21st International Conference on Computing in High Energy and Nuclear Physics **CHEP2015** Okinawa Japan: April 13 - 17, 2015

Contribution ID: **438** Type: **poster presentation**

Maintaining Traceability in an Evolving Distributed Computing Environment

The management of risk is fundamental to the operation of any distributed computing infrastructure. Identifying the cause of incidents is essential to prevent them from re-occurring. In addition, it is a goal to contain the impact of an incident while keeping services operational. For response to incidents to be acceptable this needs to be commensurate with the scale of the problem.

The minimum level of traceability for distributed computing infrastructure usage is to be able to identify the source of all actions (executables, file transfers, pilot jobs, portal jobs, etc) and the individual who initiated them. In addition, sufficiently fine-grained controls, such as blocking the originating user and monitoring to detect abnormal behaviour, are necessary for keeping services operational. It is essential to be able to understand the cause and to fix any problems before re-enabling access for the user.

The aim is to be able to answer the basic questions who, what, where, and when concerning any incident. This requires retaining all relevant information, including timestamps and the digital identity of the user, sufficient to identify, for each service instance, and for every security event including at least the following: connect, authenticate, authorize (including identity changes) and disconnect.

In traditional grid infrastructures (WLCG, EGI, OSG etc.) best practices and procedures for gathering and maintaining the information required to maintain traceability are well established. In particular, sites collect and store information required to ensure traceability of events at their sites.

With the increased use of virtualisation and private and public clouds for HEP workloads established procedures, which are unable to see 'inside' running virtual machines no longer capture all the information required. Maintaining traceability will at least involve a shift of responsibility from sites to Virtual Organisations (VOs) bringing with it new requirements for their logging infrastructures. VOs indeed need to fulfil a new operational role and become fully active participants in the incident response process.

We present an analysis of the changing requirements to maintain traceability for virtualised and cloud based workflows with particular reference to the work of the WLCG Traceability Working Group.

Primary author: COLLIER, Ian Peter (STFC - Rutherford Appleton Lab. (GB))

Co-author: Mr WARTEL, Romain (CERN)

Presenter: COLLIER, Ian Peter (STFC - Rutherford Appleton Lab. (GB))

Track Classification: Track7: Clouds and virtualization