GridPP
UK Computing for Particle Physics

WLCG SOC WG

Ian Neilson
GridPP Security Officer

Background

- Set up at March GDB after WLCG Workshop, Lisbon 2016

- Jointly chaired -
  - David Crooks (Glasgow) and Liviu Valsan (CERN)
  - **acknowledged for (almost) all slide content that follows !**
    - http://indico.cern.ch/event/394782/
    - https://indico.cern.ch/event/394831/

- Wikipedia SOC
  - An information security operations center (ISOC) is a dedicated site where enterprise information systems (web sites, applications, databases, data centers and servers, networks, desktops and other endpoints) are monitored, assessed, and defended.

Context

- A need for external observability of systems & networks
  - Increasingly opaque execution environments
    - VMs, Containers
- Increasing amount of security monitoring data being produced
  - + reductions in manpower to cope with this
- Identify tools available to WLCG sites of different sizes and provide appropriate guidance
- Leverage data analytics and Big Data frameworks used within our communities to provide security alerting, traceability and forensics information

Scope I

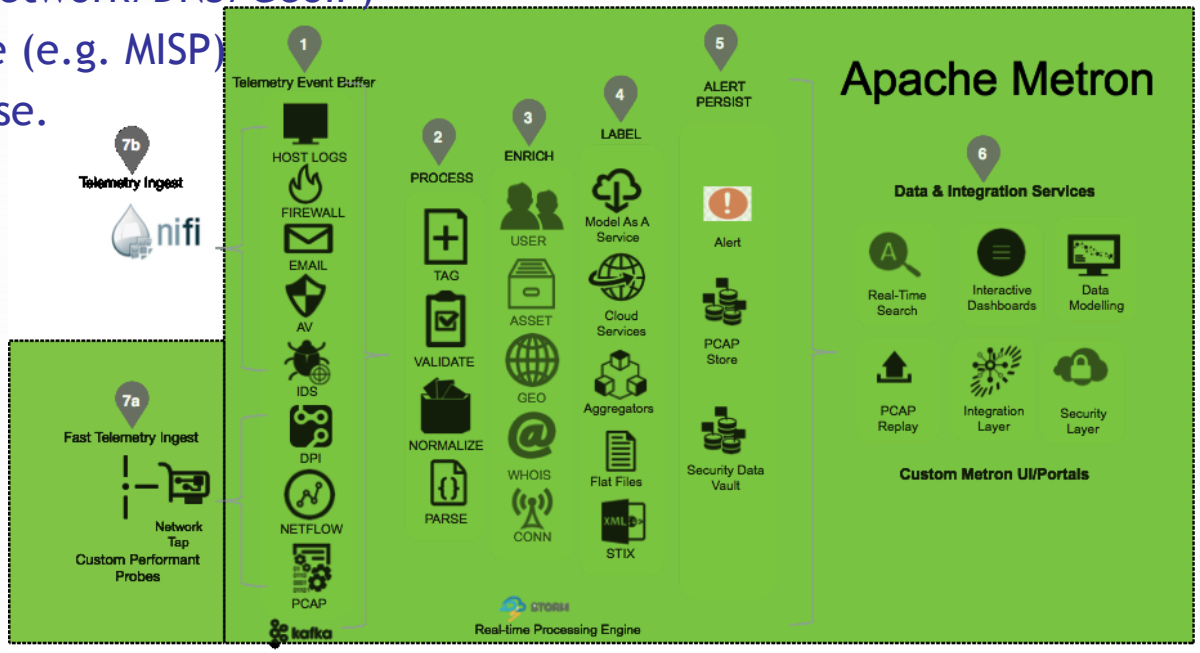
- Identify data that SOCs can / should provide
  - Both in terms of sources from which the SOC ingests data as well as necessary outputs
- Identify necessary components of a SOC for typical WLCG sites of different sizes
  - Recognising that local needs will be likely to vary
- Reference designs for SOCs of different sizes which could include installation guidance or appliances

