Logging, Traceability, Threat Intelligence and SOCs

who, what, when, where, how … why?

David Crooks

UKRI STFC

EGI CSIRT/IRIS Security team

david.crooks@stfc.ac.uk

Introduction

- Logging basics
- Central logging
- Data Protection
- Network logging
- Threat Intelligence
- Security Operations Centres

Preamble

- Assessing your risk and having visibility of your services and systems is absolutely essential

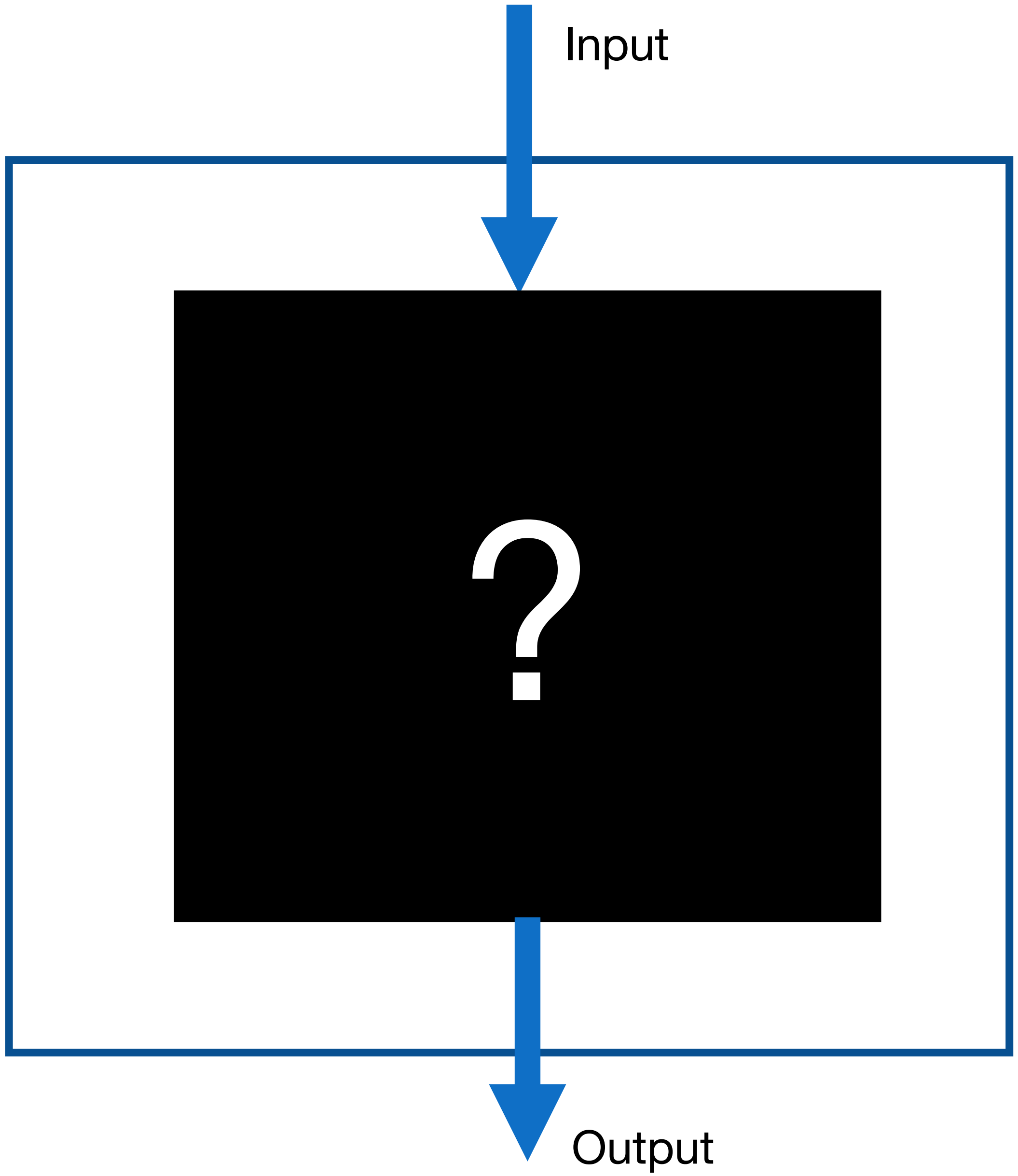
- Everything we’re about to discuss assumes that - to some extent – our area has been assessed for risk

Why do we log?

- To know what happened **in as much detail as necessary**

- Often, security concerns are an extension of operations
  - **What** happened?
  - **When** did it happen?
  - **Where** did it happen?

  - **How** did it happen?
  - **Why** did it happen?



Examples

- Why did this data transfer fail?

- Why did this job only complete partially?

- Which endpoints were involved in this process?

- What did the attacker do?

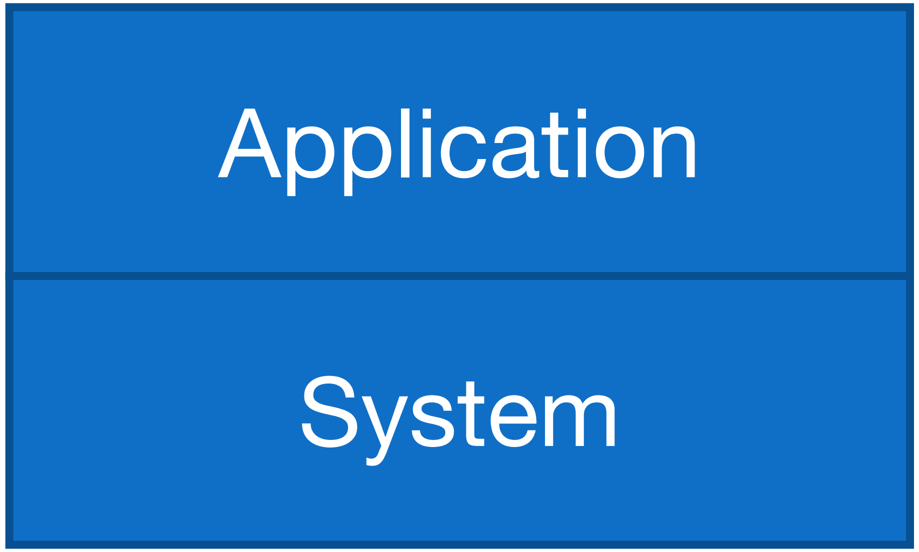
Day to day life

- Logs are an integral part of our technical lives

- But as we head heard yesterday, with this ubiquity comes careful consideration

Host/service logs

- Application logs
- System logs

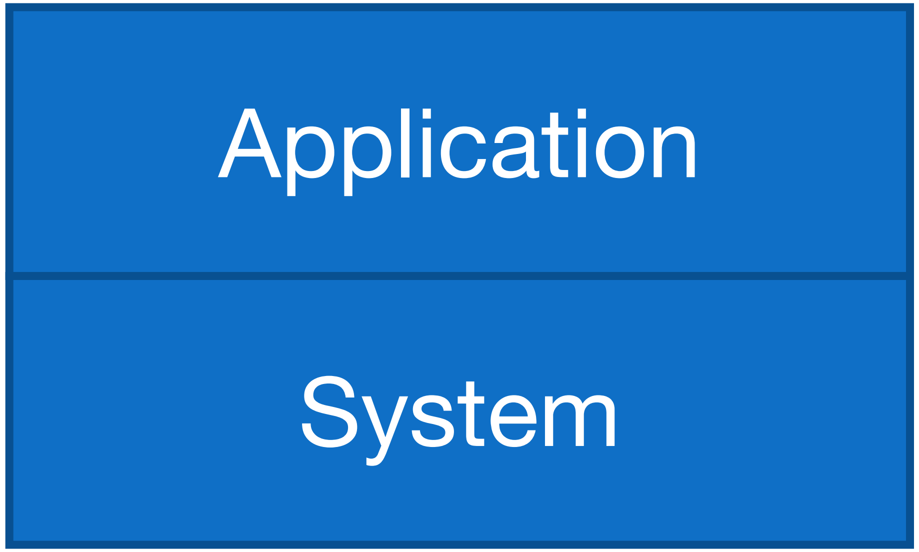


Application

System

Host/service logs

- Application logs
  - Apache
  - Drupal
  - Ceph
  - Dcache
  - ...

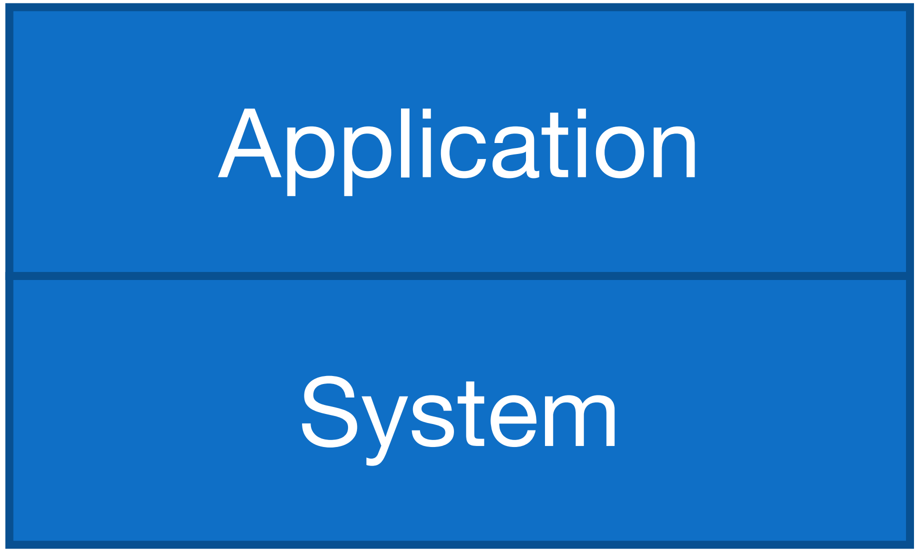


Application

System

Host/service logs

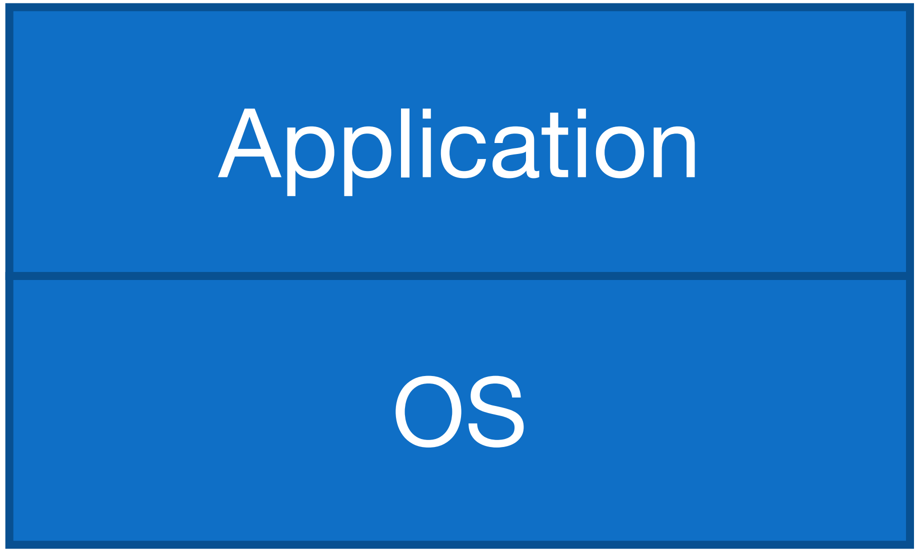
- Application logs
- These depend on the service
- Talk again in traceability, but: service owners are best placed to understand what is useful!


