Auditing in shared virtualized environments

Adrian Baldwin, Simon Shiu, Yolanta Beres
Trusted Systems Laboratory
HP Laboratories Palo Alto
HPL-2008-4
January 16, 2008*

audit,
virtualization,
security,
assurance

The introduction of virtualization into the data centre can help provide a more manageable platform for the enterprise by reducing the number of physical machines and increasing utilization as well as enabling automation via virtual machines, networks and storage elements. Enterprises are required to demonstrate good IT governance and that business risks associated with IT systems are appropriately mitigated; companies have audit programs to provide such assurance. This paper examines the changes to risks that emerge due to the usages of virtualization within the data centre and we then discuss the requirements for an audit framework capable of providing automated assurance.

Internal Accession Date Only Approved for External Publication

Auditing in shared distributed virtualized environments

Adrian Baldwin, Simon Shiu, Yolanta Beres

Trusted Systems Laboratory, Hewlett Packard Labs, Bristol, UK
{Adrian.Baldwin, Yolanta.beres, Simon.Shiu}@hp.com

Abstract. The introduction of virtualization into the data centre can help provide a more manageable platform for the enterprise by reducing the number of physical machines and increasing utilization as well as enabling automation via virtual machines, networks and storage elements. Enterprises are required to demonstrate good IT governance and that business risks associated with IT systems are appropriately mitigated; companies have audit programs to provide such assurance. This paper examines the changes to risks that emerge due to the usages of virtualization within the data centre and we then discuss the requirements for an audit framework capable of providing automated assurance.

1. Introduction

IT governance includes ensuring that IT associated business risks are appropriately managed and mitigated. Having assessed risks, organisations will design controls to mitigate these risks. Usually these controls are augmentations to operational processes, and assurance usually entails auditors assessing the design and testing the effectiveness of these controls. This lifecycle, especially the assurance piece, is people and process intensive and so is very expensive.

In most industries assurance is a required part of the business. For example, the Sarbanes Oxley Act (SOX), see [1], makes assurance on the IT controls associated with financial reporting a cost of doing business. This has lead to a huge increase in the number of internal and external auditors testing financial processes and the supporting IT environments. This in turn has lead to a number of approaches to use technology to make the assurance process more efficient [2]. Out of scope for most approaches, is how architecture might be changed to make this lifecycle more

efficient. This is not uncommon as most changes are driven by business efficiency and agility reasons, with security and risk often an afterthought.

There are many current initiatives and changes to the way IT is managed by businesses, including service orientation, data centre consolidation, virtualization and automation. This paper examines consequences on audit and assurance likely to emerge from the changes brought about by virtualization in distributed (data centre) environments. It is aligned with broader research looking at how to use virtualization to improve the base security properties of infrastructure [3,4]. A vision for this work is that this will provide a trustworthy foundation for federated, distributed and shared computing environments. Moreover, it will do so in a way that does not rely on continually improving operational excellence and a growing number of highly skilled people.

More specifically this paper examines some of the different ways virtualization could be used in data centres, and provides an analysis both of the changes in the kind of information analysis that needs to be done (for example, perhaps less reliance on reconciliation of process events, and more reliance on attestation of component integrity) and risk trade offs for relying on "trusted infrastructure" as opposed to best practice operations.

The next section provides a top-down overview of the audit process, giving concrete examples of the kind of tests and information used in (today's non-virtualized) data centre audits. Section 3 provides a series of potential data centre changes relating to virtualization, and discusses them from a security, risk and audit perspective. Section 4 provides a lengthier example of automated change based on virtualization, the kind of information needed for assurance and contrasts this with the traditional approach. Section 5 uses the previous sections to draw out principles and design assumptions that we have made as we continue to prototype and research the assurance architecture and audit system to operate in a virtualized data centre. Finally, section 6 summarises and draws conclusions.

2. Audit, Assurance and the Enterprise Risk Lifecycle

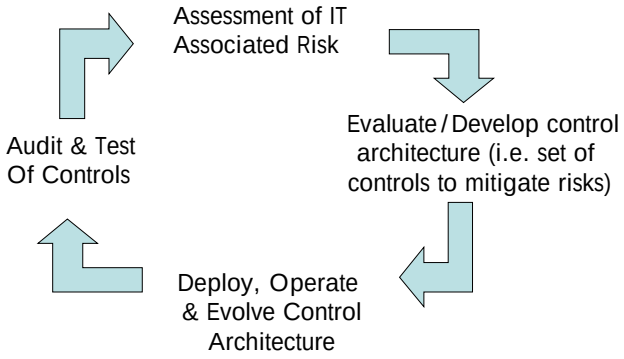

Fig. 1. An enterprise risk life-cycle

The figure shows a simplified version of the way most enterprises govern IT associated risk. First there is an assessment of the risks, for each risk a judgement is made whether and how to mitigate the risk. Most mitigations (i.e. controls) are process additions or augmentations. For example, ensuring changes in user privileges are always appropriately approved. The controls then have to be deployed and operated across heterogeneous application and operational infrastructures. Finally, the auditors have to test these controls and report their local and overall effectiveness. For an overview of standard frameworks and approaches for controls and assurance see [5,6].