Scope II

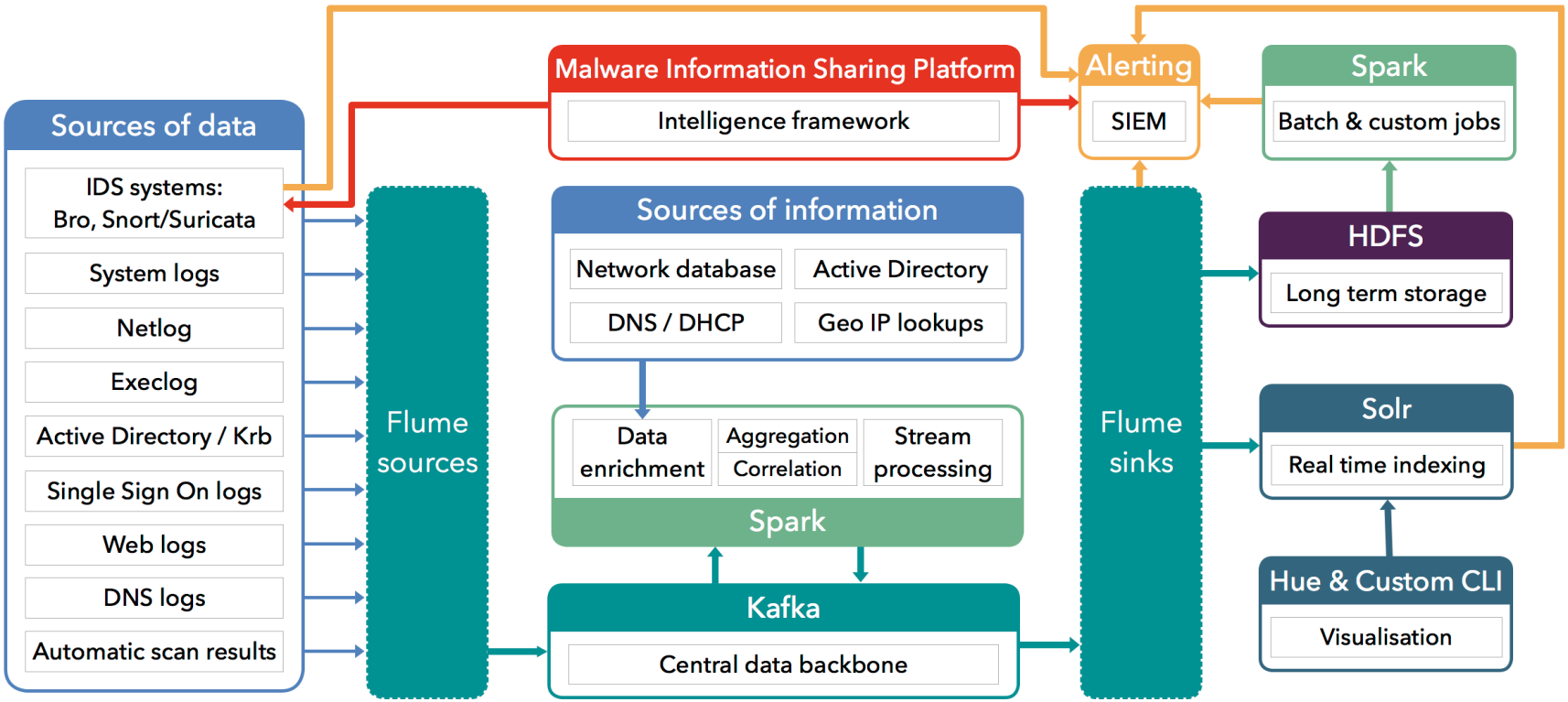
- Identify key stakeholders to be considered in the deployment of a typical SOC, including but not be limited to:
  - Local sysadmins
  - Local security teams
  - Campus security teams
  - NGI security teams/officers
  - VO Security teams
- Data protection / privacy and information sharing policies
- Timeframe for delivery (differentiated between outcomes)

CERN SOC

- CERN quite far advanced with custom campus SOC
  - Similar architecture/flow to OpenSOC -> now Apache Metron
- Ingest monitoring data
- Enrich with information (network/DNS/GeoIP)
- Correlate with intelligence (e.g. MISP)
- Store, index, alert, visualise.



CERN SOC


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
Flume sources
Malware Information Sharing Platform
Intelligence framework
Sources of information
Network database
Active Directory
DNS / DHCP
Geo IP lookups
Data enrichment
Aggregation
Correlation
Stream processing
Spark
Kafka
Central data backbone
Alerting
SIEM
Spark
Batch & custom jobs
HDFS
Long term storage
Flume sinks
Solr
Real time indexing
Hue & Custom CLI
Visualisation

Progress and Challenges

- Metron and CERN SOCs
  - Large, complex (=? fragile) stacks
  - Applicable at campus level
  - Can components be used in isolation for smaller sites?
  - Can a SOC-in-a-box be created for easy deployment and use with less manpower?
- WG starting “at each end”
  - Bro IDS and MISP could be used as minimal framework
  - Create a small peering of MISP instances (currently UK NGI-based)
    - Test instances running at Glasgow and RAL
    - To understand the tool and how it might be used
    - Challenge is probably not the tool but the human trust networks.
  - Configure Bro instances to gain experience

Thank You.