Application

System

Host/service logs

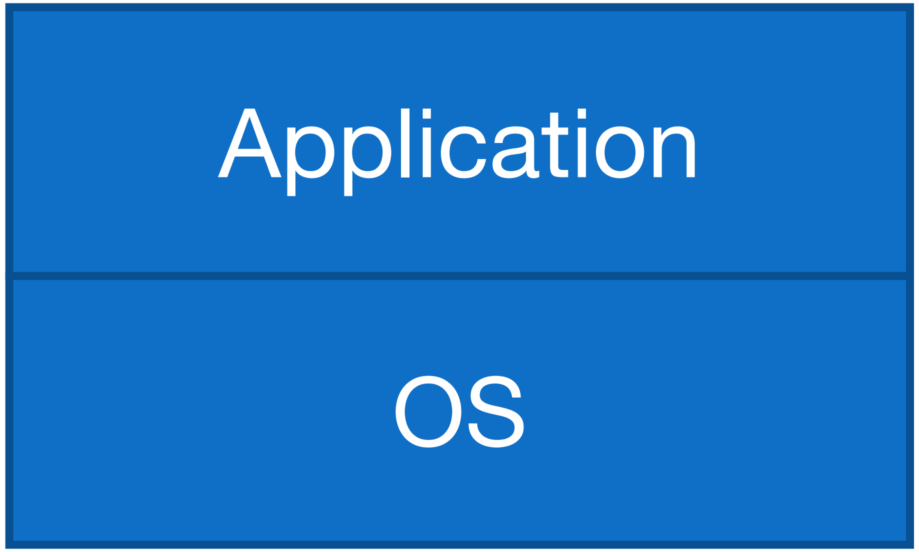
- System logs
- Give us an understanding of the behaviour of the system itself
  - Direct access via ssh
  - System behaviour
  - Auditing over time
- These paths will be for RHEL Distros



Host/service logs

- System logs
  - /var/log/audit.log

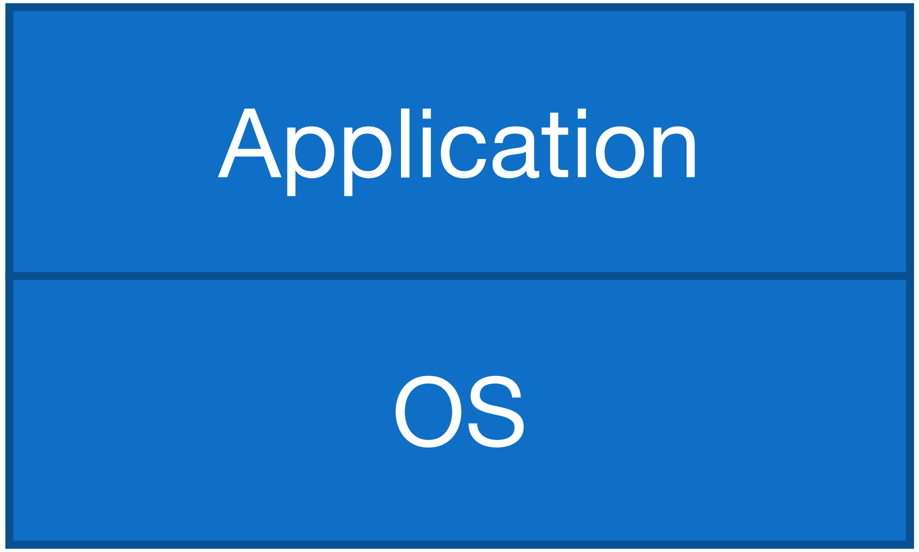
```
type=USER_AUTH msg=audit(1655751006.984:3758): pid=26347 uid=0 auid=4294967295 ses=4294967295
subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 msg='op=pubkey_auth rport=35186 acct="centos" exe="/usr/sbin/sshd" hostname=?
addr=A.B.C.D terminal=? res=success'
type=USER_AUTH msg=audit(1655751006.984:3759): pid=26347 uid=0 auid=4294967295 ses=4294967295
subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 msg='op=key algo=ssh-rsa size=4096
fp=SHA256:48:43:a1:08:47:36:a3:69:1a:d0:72:24:58:f3:e3:07:7d:99:ce:0b:bd:d5:cd:fb:10:bc:37:18:cf:f8:4a:a4 rport=35186 acct="centos"
exe="/usr/sbin/sshd" hostname=? addr=A.B.C.D terminal=? res=success'
type=USER_ACCT msg=audit(1655751006.994:3760): pid=26347 uid=0 auid=4294967295 ses=4294967295
subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 msg='op=PAM:accounting grantors=pam_unix,pam_localuser acct="centos"
exe="/usr/sbin/sshd" hostname=X.Y.Z addr=A.B.C.D terminal=ssh res=success'
type=CRYPTO_KEY_USER msg=audit(1655751006.994:3761): pid=26347 uid=0 auid=4294967295 ses=4294967295
subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 msg='op=destroy kind=session fp=? direction=both spid=26348 suid=74 rport=35186
laddr=A.B.C.D 6 lport=22  exe="/usr/sbin/sshd" hostname=? addr=A.B.C.D terminal=? res=success'
type=USER_AUTH msg=audit(1655751006.996:3762): pid=26347 uid=0 auid=4294967295 ses=4294967295
subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 msg='op=success acct="centos" exe="/usr/sbin/sshd" hostname=? addr=A.B.C.D 6
terminal=ssh res=success'
type=CRED_ACQ msg=audit(1655751006.996:3763): pid=26347 uid=0 auid=4294967295 ses=4294967295
subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 msg='op=PAM:setcred grantors=pam_env,pam_unix acct="centos" exe="/usr/sbin/sshd"
hostname=X.Y.Z addr=A.B.C.D terminal=ssh res=success'
type=LOGIN msg=audit(1655751006.996:3764): pid=26347 uid=0 subj=system_u:system_r:sshd_t:s0-s0:c0.c1023 old-auid=4294967295
auid=1000 tty=(none) old-ses=4294967295 ses=215 res=1
type=USER_ROLE_CHANGE msg=audit(1655751007.128:3765): pid=26347 uid=0 auid=1000 ses=215 subj=system_u:system_r:sshd_t:s0-
s0:c0.c1023 msg='pam: default-context=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 selected-
context=unconfined_u:unconfined_r:unconfined_t:s0-s0:c0.c1023 exe="/usr/sbin/sshd" hostname=X.Y.Z addr=A.B.C.D terminal=ssh
res=success'
type=USER_START msg=audit(1655751007.145:3766): pid=26347 uid=0 auid=1000 ses=215 subj=system_u:system_r:sshd_t:s0-s0:c0.c1023
msg='op=PAM:session_open
grantors=pam_selinux,pam_loginuid,pam_selinux,pam_namespace,pam_keyinit,pam_keyinit,pam_limits,pam_systemd,pam_unix,pam_lastlog
acct="centos" exe="/usr/sbin/sshd" hostname=X.Y.Z addr=A.B.C.D 6 terminal=ssh res=success'
```



Host/service logs

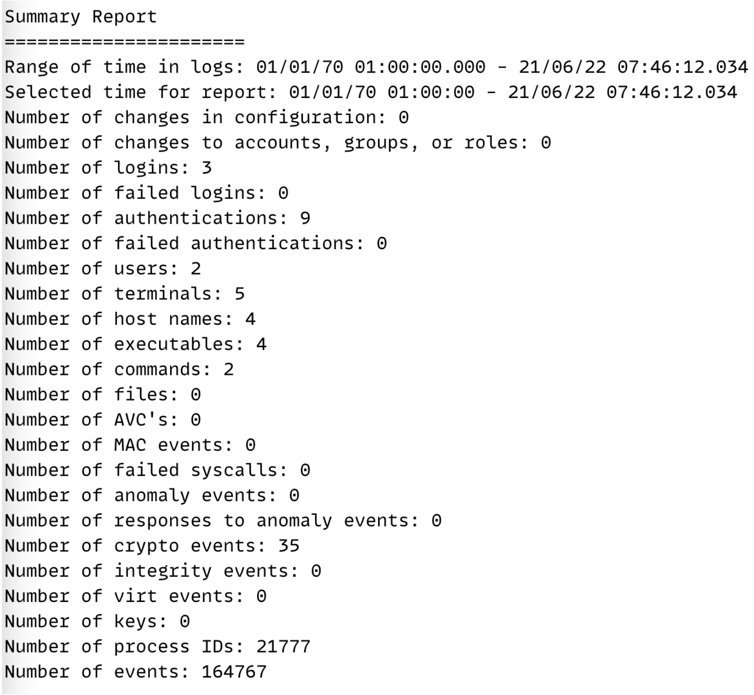
- System logs
  - /var/log/audit.log

- `aureport` can be used to get summary information

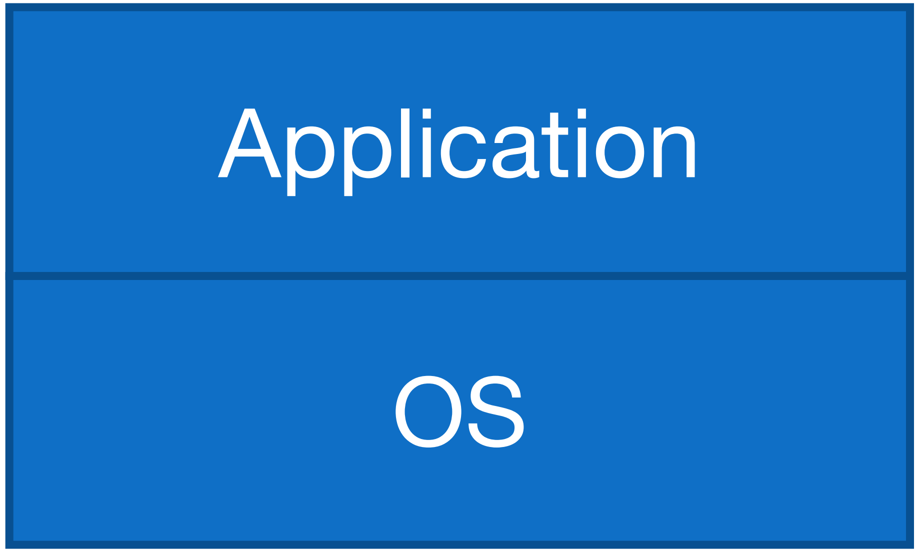


Host/service logs

- System logs
  - /var/log/audit.log



```
Summary Report
======================
Range of time in logs: 01/01/70 01:00:00.000 - 21/06/22 07:46:12.034
Selected time for report: 01/01/70 01:00:00 - 21/06/22 07:46:12.034
Number of changes in configuration: 0
Number of changes to accounts, groups, or roles: 0
Number of logins: 3
Number of failed logins: 0
Number of authentications: 9
Number of failed authentications: 0
Number of users: 2
Number of terminals: 5
Number of host names: 4
Number of executables: 4
Number of commands: 2
Number of files: 0
Number of AVC's: 0
Number of MAC events: 0
Number of failed syscalls: 0
Number of anomaly events: 0
Number of responses to anomaly events: 0
Number of crypto events: 35
Number of integrity events: 0
Number of virt events: 0
Number of keys: 0
Number of process IDs: 21777
Number of events: 164767
```



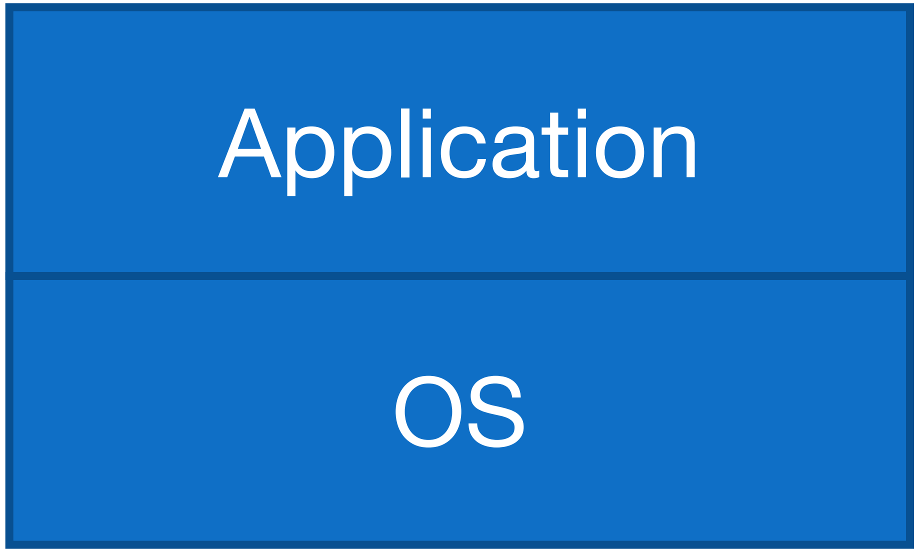
Application

OS

Host/service logs

- System logs
    - Auditbeat

- Part of the elasticsearch Beats set of tools that can also extract and effectively parse audit data