Audit programmes typically operate at 3 levels: the business (or for SOX, financial) process, application management and infrastructure operations (i.e. the data centre). As such there is a lifecycle associated with each of these layers, but also inter-connection between the layers. For SOX the applications in scope are those involved in financial processes, and the controls are typically dictated by the financial process, for example the process may rely on the application to manage and enforce segregation of duties between certain users/roles. In turn the applications in scope determine which IT operations and infrastructure needs to be audited, but since operations often work across the data centre, it is often the whole data centre that is audited.

The process of assurance also involves 3 different types of audit and auditors. There is self assessment by the process, application and operations manager/owner; there are the enterprises internal auditors; and the external

auditors. The external auditors provide the required 3rd party attestation that risk is being managed. They will perform a series of audits themselves, but this level of inspection can be reduced if the internal audit organisation and self assessment processes are sufficiently effective.

A mature internal audit department will have documented the critical controls that they will be testing. Included in this documentation will be a description of the control together with a description of the risk being mitigated. There may also be some guidance on how the control should be tested (e.g. sample 25 changes and check that all due diligence was done and documented prior to approval). Each time an audit is performed the auditors follow the documentation, thus ensuring consistent reporting. Since many systems vary in the way log and state information is preserved, the documentation is not usually prescriptive to this level of detail, and there is much reliance on the skill of the auditor. Typically though, the analysis involves:

- Reconciliation of chains of (workflow) events, e.g. for each object changed at this time, what is the context for this change, and which approved work item does this relate to;
- Configuration and state checks, e.g. is the system configured to enforce password changes every 60 days.
- Statistics and trends over time, e.g. the number of emergency changes occurring each month (a high number indicating a potential control issue, and spikes in emergency activity warranting further investigation).

At the higher level the audit director or programme manager is responsible for resourcing and scheduling the audits. It is here that an aggregated view and an overall judgement on control can be made. Key risk indicators [7] are often used at this level to create an overall and comparative view of how well different environments are controlling risk. For example by regularly recording the number of emergency changes in each month, across all environments the director is able to (indirectly) see variance in performance over time and between systems, perhaps scheduling the audit based on these results. The recent PCAOB guidelines [1], makes specific provision for allowing benchmarking through appropriate KRI's as an acceptable practice in mature environments.

2.1 Audit in a data centre

A typical data centre audit involves identifying risks to the business that result from failures within the data centre and for each risk a number of controls are identified

and auditors test the effectiveness of these controls. Within this section, we identify the three main areas within a data centre audit and briefly outline some of the audit testwork.

1. Data Centre Operations: That is looking at the overall management of the data centre, which includes auditing the physical aspects of the data centre as well as looking at management processes.
   - Data Centre Security: Checks here include looking at existence of door locks and people who are given access have access requests along with corresponding approvals.
   - Operations Management: Checks here look at the physical properties of the data centre for example checking temperature; humidity and power logs and the results of fire drills
   - Incidents and Service Requests: Checks are made on effectiveness of the incident management process by looking at a sample of 25 incident tickets and all open tickets.
   - Disaster recovery plan: Checks here often involve a check of the existence of plans and that plans are tested.
   - Availability: Review of availability metrics as compared with SLAs along with a review of exceptions, and incidents of down time.
2. Servers: That is looking at risks that exist due to the way servers are managed within the data centre. Again the audit checks fall into a number of categories.
   - Antivirus: Check that virus definitions are up to date and review last scan times and last virus found times.
   - Patch management: Review the patch level on a sample of servers and verification of change records for patch testing.
   - Password: Check the password policies and settings on all critical servers.
   - Change: Review of all change records on critical servers.
   - Capacity: Review the last two months of usage of cpu, memory, disk space and utilisation along with monthly exception reports. Check for evidence that the customers are kept informed.
3. Data Centre Services: There are a number of services run within the data centre that are used by various systems and upon which many systems rely.
   - Active Directory: Risks relate to user management and checks are made on a selection of new and terminated accounts,
   - User Management: Checks made on all new privileges and that there is a regular review of the privileges. Checks on all terminated employees.

- Security: Check that security event information is reviewed and incidents dealt with. Checks on security sensitive configurations.
- Backup: Check backup schedule for a sample of servers. Checks of requests for restoration of data and people who have access to the backup data.
- Config: Check on mandatory fields within the CMDB are filled. Review accuracy of a sample of entries within the CMDB,

3. Audit requirements for a virtualised data centre

Many are proposing the use of virtualisation within the data centre to achieve a number of goals from improved utilisation through to the construction of an automated shared flexible data centre. Clearly, such a range of different options will lead to a wide variance in the change in the risk profile; controls and hence the audit requirements. Here we identify a number of ways that virtualisation can be used in the data centre as well as areas of additional risk.

3.1 Machine Virtualization

A major driver for virtualization is simply to reduce the number of servers in our data centres. That is, if applications can share a server then the number (and hence cost) of server and server maintenance goes down. Such an approach is already happening in test and development environments and in some cases; virtualisation is being used to reduce costs in failover/recovery systems. The next step is to have production systems running within the virtualisation layers.

