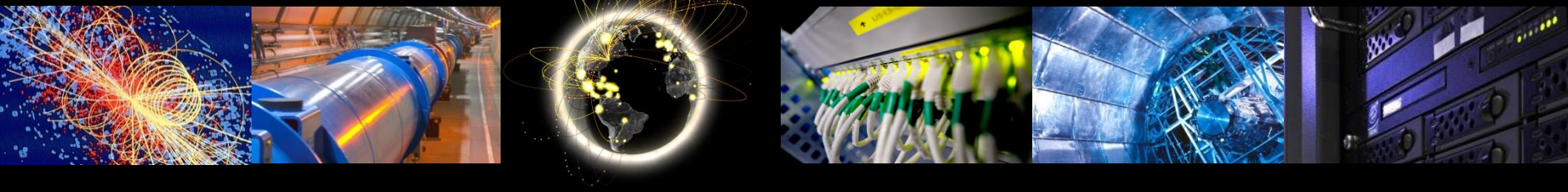
Evolving Security in WLCG

Ian Collier, STFC Rutherford Appleton Laboratory

1st February 2016, WLCG Workshop Lisbon



WLCG
Worldwide LHC Computing Grid

CERN

Overview

- Current model
- Changing technology
- Evolving trust fabric
- Changing threat landscape
- Possible ways forward

Current model

- Traceability not control
- Relationship between sites and VOs and users is based on trust
  - This works pretty well
- Incident response based on sites & collaboration CSIRT teams
- All traceability information (in theory) at sites
  - Central loggers
  - But in practice, must contact VO to identify so be able to suspend credential associated with problematic activity
- Anonymous pilot jobs but separation & traceability supported by glexec – again *in theory*
  - After 10 years *still* not used universally
  - Not really 'loved' by either sites or VOs

Changing technology

- Increasing use of virtualisation and containers
  - VMs on cloud platforms
  - Containers within 'traditional' batch systems
  - Who maintains the VMs/containers?
    - Should strive for best management – tools to streamline
- Offers alternate route to job separation
- Removes direct access to & trust of execution environment.
  - May no longer be able to trust logs
- Makes maintenance of underlying OS easier for sites
  - But they pick up complexity of cloud management frameworks
    - On plus side these have much larger communities behind them than grid software.
- Technology & VO workflow changes create constant pressure on incident response teams
  - Emergence of cloud technologies a particular challenge

Evolving trust fabric

- Federated identity management promises huge potential benefits – as well as bringing with it not a few challenges
- Will take significant changes at all levels
- But mixing assurance from different sources (not just CAs) will bring benefits
  - Not least making it easier to co-exist in a world where WLCG is one among many large users of distributed infrastructure
- Eduroam example is instructive
  - The benefits are now obvious - it is *really* convenient
  - But it has been quite a bumpy ride

Changing threat landscape

- Rise of organised, very businesslike cybercriminals
  - They no longer ignore us
  - Sophisticated, targeted attacks – especially phishing
- Identity/personal information is now the major target
  - Federation just increases the cost of compromised credentials
- Our infrastructure itself may be 'secure enough'
  - The challenge now to protect our people
- As we move to more standard software & interfaces attack surface is also more standard
  - Much of the effort that used to go into making our bespoke software more secure will be needed protecting 'standard' software & interfaces
- Most incidents are discovered through external reports
  - Must improve our ability to exchange intelligence with other communities (industry, law enforcement, etc.)

Ways forward I

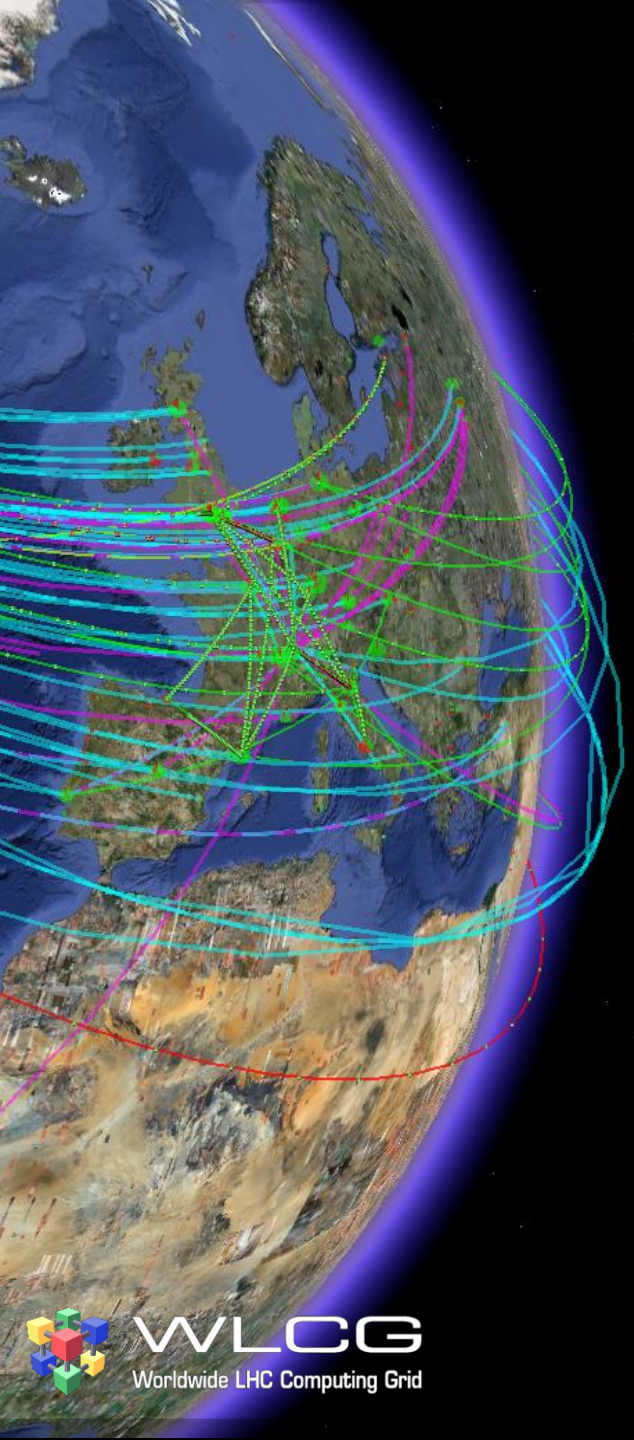
- Treat VMs/containers as processes
  - Shift focus to externally observable behaviour
    - Logs from inside VMs not as trustworthy
  - Better logging of network flows – often neglected – may have implications for network hardware choices and costs
- 'Big data' tools for storing, aggregating, searching larger volumes of data
  - Security Operations Centre
  - Significant effort to deploy
  - Can we develop an 'appliance' for this (similar to what Perfsonar does for network monitoring)
- Can we then 'forget' about glexec?

Ways forward II

- Bring VOs more fully into incident response process
- Improve the capability for collaboration traceability?
  - Better intrumenting VO frameworks to (centrally?) log data
  - Can we take advantage of VOs being based at CERN to ingest appropriate traceability data directly into the SoC as it develops?
- More emphasis on protecting people in order to protect our infrastructure
  - Phishing, sharing threat & incident intelligence
- Put effort in to supporting & exploiting federated identity management
  - not forgetting impacts on traceability and incident response

Summary

- Separation via VMs/containers – drop glexec
- Invest in deploying ‘big data’ tools for managing traceability data
- Invest in better intelligence/trust links with other communities
- Embrace global & federated identity management



- Over to the others

WLCG
Worldwide LHC Computing Grid