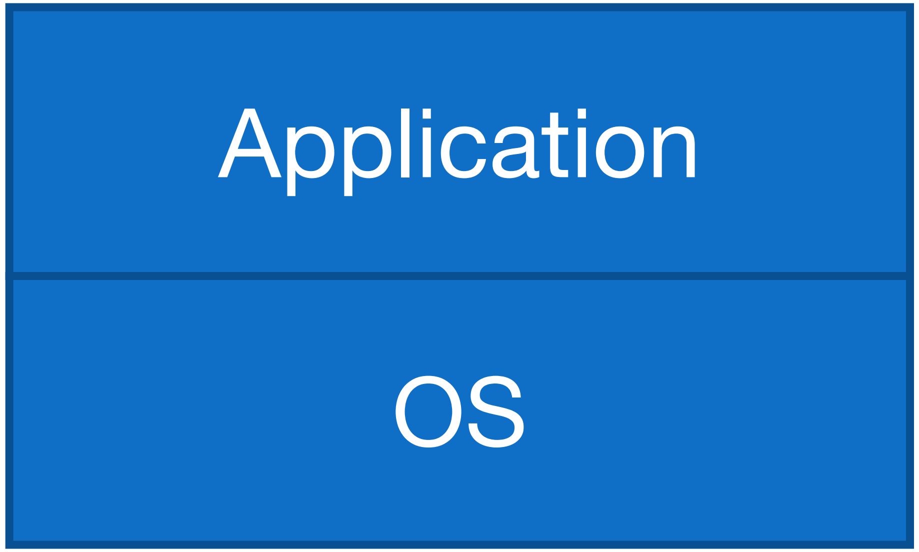
Application

OS

Host/service logs

- System logs
  - /var/log/messages

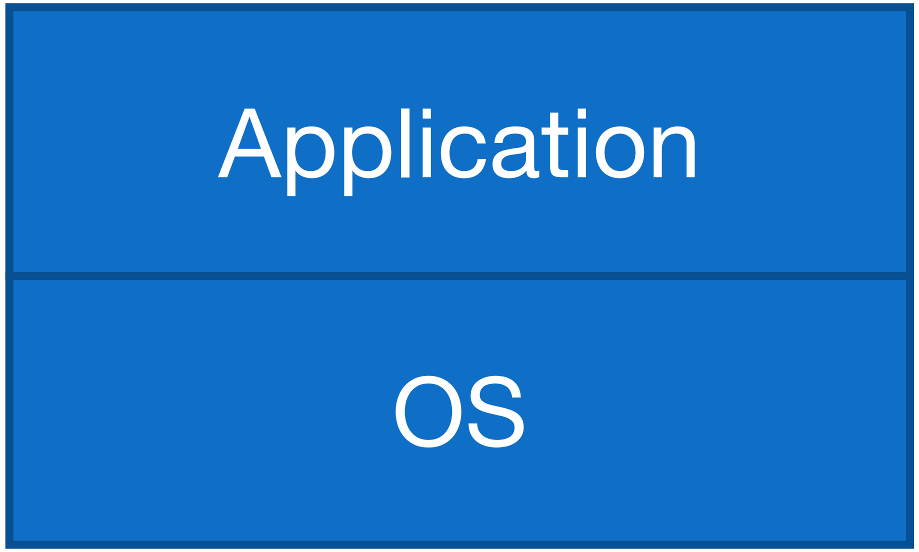
Records global log messages, system notifications including those during boot



Host/service logs

- System logs
  - /var/log/secure

Records successes and failures for users using `ssh` to access the system

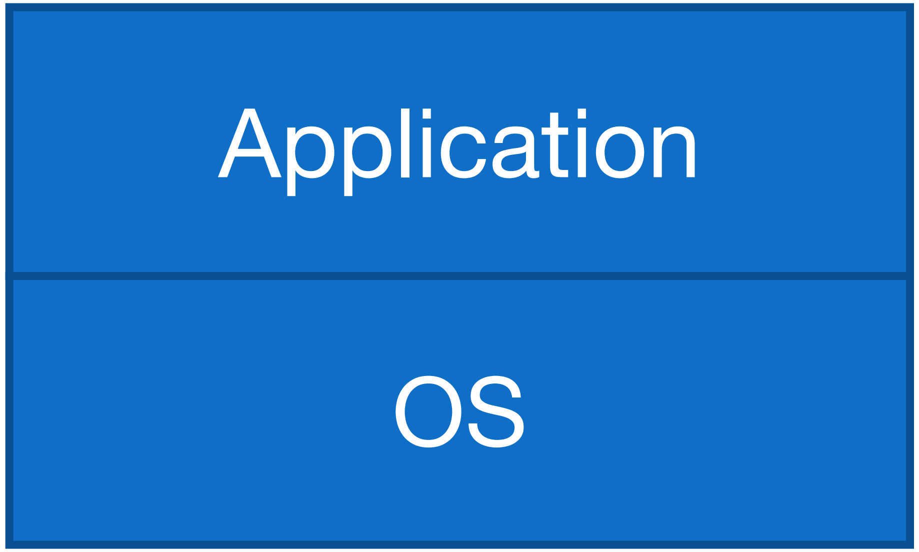


Host/service logs

- System logs
  - /var/log/secure

```
Jun 19 22:18:36 hostname
sshd[26877]: Accepted
publickey for user from
A.B.C.D port 60096 ssh2: RSA
SHA256:…
```

Success!



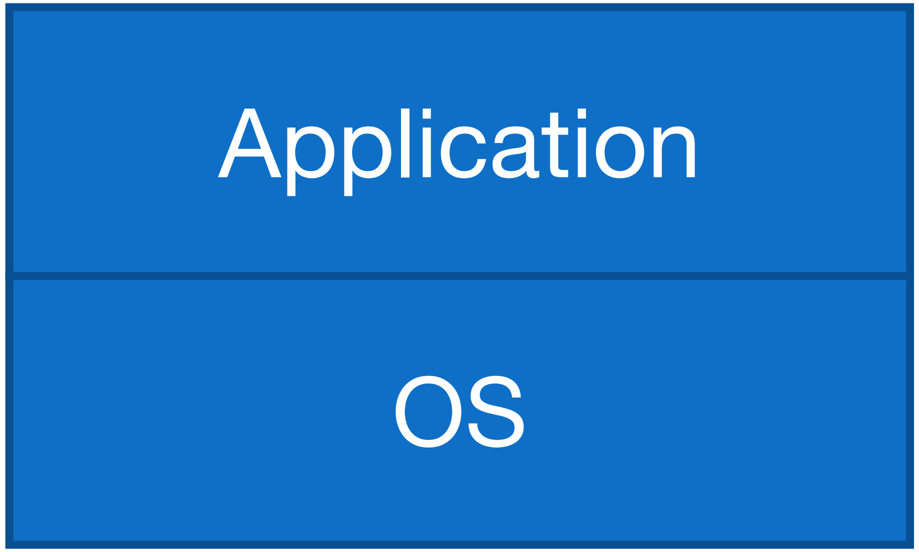
Application

OS

Host/service logs

- System logs
  - /var/log/secure

```
Jun 20 19:08:58 hostname
sshd[7555]: Invalid user admin
from A.B.C.D port 36844
```



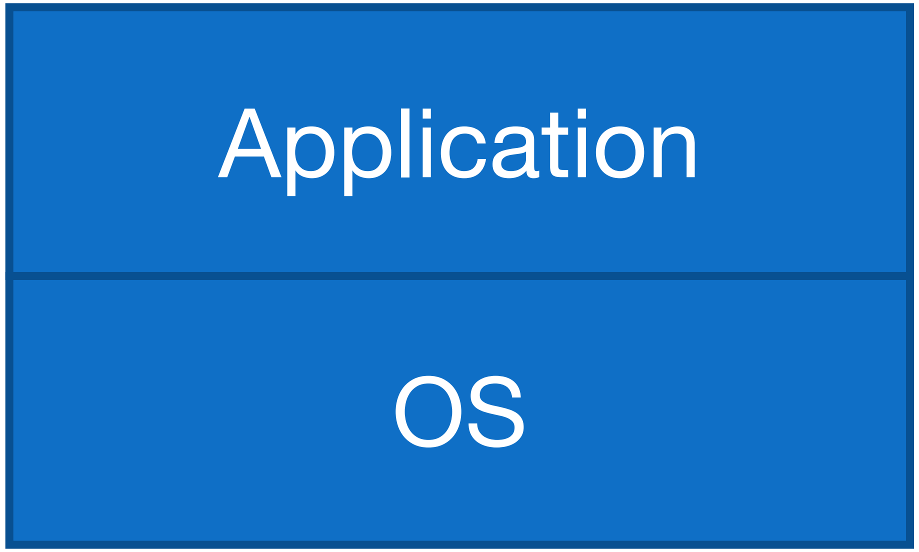
Application

OS

Host/service logs

- System logs
  - /var/log/secure

```
Jun 20 19:08:58 hostname
sshd[7555]: Invalid user admin
from A.B.C.D port 36844
```



Application

OS

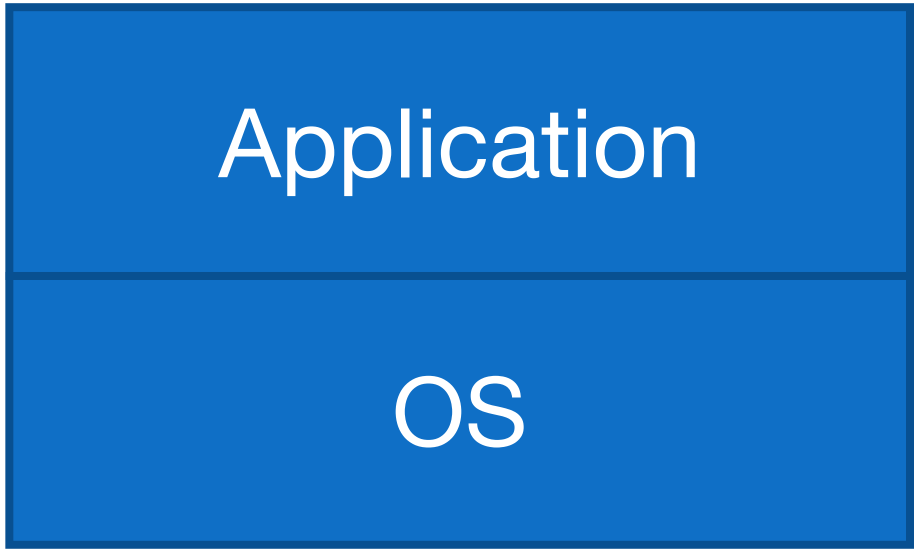
Host/service logs

- System logs
  - /var/log/secure

… this is why you harden your systems (although only a *real* problem if they succeed)

A primary source of checking for malicious access

Unless?



A successful attacker

- Gains access via a weak password (`password2022-2`)

- Installs a compiler, builds some code…

- … hides their tracks by truncating the logs

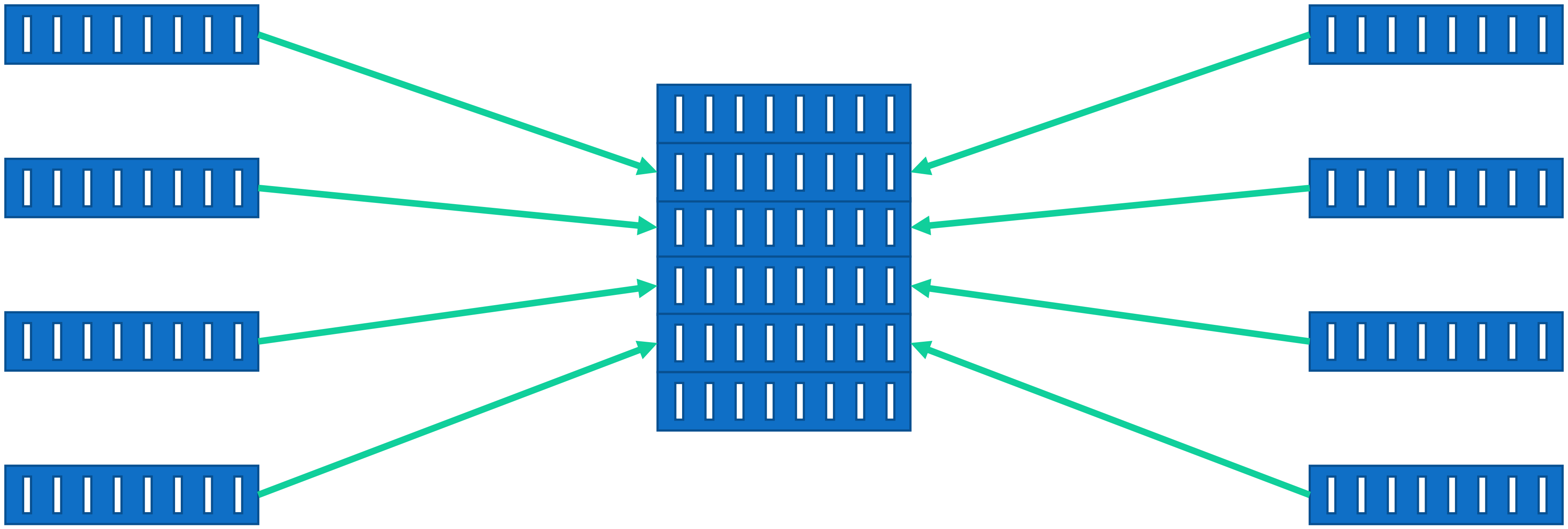
Central logging

- Logs are data

- Vulnerable to deletion or corruption

- Back them up!

