Maintaining Traceability in an Evolving Distributed Computing Environment

I Collier[1], **R Wartel**[2]

[1] Science & Technology Facilities Council, Rutherford Appleton Laboratory, Harwell Oxford, DIDCOT OX11 0QX, UK

[2] CERN, CH-1211 Geneva 23, Switzerland

E-mail: ian.collier@stfc.ac.uk

Abstract. 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.

1. Introduction

Security incidents are an operational reality on all computing infrastructures. These incidents include, but are not limited to attempts to break in and hijack resources, denial of service attacks and misuse of resources by valid users.

When things go wrong we need, at a minimum, to trace the source of all actions (executables, file transfers, pilot jobs, portal jobs, etc.) and the individual who initiated them. In other words we need to be able to answer the questions *WHO* did *WHAT*, *WHEN* and *WHERE*. In turn we then require the granular controls which allow us to block current and future actions initiated by questionable credentials until investigation is complete.

Across the Worldwide LHC Computing Grid (WLCG) [1] and related distributed computing collaboration sites we are used to managing the risks inherent in running grid based distributed computing infrastructures. We have well established policies [2] and well developed incident response procedures which allow us to contain the impact of incidents, preserve reputations and ensure that resources are available for their intended purposes.

However, the increasing use of virtualization and cloud computing means that in order to maintain traceability the security policies, procedures and best practices will have to evolve. We will examine the issues in more detail, discuss possible solutions currently under active investigation within WLCG and consider what may still be required.

2. Current Situation

The current procedures and policy assume sites have direct sight, and control, of the worker node operating system. This allows sites to collect detailed logs from the execution environment as well as from the compute elements, batch systems and other grid middleware to central loggers.

An authorization & traceability framework is in place using glexec [3] which ensures that jobs from different originating users are run with distinct local identities (although this is not implemented universally) and Argus [4] which provides a granular authorization service. In addition Argus allows fast emergency suspension of possibly compromised credentials at the site, national and global collaboration level. These components in turn send detailed logs to site central loggers.

Together this means that all the information required to identify the origin of particular actions at a site, or resource provider is in principle available to the administrators and incident response teams at that site.

More or less sophisticated tools (currently often variants of grep) are used to search these logs when incidents occur. (It is worth noting that the management, aggregation, indexing and searching of large volume logging data using modern data mining and analytics tools is already a fertile area of investigation in the academic grid computing community. One example being work reported by the CMS experiment at CHEP 2015 [5])

Across our collaborations we have clearly identified contact points and well established trust relationships.

All of these combine to ensure that when incidents occur all the information is readily available to support both the analysis of and commensurate response to problematic activities.

3. The emergence of virtualization and cloud computing

The increasing use of private, public and federated cloud resources brings changes to many aspects of distributed computing, which in turn will require changes both to the information gathered, the tools used to manage that information and the policies and procedures governing both the collection and use of traceability information. Most significantly, all the necessary information related to one site or resource provider may no longer be contained at that site.

This brings significant changes to workflows, with traditional grid interfaces being replaced by cloud interfaces (EC2, OCCI) and federated cloud broker portals such as the EGI Federated Cloud. Some of these changes remove complexity for users (that is certainly the aim), and more obvious changes for providers who no longer simply maintain worker node operating systems and some grid interfaces, but will now also be running more or less complex cloud management frameworks plus the components to integrate with cloud federations. These new components are being actively deployed and developed, along with an understanding of best practice.

There are many new ways for things to go wrong.

Complex interactions between components in management frameworks (see for example Figure 1, the OpenStack architecture diagram below) plus orchestration layers which are being deployed (VAC, INDIGO etc.).

