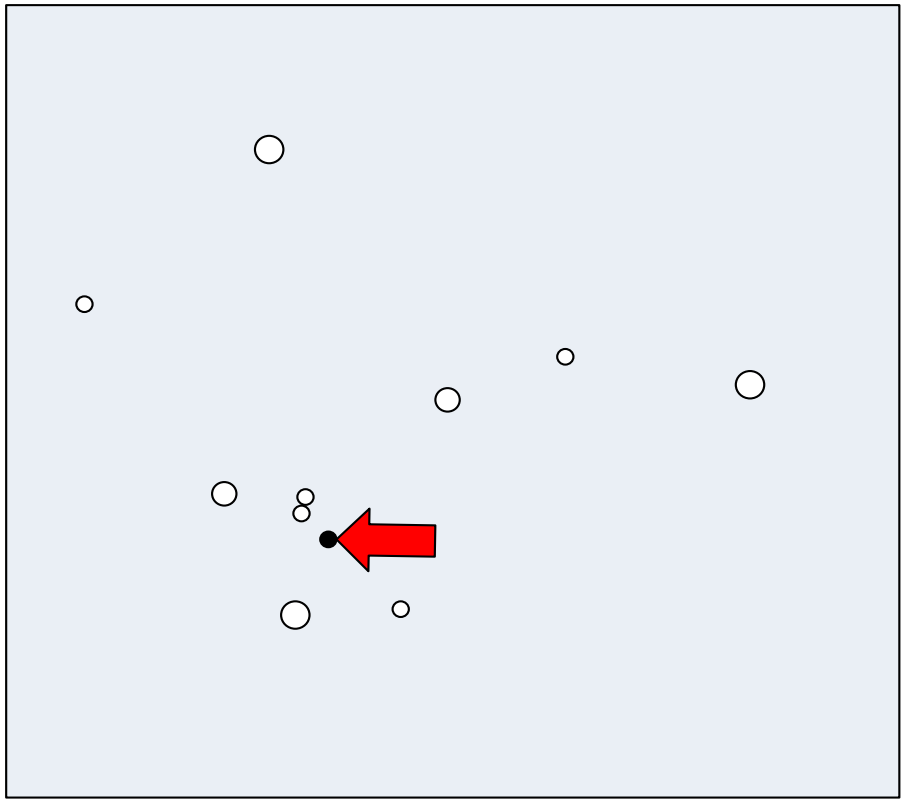
big data challenge

botprobe

Network Big Data



big data challenge



97% reduction in threat intel. data volumes

1) SOC team reacts **faster** to cyberattacks

2) protecting **business assets** and **reputation**

new opportunities

template extensibility + big data reduction =

- automated mitigation
- legal interception
- pre-event forensics
- pcap indexing [flow indexing]
- new detection algorithms [*not just for botnets*]

we need you…

botprobe

if you are interested in collaboration,
We'd love to talk with you:

adrian.winckles@anglia.ac.uk

www.botprobe.co.uk