Central logging



Central logging

- One of the single most important things to do for the security of a service

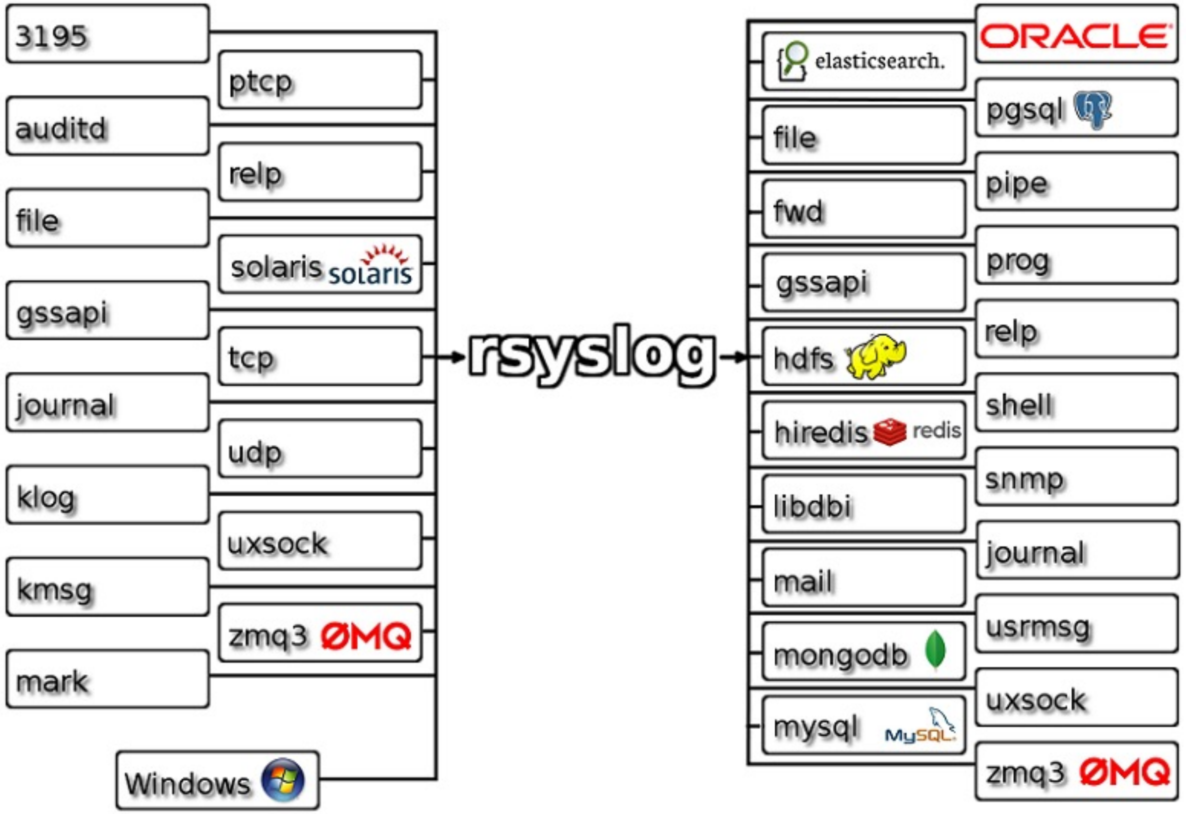
- Helps incident response

- Helps correlate logs between hosts

rsyslog

- rsyslog is a well-featured logging engine

- rsyslog and syslog-ng are both feature-rich successors to the original syslog



https://www.rsyslog.com

rsyslog and other tools

- Especially at this point, storing raw logs is not the most useful

- Use a tool like elasticsearch to allow better searching an querying of the data

OSSEC/Wazuh

- OSSEC is a very nice host-based IDS that will aggregate logs in a server/client topology
- Customisable rules
- Very flexible

https://www.ossec.net

OSSEC/Wazuh

- Wazuh is a modern development of OSSEC that integrates tightly with elasticsearch

- Important when considering defence in depth – having one exactly one tool to monitor your system is **not** optimal

wazuh.

https://wazuh.com/

Wazuh/OSQuery

- Wazuh can monitor many useful things at the host level
  - File integrity + checksums
  - Configuration Assessment
  - Extended Detection and Response

- OSQuery is a nice tool that provides an SQL interface to system information

https://wazuh.com/

https://osquery.io

System + application logs

- Discussed some key system logs

- Application logs are best understood by their service owners: how to choose what you need?

System + application logs

- We can’t store an infinite amount of logs

- And we don’t want to

“ too much data looks like noise”

Data protection

- I am not a lawyer ☺

Data protection

- We are in an era where individual privacy rights are taken particularly seriously
- This is not something that should hinder our security work

- GDPR
- CERN OC11
- Development of UK data protection laws
- Working with laws in other countries

GDPR and CSIRT activities

- In GDPR and associated findings the exchange of logs for incident response is recognized as a useful activity

- https://www.first.org/blog/20171211_GDPR_for_CSIRTs

- We **do** need to be careful about what we store, why, and for how long

Log retention

- In WLCG, for a long time 90 days was the retention period set by policy

- Now moving towards 180 days or more: why?

Log retention

- The number of incidents that have their beginning many months ago

- Only having logs for 90 or 180 days means we lose visibility

- 12 – or 13 – months is where we might set our sights

Log retention: practical matters

- Of course, there are practical matters
  - Logs take up room
- Central logging **also** makes capacity planning easier
  - Build to a set of services that are logged
- Continuous improvement is important

Log retention: practical matters

- Our architecture will suggest where and how many logs we can keep

- This can and should evolve over time

- Focus on sustainable development

Traceability

- For security, we want the logs that will help us piece together a set of events
    - When did someone gain access?
    - What did they do on the host?
    - Where did they go next?
    - What other hosts did they talk to?

Traceability

- Traceability is the ability for us to trace the activity associated with a particular user and/or particular workflow

- Want to be able to track the entire lifecycle
  - Initiation
  - Primary events
  - (External) communications
  - Closeout

Traceability

- Core system logs are essential; for application logs we want anything that helps piece these together

- Debug logs don’t help with this

- It is likely that this will **also** evolve over time

- Make a plan and iterate based on your risks and resources

Split traceability

- In our current circumstances, it is **highly likely** that the logs from a particular service – or even facility – will not be sufficient to track the activity of a user or group

- Why?

Split traceability

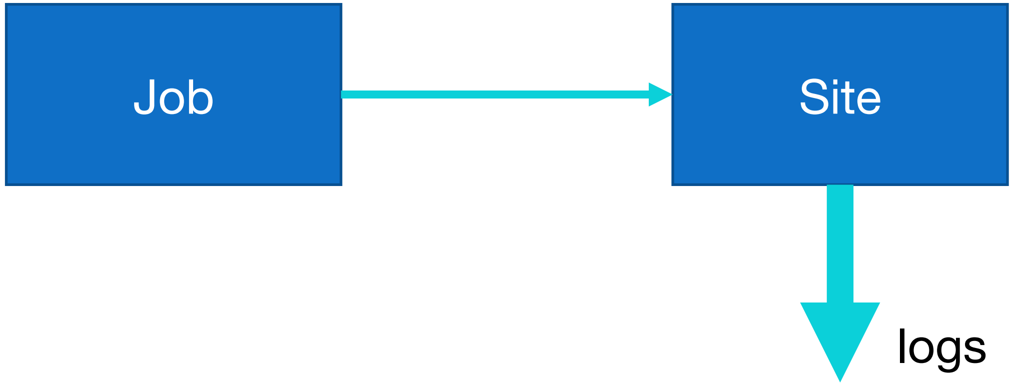
- In research and education, invariably work as part of a bigger infrastructure, federation or federation of federations

Split traceability

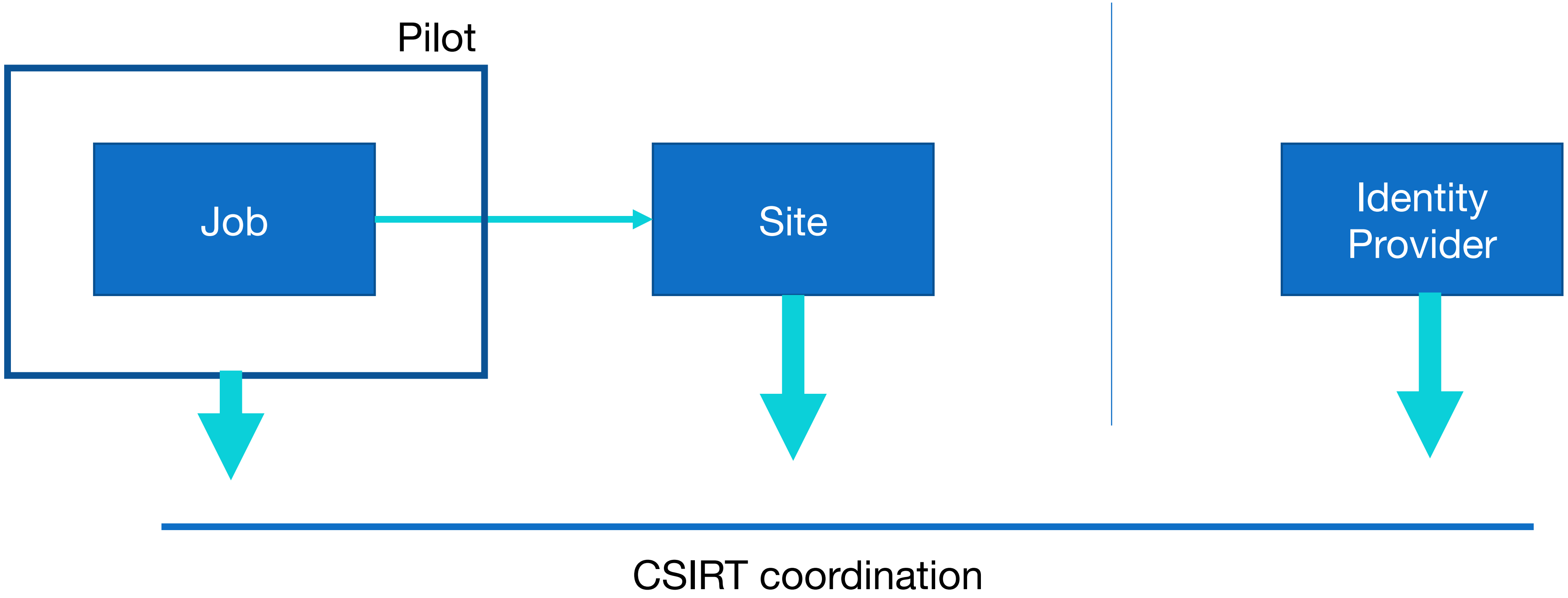
- Many (most!) of our activities involve many services composed together
  - WLCG pilot jobs
  - **Cloud services**

- We can **no longer** rely on the logs on a single host/in a single facility to assemble the full picture of a user's activity