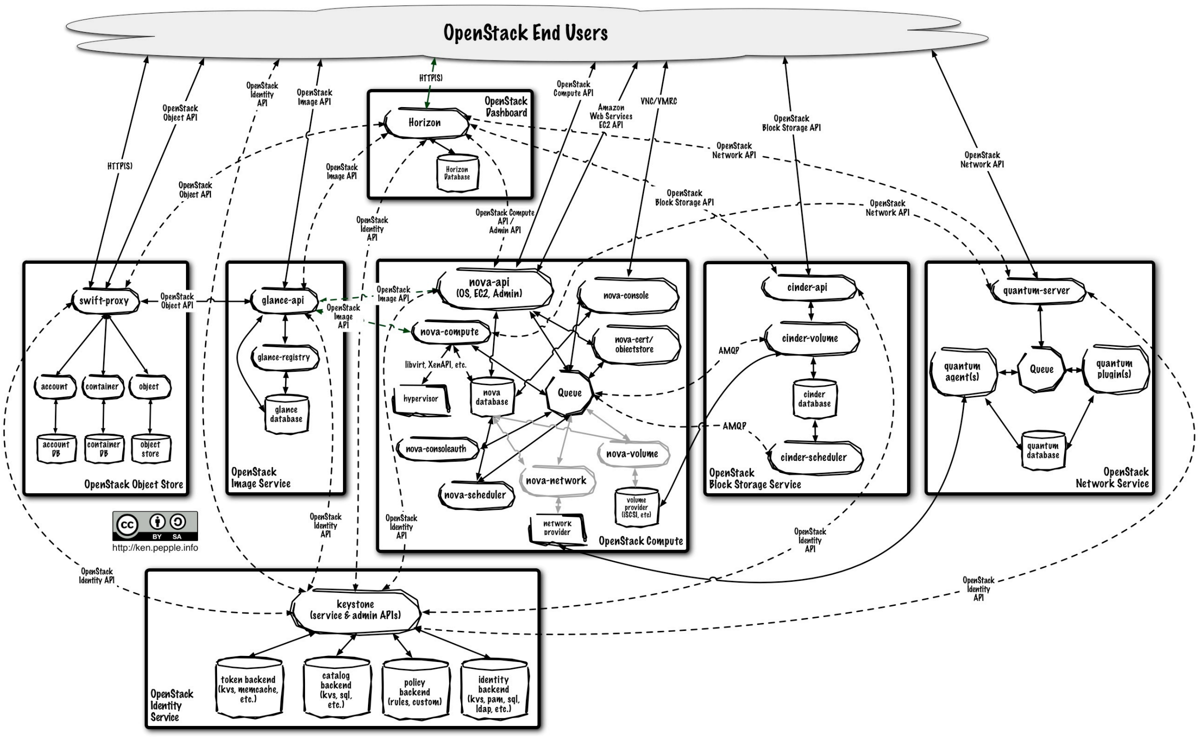

Figure 1: Logical architecture of the Folsom release of OpenStack [6]

Sites no longer have the same control of the execution environment. On the one hand this leaves resource providers with simpler and more easily secured hypervisor operating systems to manage, but on the other hand they cannot directly log activity from the execution environment, and would have to make particular arrangements with Virtual Organisations (VOs) to have root access to virtual machines (VMs) for incident response investigation. It will at the very least require the development of new trust relationships.

One comparison which is often made is with public cloud providers such as Amazon Web Services who ‘don’t care what goes on inside’, although in practice they will typically have large and well resourced security teams and they certainly will inspect VM images when they believe the risk justifies it. They also establish contractual relationships with users that place the responsibility for the financial and other costs of any security incidents with the user.

Rather than individual jobs being dispatched to batch systems at sites, we will now see VMs launched by VOs – or automatically by their workload management systems. In the absence of a mechanism analogous to glexec, sites cannot, on their own, know what user is responsible for activity in a particular VM, and consequently might have no choice but to suspend an entire VO in the event of problems. The logical extension of this would be suspending that VO globally until investigations are complete. This would cause problems for both VOs and resource providers.

As new workflows, frameworks and technologies emerge we should aim to match the traceability present in the best of current practice, not the worst. We should ensure we maintain the traceability we depend upon.

4. Investigations of possible solutions

WLCG has established a working group to carry out practical investigations of solutions to the challenges that have been identified. This section describes some of the areas addressed by this working group.