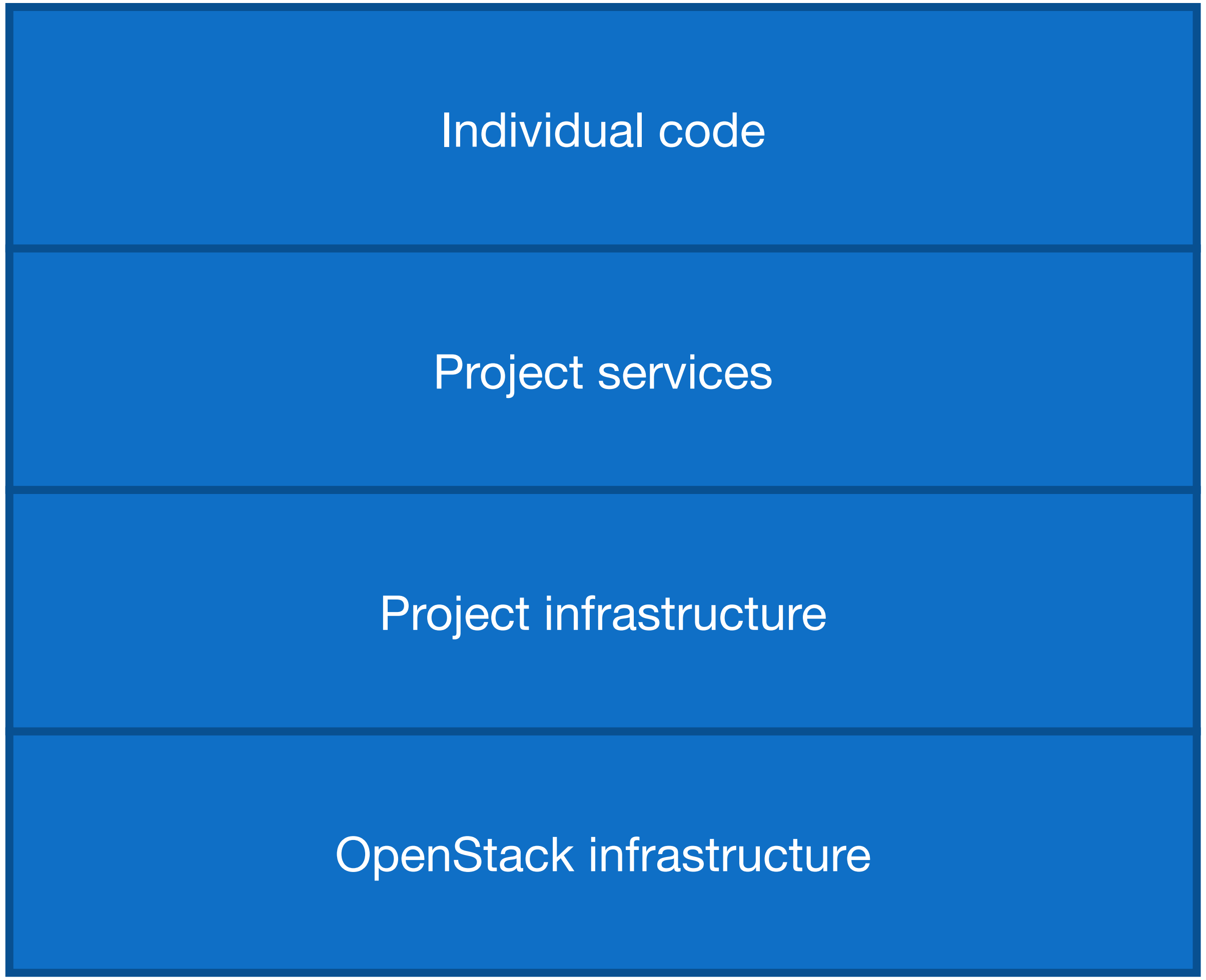
grid jobs: before



pilot jobs: after



Cloud services



Individual code

Project services

Project infrastructure

OpenStack infrastructure

How do we check we our traceability?

- Planning and policy
- Collaboration and cooperation
- **Testing**
- Find use cases that are appropriate for you and try them out!

Network logging

- We’ve talked about host based logs

- What’s happening on the network?

Sources of network logs

- Routers

- Host-based generators

- Monitoring

Netflow and sflow

- Netflow and sflow are different but similar methods of storing **metadata** about network connections
  - Endpoints/duration/…
- Most switches we'll use will generate one or the other
- Can generate on-host
  - `hsflowd`

Netflow came from Cisco

sflow came from InMon

Netflow and sflow

- Pros
  - Ubiquitous
  - Easy to generate

- Cons
  - Sampled

- In general, have **sampled** data from netflow and sflow
  - Useful for long term connections but not forensically useful

Netflow and sflow

- Netflows are especially useful at a high level
    - NRENS

- You **can** produce 1:1 data, but…

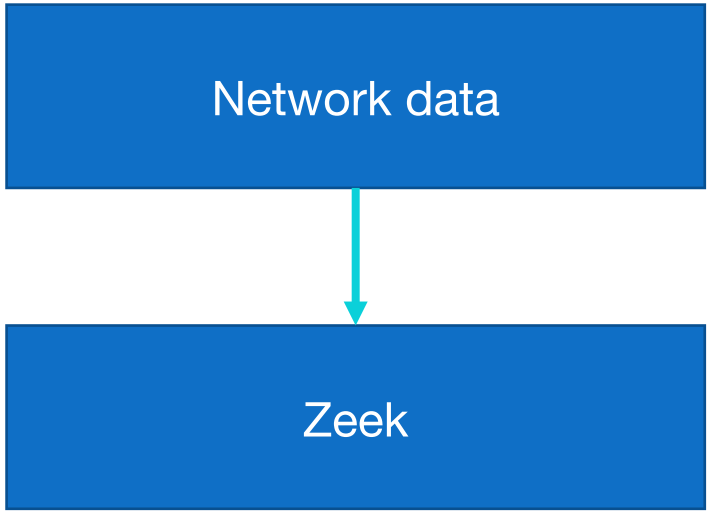
Deep Packet Inspection

- Using a tool that analyses every packet it sees will yield rich information
  - Metadata
  - File information
  - Certificate information…
- Can’t see inside encrypted streams unless you do decryption

Zeek

- Zeek is an example of a current network IDS in broad use in the US and EU
  - Ingest data by taking tap of network traffic
  - Optical or port spanning

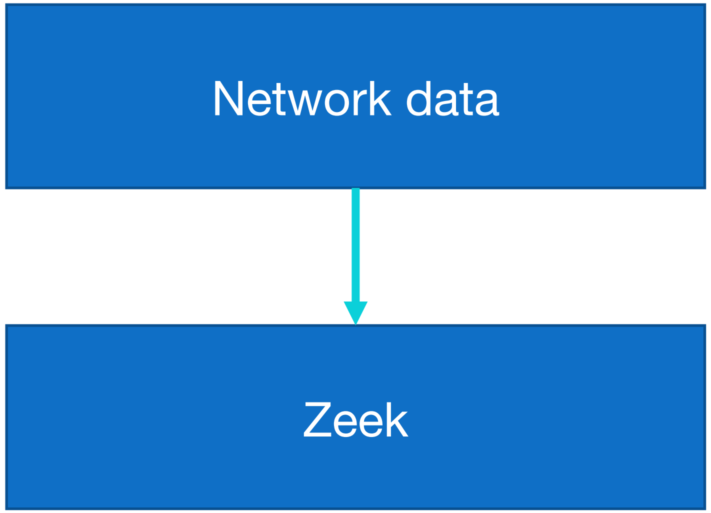
- Single threaded, works by running a set of scripts against each packet
  - Scale out by building a zeek batch farm



Zeek

- This gives us forensic level results
  - Every packet is tracked

- But this is computationally expensive
  - Need care in choosing deployment

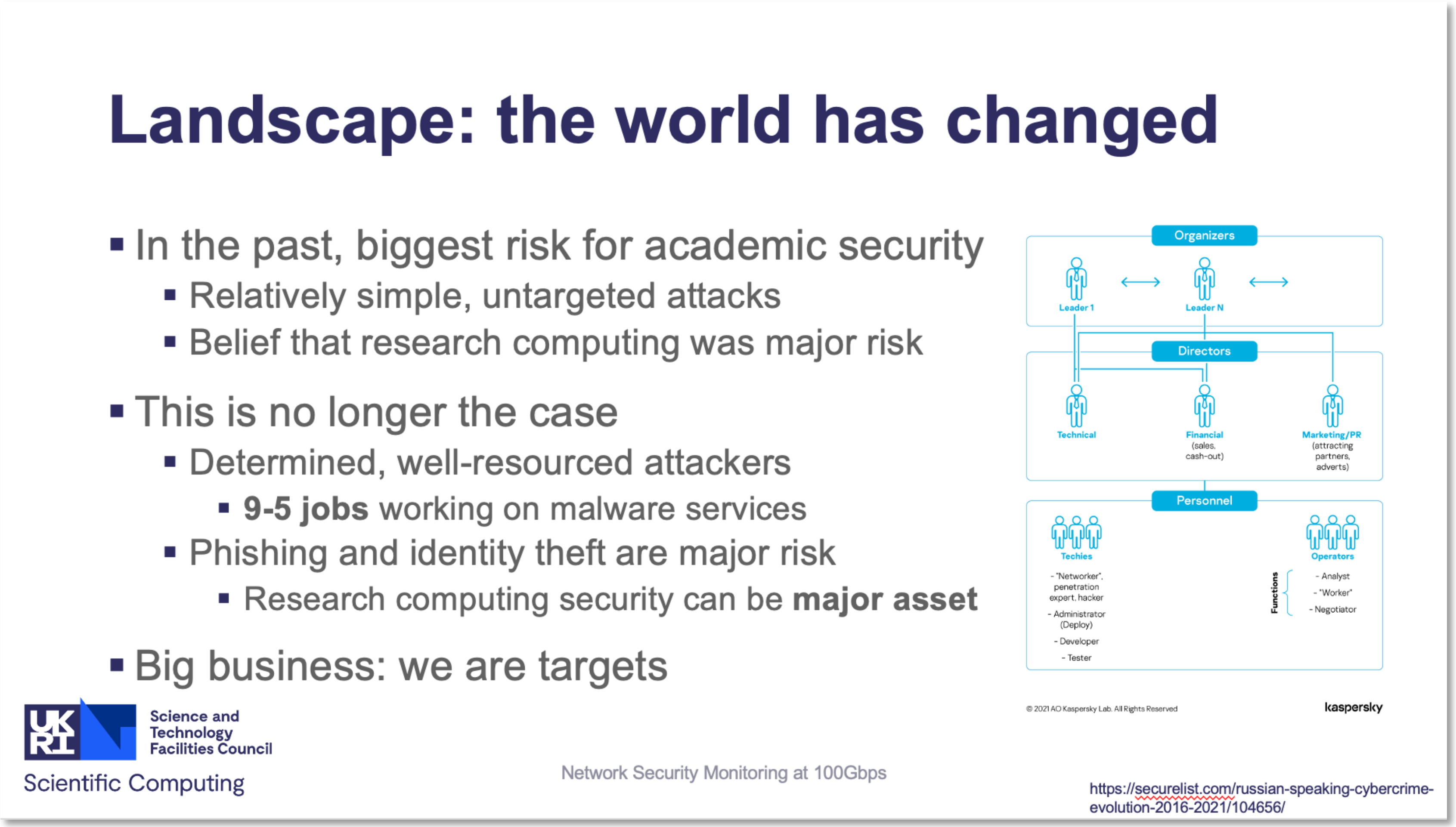


Intrusion detection with SOCs

- We need to match our monitoring capabilities and methodology to our circumstance

- Following our architecture: what are our threats, how do we defend ourselves?

Landscape



Landscape: the world has changed

- In the past, biggest risk for academic security
  - Relatively simple, untargeted attacks
  - Belief that research computing was major risk
- This is no longer the case
  - Determined, well-resourced attackers
    - **9-5 jobs** working on malware services
  - Phishing and identity theft are major risk
    - Research computing security can be **major asset**
- Big business: we are targets

Science and Technology Facilities Council
Scientific Computing

https://securelist.com/russian-speaking-cybercrime-evolution-2016-2021/104656/

Impact

- In our community (research and education) we are faced by determined attackers

- The impact of successful attacks can be **catastrophic**

- **Months** of site/facility downtime

- Major reputational and financial damage

The approach

- During incident response, we generate useful Indicators of Compromise (IoCs)
- Give a fingerprint by which to identify malicious traffic and your or another site
- We **must** share this information

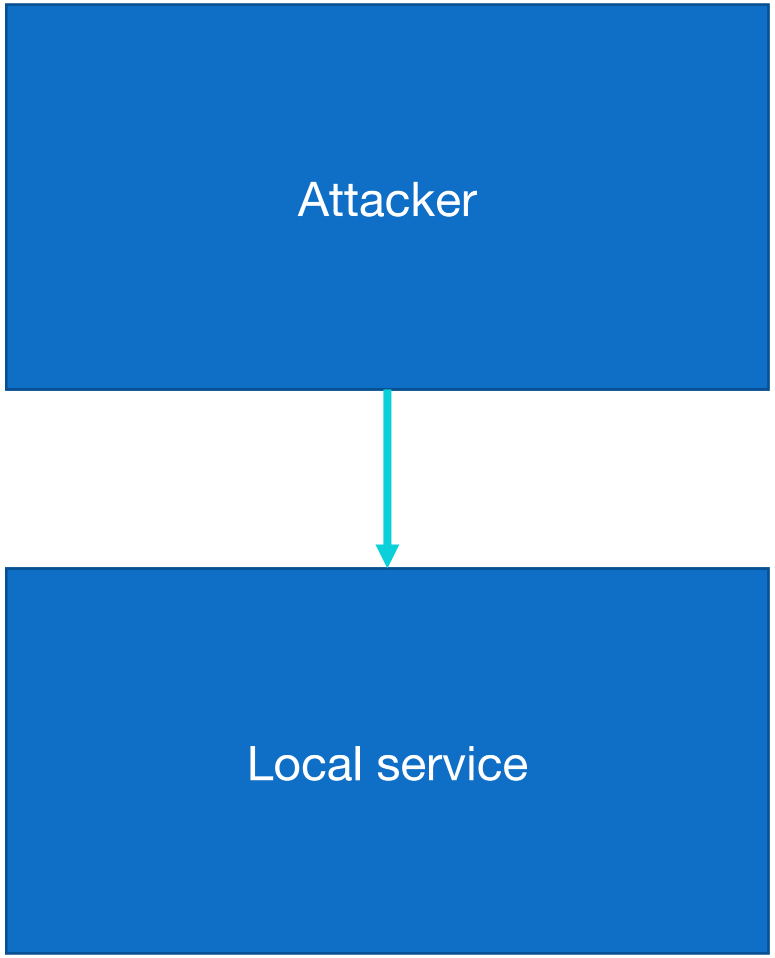
Threat intelligence

- Threat intelligence is the collection of these IoCs in a way that can help identify an attack

- It **does not** include specific information about your facility or service

Local vs attacker evidence

- Let's imagine that your Drupal CMS has been compromised via a recent unpatched vulnerability
- You're doing incident response and have a lot of information about the impact on your services
- You have some information on where the attacker came from and what actions they took on your network


Attacker

Local service

What information to share?

- The information that is useful to others are the **IoCs that identify the attacker**

- **Not** the impact on your service

- “The attacker’s IP was…”

  vs

- “My Drupal with all my group information was hacked and it’s a disaster!”

Sharing threat intelligence

- Sharing information this way means you are giving others the most important information

- **Without** giving away sensitive information
  - not in the data protection sense here

Type of IoCs

- Network
  - IP
  - Port
  - Timestamps

- Files
  - Checksums

- TTP information
  - Tactics Techniques Procedures

Who to share with

- Build trust groups
- Share with others that are similar to you
  - What is useful to me?
  - What is useful to them?
- Make the information as useful as possible

What makes good intelligence?

- Accuracy
- Timeliness
- Relevance

- Bulk lists of IPs are less useful than
  - I saw this set of indicators in active use today and these are developing
  - I saw evidence that X/Y/Z may be affected right now

Traffic Light Protocol (TLP)

- TLP is a set of 4 designations

- Designed to indicate the conditions under which information can be shared

- And with which audience

Traffic Light Protocol (TLP)

RED	Not for disclosure, restricted to participants only.	Recipients may not share TLP:RED information with any parties outside of the specific exchange, meeting, or conversation in which it was originally disclosed. In the context of a meeting, for example, TLP:RED information is limited to those present at the meeting.
AMBER	Limited disclosure, restricted to participants' organisations.	Recipients may only share TLP:AMBER information with members of their own organisation, and with clients or customers who need to know the information to protect themselves or prevent further harm. **Sources are at liberty to specify additional intended limits of the sharing: these must be adhered to.**
GREEN	Limited disclosure, restricted to the community.	Recipients may share TLP:GREEN information with peers and partner organisations within their sector or community, but not via publicly accessible channels
WHITE	Disclosure is not limited	Subject to standard copyright rules, TLP:WHITE information may be distributed without restriction.

https://confluence.egi.eu/display/EGIG/Traffic+Light+Protocol

TLP:AMBER

- For TLP:AMBER we can and **should** specify any specific restrictions
    - Only for security teams
    - Only for **this** security team, but all members of it

TLP Examples

Example	Category
Information about a vulnerability which impacts our community badly, but is not (yet) public knowledge	
I met my colleague and they had very timely information that would have an extremely high impact if it were to be generally available	
I have information that is timely and relevant about an ongoing incident that would be useful to my fellow incident responders	
I read about a critical vulnerability on The Register and $GIANTPLATFORM is impacted!	

TLP Examples

Example	Category
Information about a vulnerability which impacts our community badly, but is not (yet) public knowledge	TLP: GREEN
I met my colleague and they had very timely information that would have an extremely high impact if it were to be generally available	TLP: RED
I have information that is timely and relevant about an ongoing incident that would be useful to my fellow incident responders	TLP: AMBER
I read about a critical vulnerability on The Register and $GIANTPLATFORM is impacted!	TLP: WHITE

Data classification over time

- When determining which designation to use, what are the circumstances under which it will change?
  - We will tell you
  - After two weeks
  - …

- Specificity is at the heart of all good communication

Chatham House Rule

> “
>
> When a meeting, or part thereof, is held under the Chatham House Rule, participants are free to use the information received, but neither the identity nor the affiliation of the speaker(s), nor that of any other participant, may be revealed.

https://www.chathamhouse.org/about-us/chatham-house-rule

Threat Intelligence technology

- OK, we now have

- **Intelligence**
  - That is timely and relevant

- And we know

- **Who we want to share with**
  - Under what restrictions

Threat Intelligence technology

- How best to share this?

- Word of mouth
- Email
- …

- Specific service

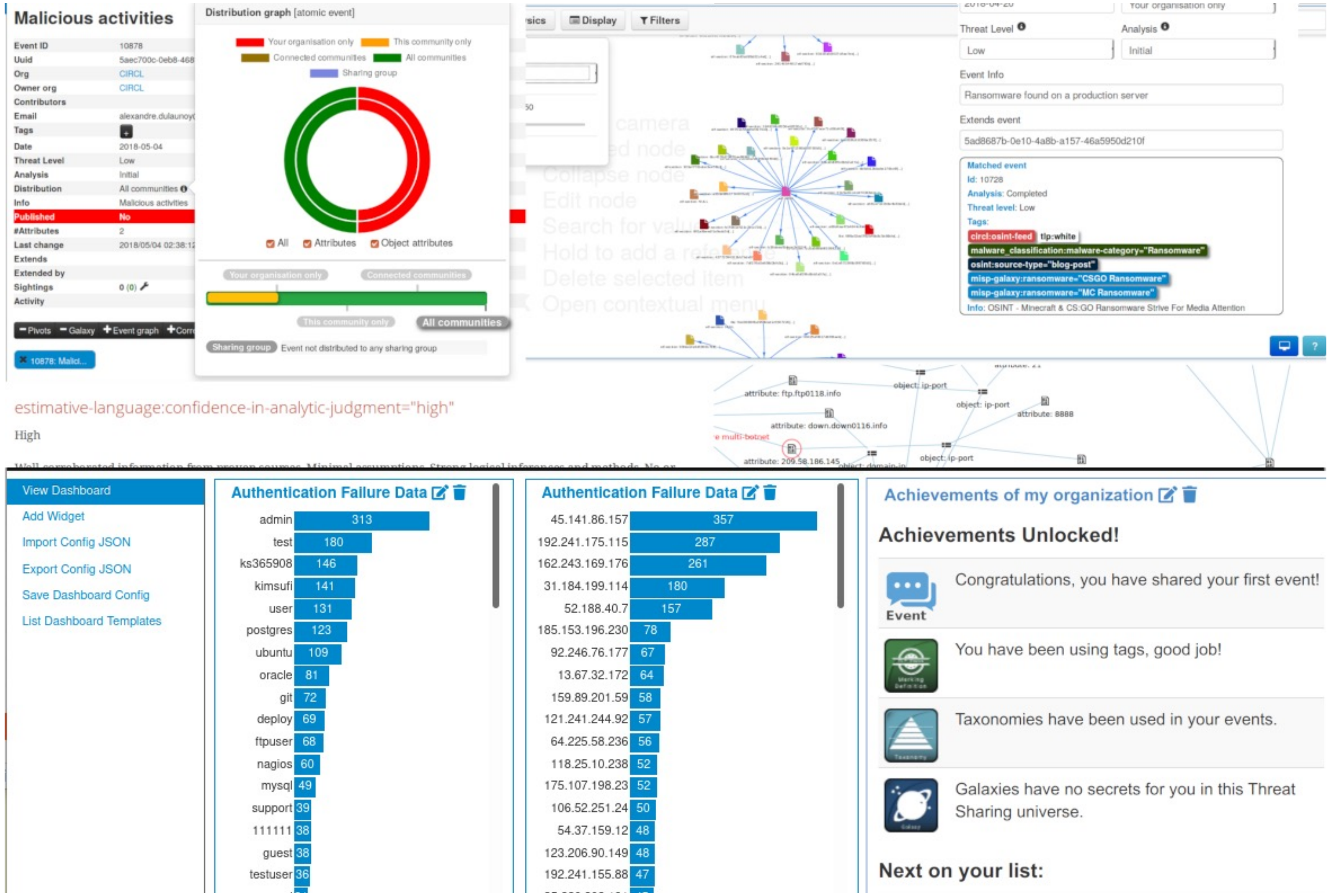
MISP

- Previously **Malware Information Sharing Platform**
- Incredibly flexible threat intelligence sharing tool developed by CIRCL.LU
- Web application with API

MISP
Threat Sharing

https://www.misp-project.org

MISP



Malicious activities

Event ID	10878
Uuid	5aec700c-0eb8-468
Org	CIRCL
Owner org	CIRCL
Contributors	
Email	alexandre.dulaunoy
Tags	+
Date	2018-05-04
Threat Level	Low
Analysis	Initial
Distribution	All communities
Info	Malicious activities
Published	No
#Attributes	2
Last change	2018/05/04 02:38:12
Extends	
Extended by	
Sightings	0 (0)
Activity	

−Pivots −Galaxy +Event graph +Corr

10878: Malici...

Distribution graph [atomic event]

🟥 Your organisation only 🟧 This community only 🟫 Connected communities 🟩 All communities 🟦 Sharing group

☑ All ☑ Attributes ☑ Object attributes

Your organisation only · Connected communities · This community only · All communities

Sharing group: Event not distributed to any sharing group

estimative-language:confidence-in-analytic-judgment="high"

High

Threat Level: Low

Analysis: Initial

Event Info: Ransomware found on a production server

Extends event: 5ad8687b-0e10-4a8b-a157-46a5950d210f

Matched event
Id: 10728
Analysis: Completed
Threat level: Low
Tags:
circl:osint-feed tlp:white
malware_classification:malware-category="Ransomware"
osint:source-type="blog-post"
misp-galaxy:ransomware="CSGO Ransomware"
misp-galaxy:ransomware="MC Ransomware"
Info: OSINT - Minecraft & CS:GO Ransomware Strive For Media Attention

View Dashboard
Add Widget
Import Config JSON
Export Config JSON
Save Dashboard Config
List Dashboard Templates

Authentication Failure Data

admin	313
test	180
ks365908	146
kimsufi	141
user	131
postgres	123
ubuntu	109
oracle	81
git	72
deploy	69
ftpuser	68
nagios	60
mysql	49
support	39
111111	38
guest	38
testuser	36

Authentication Failure Data

45.141.86.157	357
192.241.175.115	287
162.243.169.176	261
31.184.199.114	180
52.188.40.7	157
185.153.196.230	78
92.246.76.177	67
13.67.32.172	64
159.89.201.59	58
121.241.244.92	57
64.225.58.236	56
118.25.10.238	52
175.107.198.23	52
106.52.251.24	50
54.37.159.12	48
123.206.90.149	48
192.241.155.88	47

Achievements of my organization

Achievements Unlocked!