4.1. Logging

As already noted, it is essential that all services running at sites be configured to produce relevant logging information, addressing the traceability requirements discussed above. However, not all the information required is now available at the site itself. Information from the VO workflow management systems and from cloud federation frameworks may also be required in order to identify the origin of activity at a particular site.

As resource providers no longer have the same access to the execution environment itself, it will be necessary to increase focus on externally observable behaviour, in particular network flows to and from virtual machines. This will require new techniques and logging infrastructures for many sites. In addition, the cloud management frameworks themselves introduce new levels of complexity with many components and interactions between those components. (As illustrated previously by the OpenStack architectural diagram in Figure 1.)

Furthermore hypervisor and Cloud management frameworks and federated resource frameworks will all require careful configuration and monitoring.

Eventually it becomes essential to have complementary tools improving the logging of network activity and network flows. Examples would be netlog and execlog, linux kernel modules developed and maintained at CERN that allow granular logging of process and network calls [4]. These examples focus on linking a given network flow to the matching user or process. For many use cases, for instance a hypervisor sharing a network addresses between several VMs, or on multi-user systems, such tools provide crucial evidence during investigation. They enable the origin of a particular network activity to be accurately identified.

Once services are correctly configured, it will also be necessary to connect VMs to central loggers, which may requires standardised hooks to be provided in VM images.

In many cases, there is no single location containing all the details of a particular activity, and aggregation of and cross checking between multiple sources becomes vital. Investing resources in improved tools for storing, aggregating and searching is increasingly critical.

4.2. Traceability information infrastructure design

With emerging clouds and the ever growing size of computing infrastructures, the volume of logged information may very rapidly overwhelm traditional logging infrastructures.

In addition, the amount of threat indicators and malicious patterns shared within organisations has increased dramatically in the recent years. A growing field of computer security is threat intelligence, aiming at producing actionable data from existing experience and expertise in the community. International collaboration is a key aspect there.

Two essential aims of traceability include the capability to raise alerts rapidly when suspicious entries are detected and to search the details of past records for specific entries as part of incident response

Achieving these goals will require a specific traceability infrastructure, enabling both storing and indexing large amounts of information, as well as the ability to crawl live and existing data quickly.

Many organisations in the computing industry are faced with this challenge, and after several years of home grown solutions, there seems to be a convergence in the community to use a similar design and set of tools, drawing upon on the experience of big data management.

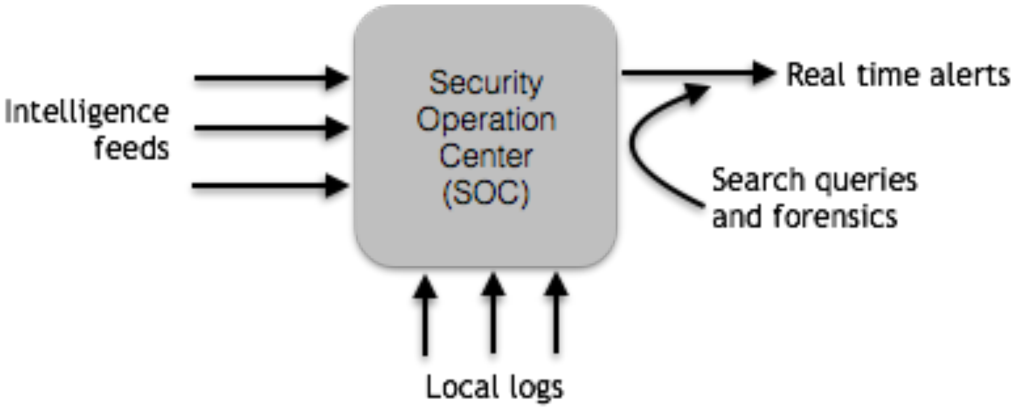

Figure 2. OpenSOC schematic diagram

A notable example is OpenSOC [7], which is built upon Elastic Search and Hadoop. Illustrated in Figure 2, OpenSOC aggregates logs from local systems, network flows if available, plus almost any other available information feeds and indexes them, thus facilitating sophisticated search queries and real time alerts. To do this it makes use of technologies developed in the wider 'big data' industries. A growing number of organisations go further and integrate their intrusion detection systems directly in such an infrastructure. One example of this is Bro IDS [8].