Event — Congratulations, you have shared your first event!

You have been using tags, good job!

Taxonomies have been used in your events.

Galaxies have no secrets for you in this Threat Sharing universe.

Next on your list:

https://www.misp-project.org

MISP

- Technical expression of trust
- Share information within a pre-defined set of sites / other MISP instances
    - Tags/comments/…

MISP

- One of the most important tools we are using and will use

- Genuinely broad usage across gov/commerce/academia

- Lots of training and documentation is available

R&E threat intelligence + EGI CSIRT

- R&E threat intelligence instance hosted by CERN
- Grew from activity for WLCG, available to the sector
- Either sync or use API

R&E threat intelligence + EGI CSIRT

- EGI CSIRT currently distributes IoCs via broadcasts to our sites

- Now working on incorporating threat intelligence sharing directly into our procedures

- Highly relevant intelligence on ongoing incidents to our scope

Security Operations Centres

- We have a great source of intelligence: what now?

- We need to understand what is happening in our service/facility/network
    - Host/network logging

- Let’s integrate these

Security Operations Centres

- From a technology standpoint, a SOC is the combination of
  - threat intelligence
  - fine-grained logging information
  - storage and visualization
  - alerting

Security Operations Centres

- From a high level, however, a SOC is the combination of
    - Technology
    - People
    - Processes

- Developing the team that uses a SOC and develops good information from it

SOC roles

- Key roles
  - SOC Service Manager
    - Deployment/Maintenance
  - SOC Analysts
    - Making sense of the data
  - Incident Responders

- These roles can spread across several people!

Security Operations Centres

- Developing the processes by which you disseminate and coordinate the alerting from the SOC

- Are **equally important to the tooling**

Teams and Processes

- Who maintains the SOC?
  - Next year?
- Where does the next tranche of hardware come from?
- Who analyses the alerts?
- …

How to deploy a SOC

- First question: what is the scope?
  - Individual batch farm?
  - Single site organisation?
  - Multi-site organisation?
  - Country?

- What considerations might come into play?
  - Effectiveness of intelligence
  - Network logging

How to deploy a SOC

Example	Deployment
Individual batch farm	Not clear that intelligence will be most useful
Single-site organisation	Identify network choke points
Multi-site organisation	How do we ship data around?
Country	Can’t use DPI for the backbone of a country

How to deploy a SOC

- Understand what scope you need to cover
- What outcome do you want?
- What logging capabilities do you already have?
- What staffing is available to you?

- **Start small enough to be useful**
  - MVP (minimum viable product)

Considerations

- Important to identify a realistic starting point

- Your capabilities with the tools will grow with experience

- Want to make your processes effective rather than throw hardware at the problem
  - You do need some of that!

Specific questions

- Where are the network choke points that are most relevant?

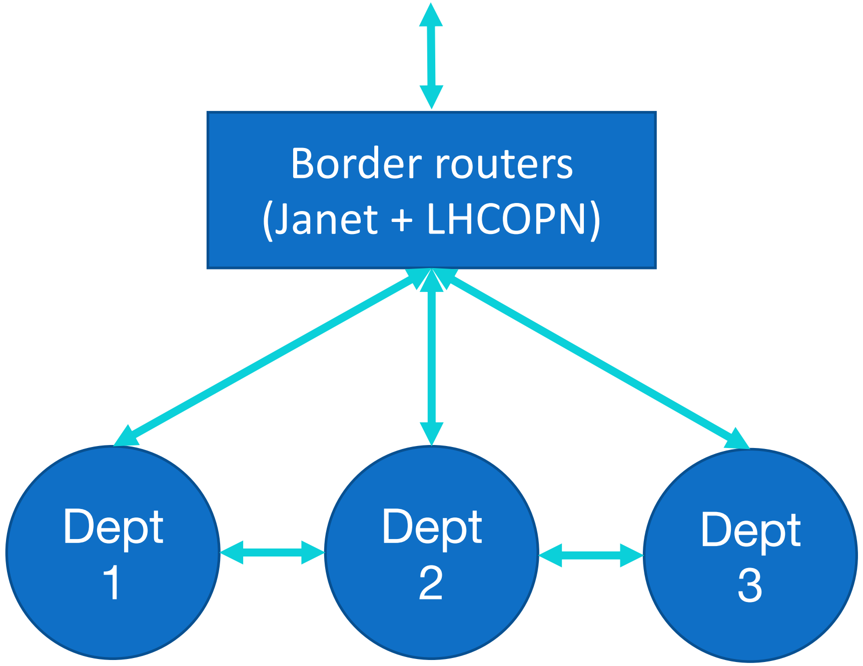
Example

- STFC is a multi-site organisation and we are deploying a SOC against the RAL campus

Specific questions

Example

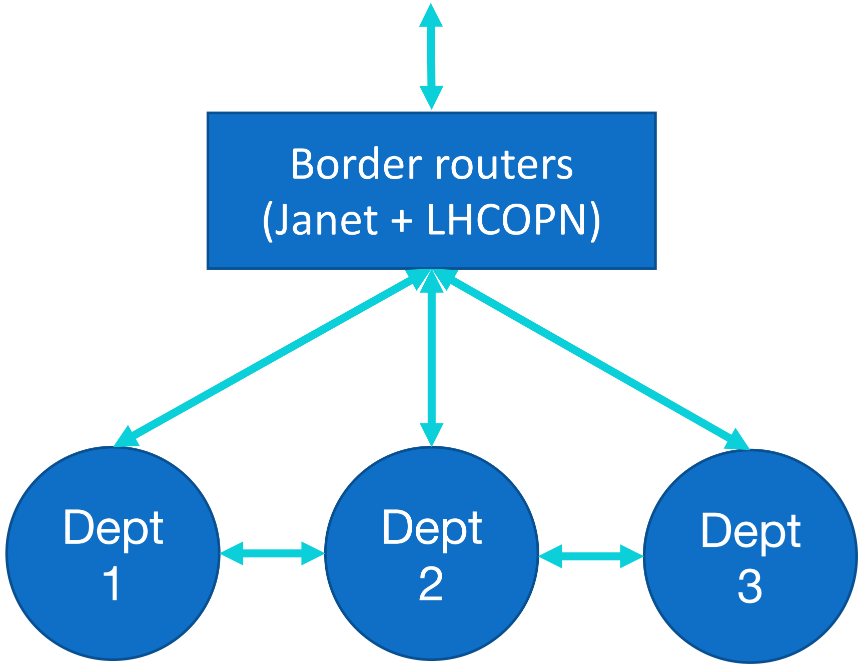
- STFC is a multi-site organization and we are deploying a SOC against the RAL campus



Specific questions

Example

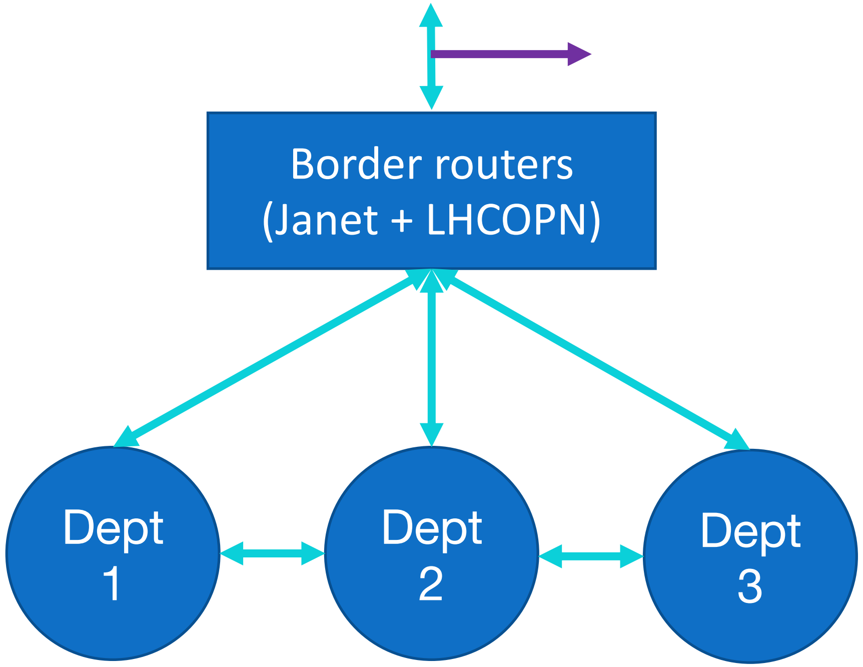
- Where do we put the network tap?



Specific questions

Example

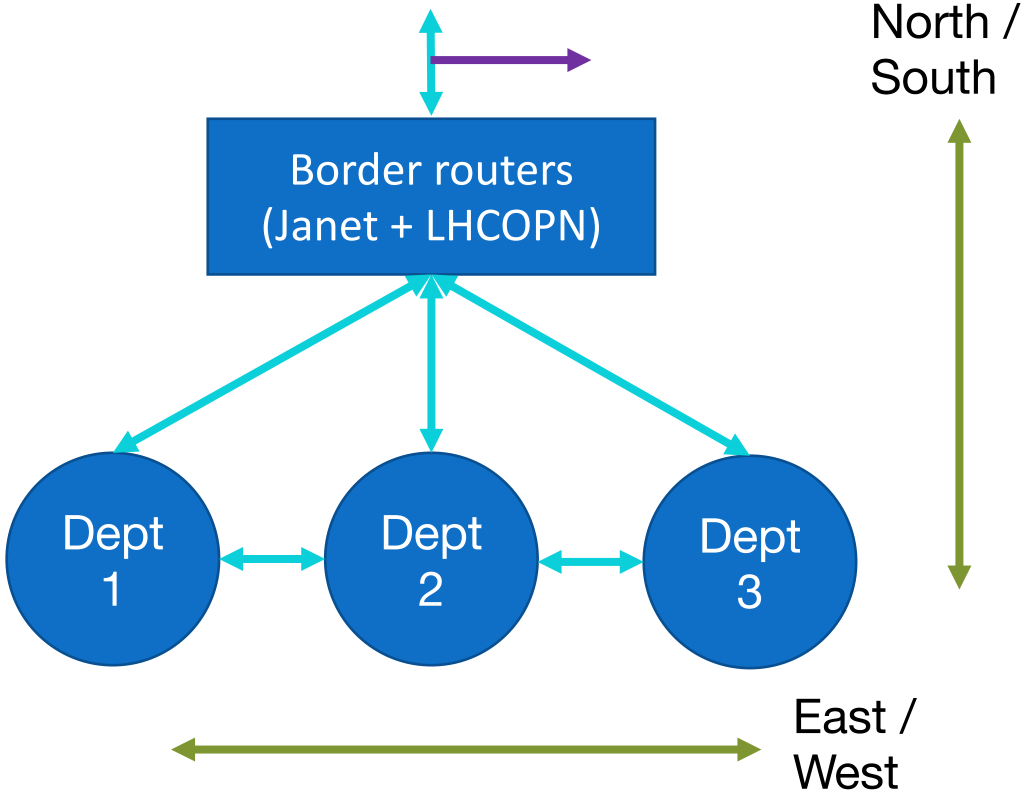
- Where do we put the network tap?



Specific questions

Example

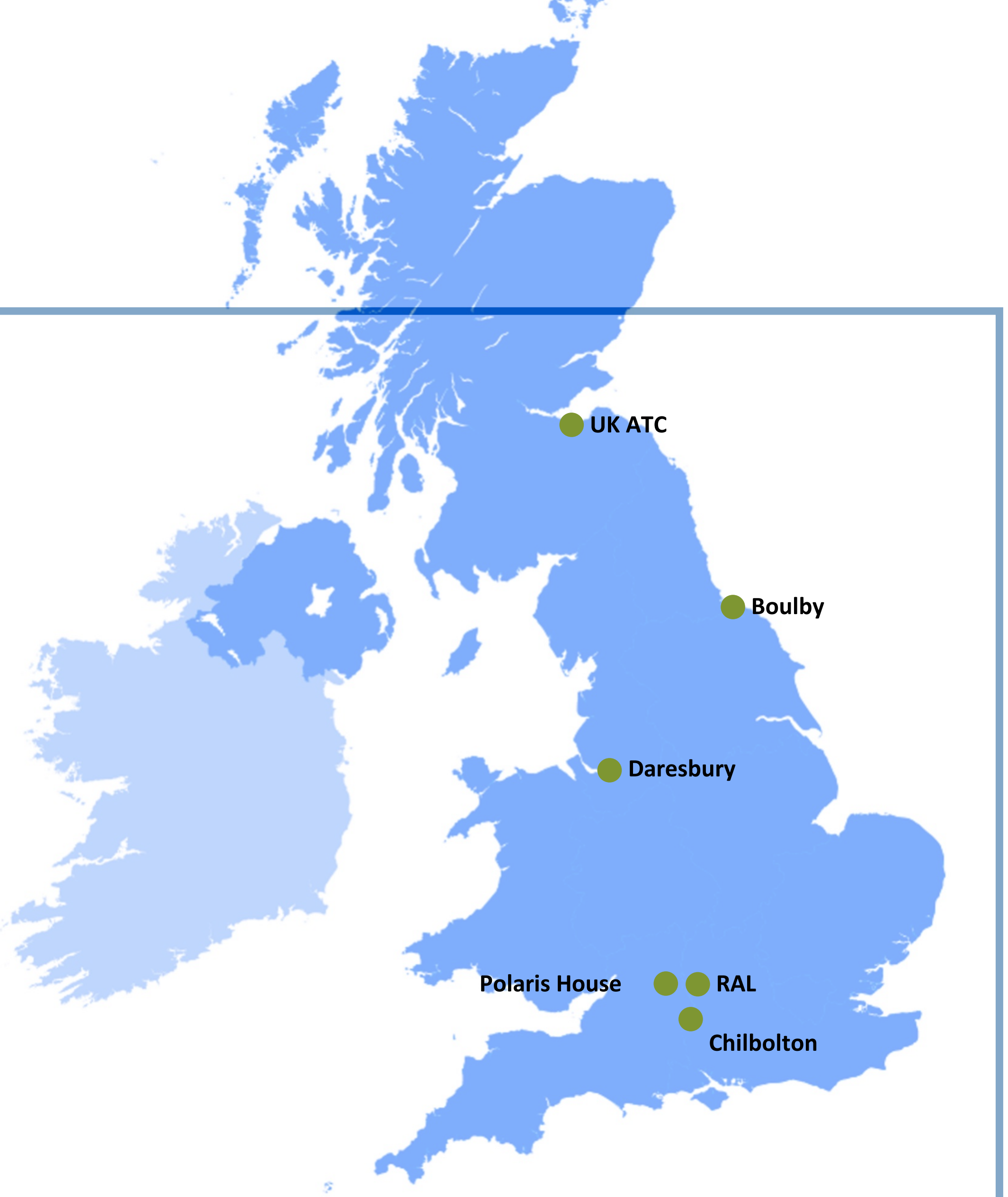
- North South traffic
  - Into and out of a site
- East West traffic
  - Traffic within a site



Specific questions

Example

- STFC Multi-site
  - What could our approach be?



SOC Components

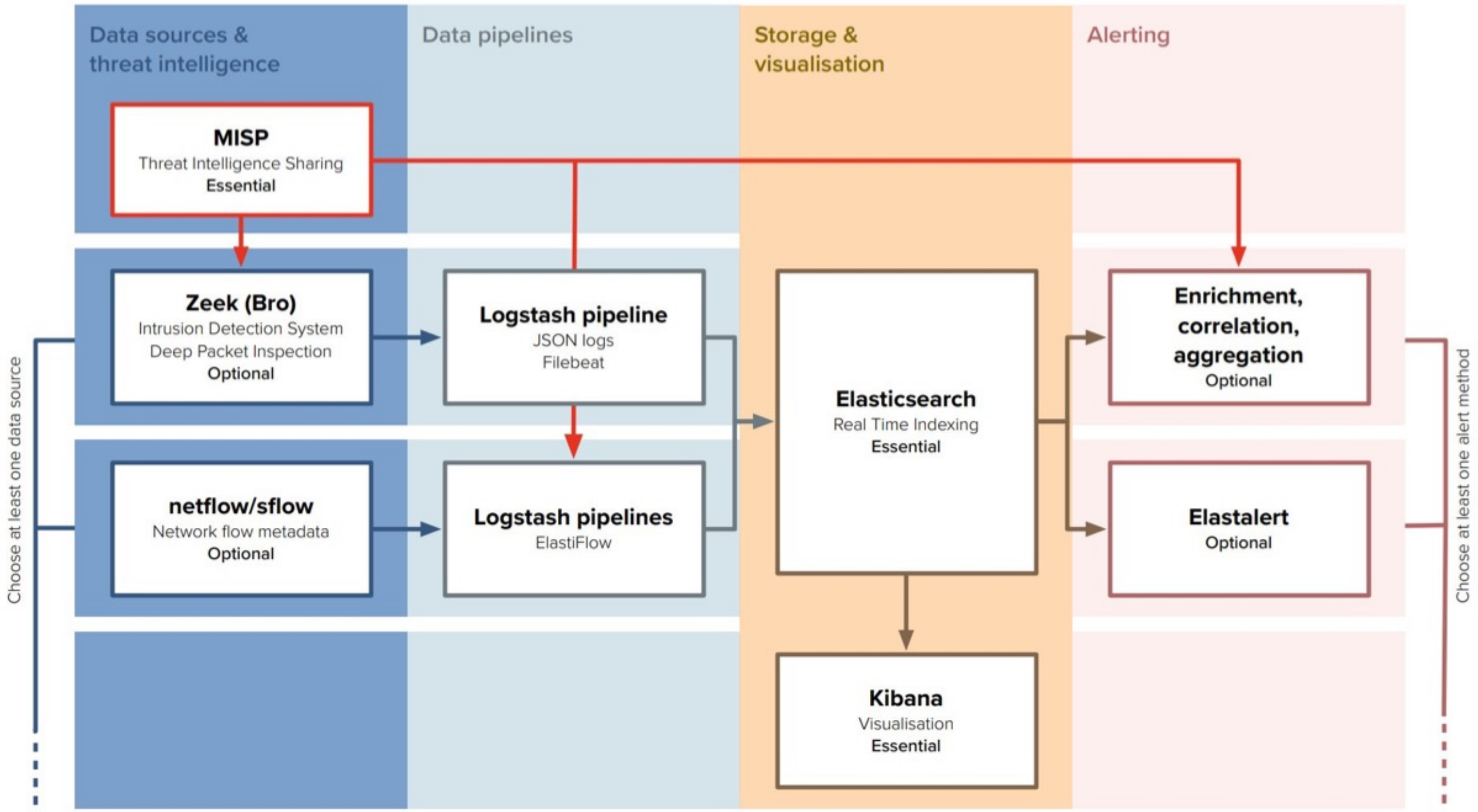
- Talked about some of the key components
    - Threat intelligence
    - Fine-grained network monitoring

- Let’s look at an overall structural diagram

SOC Components

- **NOTE**: this is the reference design created by the SOC WG
  - Coordinated by WLCG but open to R&E
  - Not the only way of going forward
  - Contains the necessary core elements

SOC Components



Data sources and threat intelligence

- Already discussed
  - MISP: threat intelligence
  - Zeek: network monitoring
  - Net/sflow: network monitoring
  - +host logs

- Start with one and grow from there

Data pipelines

- Logstash works as part of the standard elastic stack
    - Starting point

BUT

- Is typically not performant enough at high load
    - Kafka, …

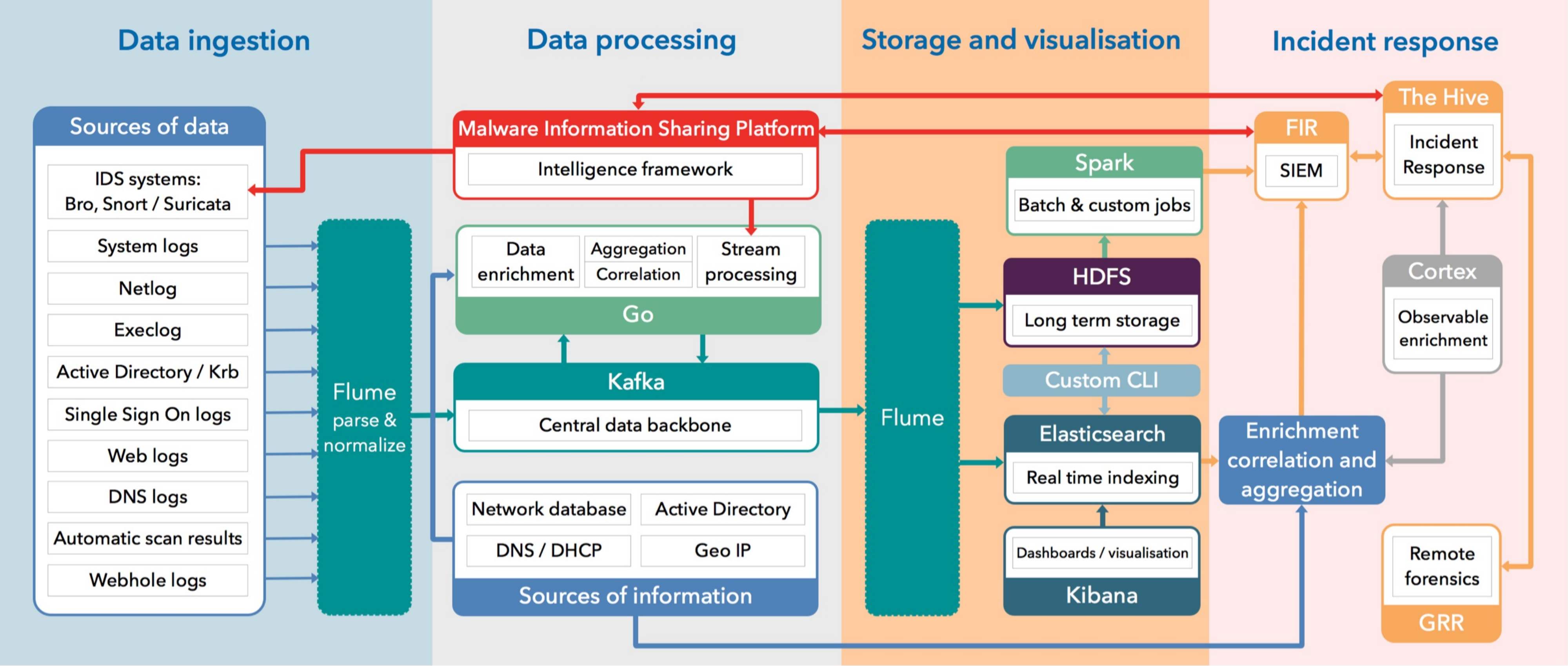
Storage and visualisation

- Elasticsearch + Kibana
  - Common, well understood components
- CERN has Elasticsearch service
- OpenSearch is a useful new distribution
  - Includes security plugins from the outset

Alerting

- Alerting directly from Zeek (see this during the week)
- Alerting from elasticsearch
- Aggregation of information into emails
  - In use in CERN SOC

CERN SOC



Data ingestion
Data processing
Storage and visualisation
Incident response
Sources of data
IDS systems: Bro, Snort / Suricata
System logs
Netlog
Execlog
Active Directory / Krb
Single Sign On logs
Web logs
DNS logs
Automatic scan results
Webhole logs
Flume
parse & normalize
Malware Information Sharing Platform
Intelligence framework
Data enrichment
Aggregation
Correlation
Stream processing
Go
Kafka
Central data backbone
Network database
Active Directory
DNS / DHCP
Geo IP
Sources of information
Flume
Spark
Batch & custom jobs
HDFS
Long term storage
Custom CLI
Elasticsearch
Real time indexing
Dashboards / visualisation
Kibana
The Hive
Incident Response
FIR
SIEM
Cortex
Observable enrichment
Enrichment correlation and aggregation
Remote forensics
GRR

Conclusions: Detection

- This afternoon we’ve looked at
  - The basics of logging, and logging technologies
  - The importance of identifying the most useful logs to avoid “data as noise”
  - The difference between flow based and deep packet inspection network monitoring

Conclusions: Detection

- We’ve also discussed
  - The importance of sharing threat intelligence for our community
  - Tools to help share intelligence responsibly
  - The MISP platform