No matter what design or solution is being considered, it must be able to manage both the increased volumes of logging information and the complex interactions of information from different sources, sometimes across organisations and locations. But it is the analysis and data processing power which are likely to become the bottleneck requiring significant computational resources to be deployed simply to ensure the traceability we depend upon.

4.3. Quarantining Virtual Machine Images

Virtualisation does provide some significant advantages over traditional systems. During investigations, it allows VM images to be fully captured without damaging evidence for forensics purposes. At a minimum this means that if intercepted while problematic activity is taking place, resource administrators are readily able to secure copies of running machines. However, tracking all VMs is extremely challenging, in particular in cases where the attacker deliberately uses short lived VMs, or more generally when incidents are discovered after a VM has been terminated.

This risk may be mitigated by deferring deletion of VM images for a quarantine period after they have been terminated.

Some cloud management frameworks already provide this functionality as standard, one example is StratusLab [9] which has quarantining of VM images configurable by its storage management service. Deferred deletion, or quarantining, of images should be highly encouraged, as it can often take weeks to discover compromises. The storage management services associated with OpenStack and OpenNebula, however, do not by default have this feature. Implementation is not completely trivial without obstructing existing storage lifecycle management.

Members of the WLCG Cloud Traceability Working Group are actively investigating ways to achieve deferred deletion of images in OpenStack and OpenNebula using Ceph.

4.4. Roles and responsibilities

A high proportion of computer security incidents, at both academic and commercial infrastructures, involve more than one organisation.

Typically, each organisation is responsible for its own security, and organisations collaborate in a peer-to-peer fashion, with a varying degree of collaboration depending on the community and trust agreements in place. In the future, however, communities, or groups of organisations, may need to adjust this model to acknowledge the ever increasing complexity of attacks.

It will indeed soon become extremely difficult or perhaps even not economically viable for a single organisation to maintain sufficient expertise in-house, in order to deal with advanced criminal or government-based attacks. However, as a community, mutualising resources, efforts and expertise may provide an adequate response to these sophisticated threats. In such a configuration, each community would appoint a small number of computer security investigators. Then, each organisation would act as a "traceability provider", simply feeding trusted remote investigators with the adequate traceability information needed to resolve an ongoing security incident. This involves an appropriate trust model and procedures in order to identify and trust the investigator, and implement adequate mechanisms to share traceability information.

This long term transition is also motivated by the knowledge that, in more and more cases, no one administrative entity will own sufficient traceability information to fully resolve a security incident: it will become even more crucial for organisations to collaborate and share information to reconstruct the trail of activity of a particular user or job. This is particularly true for grid computing, where security teams must actively involve the VOs in the incident response process, in order to identity, and possibly suspend, the compromised or malicious users behind specific jobs.

This is usually done by simply re-using the detailed trail and job workflow used otherwise by the VOs for debugging and workload management purposes.

Whenever it is needed to conduct investigations, efforts needed and most important, overall effectiveness, depend critically upon the implementation details within the VO.

A very useful and efficient mean to identify limits of the current logging and suggest enhancements, is to run regular "security challenges". Payloads and challenge methodologies for "traceability service challenges" are currently in development, while the approach has been used for several years already.

Such drills constitute an excellent opportunity to formally recognise the existing reality that active participation of VOs is needed in order to maintain traceability, and to identify where they best fit in a modern incident response workflows.

Finally if sites no longer have any site of the originating identity of workloads, it may be that VO workflow management frameworks need to be integrated with the Argus authorization system, or something very like it, in order to support fast centralised suspension of credentials associated with inappropriate activity until investigations are complete.

4.5. Data Protection

Managing both traceability and privacy requirements is critical to any organisation. Traceability information would usually contain personal data, which should be handled and processed with due care.

Organisations or collaboration would typically have data protection policies, detailing how to ensure both legitimate interest of the user and effective incident response, in line with existing directives. This is in particular important for EU countries, where *Directive 95/46/EC* [10] enforces a number of specific points.

Beside informing the user about how their personal data and relevant logging information will be handled and processed, existing directives typically require organisations to document explicitly and in details what information is being processed and by whom, always insisting that each processing of personal information must be an absolute necessity to operate the service offered to the user.

As a result, it is essential to implement and document a framework with the relevant internal procedures, detailing roles and responsibilities, as well as ensuring the controls are in place to enable appropriate access to personal information during routine security operations and incident response.

5. Conclusions & Future work

The WLCG collaboration has established a Cloud Traceability Working Group to carry out practical investigations and development of possible solutions. Practical experiments will demonstrate the viability of different technical solutions and service challenges will identify gaps in our traceability capability. The results from these areas of investigation will be used to inform the development of updated policies setting out requirements for running these new forms of distributed computing infrastructures without compromising traceability, as well as best practice recommendations for how to gather additional logging information and how to configure management frameworks and VM images.

While this work is focused in the already well developed WLCG collaboration, the policy and best practice we produce can provide a model for emerging cloud & virtualization based distributed computing infrastructures (EGI FedCloud, INDIGO DataCloud, HNSciCloud, etc.).

It remains an open question how to develop trust frameworks that will allow sites to accept VOs as full partners in incident response. A useful comparison can perhaps be made with the work done to establish WLCG that allowed workloads to be distributed across grids in the first place. We can be certain that security incidents will occur. It is vital that as we adopt new technologies and ways of working we do not lose our ability to effectively investigate security incidents or the ability to respond quickly and appropriately.

References

[1] The LCG TDR Editorial Board 2005 *LHC Computing Grid: Technical Design Report* (PDF). *document LCG-TDR-001, CERN-LHCC-2005-024* ISBN 92-9083-253-3. Available http://cdsweb.cern.ch/record/840543/files/lhcc-2005-024.pdf [Accessed 17 05 2015]

[2] Kelsey D 2011 *EGI Security Incident Response Policy* [online]. Available: https://documents.egi.eu/public/ShowDocument?docid=82 [Accessed 16 05 2015]

[3] Groep D 2007 *gLExec: gluing grid computing to the Unix world* CHEP 2007 [online] Avalailable: http://www.nikhef.nl/grid/lcaslcmaps/glexec/glexec-chep2007-limited.pdf [Accessed 16 05 2015]

[4] Activity_klog https://github.com/CERN-CERT/activity_klog [accessed 17 05 2015]

[5] Morovic S et al. 2015 *A scalable monitoring for the CMS Filter Farm based on elasticsearch* Oral Presentation Available: http://indico.cern.ch/event/304944/session/1/contribution/217 [Accessed 17 05 2015]

[6] Popple, K 2012 *OpenStack Folsom Architecture* [online]. Available: http://ken.pepple.info/openstack/2012/09/25/openstack-folsom-architecture/ [Accessed 16 05 2015]

[7] *OpenSOC Home Page* [online]. Available: http://opensoc.github.io/ [Accessed 16 05 2015]

[8] Grutzmacher K 2014 *Cisco OpenSOC Hadoop Design with Bro* [online]. Available https://www.bro.org/brocon2014/brocon2014_grutzmacher_opensoc.pdf (BroCon 2014) [Accessed 16 05 2015]

[9] Loomis C et al. 2012 *StratusLab Cloud Distribution* [online] Available: http://www.researchgate.net/publication/229441505_StratusLab_Cloud_Distribution European Research Activities in Cloud Computing, Chapter: 10, Publisher: Cambridge Scholars Publishing, pp.260-282

[10] *EU Directive 95/46/EC* [online] Available: http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:31995L0046:en:HTML