Further complication within the use of virtualisation comes where virtual images can be suspended for use when needed, e.g. if other systems fail or as loads increase. The ability to migrate a virtual machine from one physical machine to another for example to rebalance server loads as needs change again introduces further risks.

Here we examine some of the risks that have been introduced as they relate to the audit programme.

1. Data Centre Operations:
   - Virtualisation introduces a layer of indirection between the running OS and the physical machine. This makes it harder to check that the virtual machines are running within a protected environment. When migration and suspension

are introduced, we need to know that the suspended image and the migrated VM both remain within a physically protected boundary. This implies the need for detailed logging and correlation of the lifecycle events for the VM.

- Availability: Availability metrics can now become more complex in that we need to look at the availability of individual VMs and the virtualisation layer. Migration can help administrators manage availability but introduces additional risks in the way it is managed and hence availability management records need further review.

2. Servers: Many of the audit checks that apply to individual servers will remain the same whether they are on physical or virtual servers. An additional complexity that requires audit is that the state of the virtualisation layer (the VMM and management domain) needs the same processes running on them as the critical servers. Audit needs to check that these are functioning effectively. Such checks are particularly critical in that multiple business services may be dependent on this infrastructure layer.

   Additional risks exist due to the ability to suspend machines and hence live checks on machines may not be possible (for example, there are many tools (e.g. Symantec's ESM) that have agents running on each machine to check that the machine is in a compliant state). Audit needs to check the process for ensuring the update and checking of suspended machines before becoming live.

   Capacity management becomes more complex in that cpu and memory resources are shared and it is important that the physical to virtual mapping is understood in ensuring adequate capacity for each VM. The ability to migrate VMs over the data centre enables better capacity management and audit need to check that there are processes for ensuring that there is sufficient capacity as a whole within the data centre.

3. Data Centre Services
   - Security: There are additional risks introduced in the configuration and potential attack of the virtual infrastructure layer. As such, auditors need to pay particular attention to check that configuration and security events from this layer are well managed.
   - Backup: Again, there are additional risks introduced by the additional layer of virtual infrastructure that needs to be backed up and auditors need to check that this virtual infrastructure can be recreated. Additional backup issues occur with suspended VMs where checks need to be made that there is an adequate backup plan.

- Configuration: The CMDB maintains the description of the infrastructure and should maintain the physical to virtual mappings. Audit checks should establish the correctness of these mappings even when migration introduces fluidity. The CMDB also needs to hold the location of suspended images and audit needs to ensure these records are accurate and that when a suspended machine becomes live it is the correct image.

3.2 Virtual infrastructure

Further out it is possible to build a virtual data centre using a combination of virtual resources such as virtual storage and virtual networking. This has the advantage that distributed components that normally rely on separate infrastructure and have their own internal communication constraints (e.g. 3-tier web architecture) can gain flexibility and share physical resources. The introduction of further virtual resources adds complexity to the physical to virtual mappings.

Such a change will exaggerate even further the changes discussed in section 3.1. For example, now the VMM is being relied on to control the view each component has of their virtual network. In doing so it must both maintain separations to manage which machines can access a given virtual resource (such as a network) and enforce policy determining the manner of access. The management domain drives the configuration of the way the virtual resources are used and hence ensuring the security of this system is critical.

Again, we examine some of the additional risks that have been introduced as they relate to the audit programme

1. Data Centre Operations:
   - Data Centre Security and Data Centre Ops: As before, there are issues around knowing that all virtual resources are contained within a safe physical boundary. For example, it is important to understand that connections to the network can be controlled within the vmm, rather than at the switch (and similarly for storage).
   - Availability: In understanding the availability of all virtual resources, it is necessary to ensure that the physical to virtual mapping is understood and used when considering the ability to meet SLAs.
2. Servers:
   - Anti-Virus and Patching: The use of a virtualised system allows us to create better segmentation between machines. For example, isolating a critical server with limited network connectivity; and this strategy may be used instead of

patching where the risks of changing the systems are considered high. In these cases, exception requests need checking and auditors need to check that there are appropriate policy controls in place to limit access to these resources.

- Data Scrubbing: With the use of virtualized storage additional assurance information needs to be gathered about data cleanup/scrubbing between allocations to different VMMs. This should provide evidence that data is not leaked across services running on VMMs. Also this might require changes in the way backup is managed and performed.
- Capacity Management: As more resources such as disk and network are virtualised, there is a greater ability to ensure flexibility to manage capacity but ensuring each machine has sufficient capacity involves understanding the mappings from the virtual physical world. We then need to understand that there is sufficient overall capacity and that it can be adequately shared between the resources.

3. Data Centre Services:
   - Configuration: As before, the CMDB must contain records for all virtual resources along with their physical mappings.

3.3 Automation

Data centres that produce virtual infrastructures using a variety of virtual resources will be complex to configure but highly flexible and therefore lend themselves to the use of automation. This could manifest itself in full-blown desired state automation, or it could simply mean automation of some of the better-understood configuration tasks (codebook automation). Clearly, automation brings tremendous advantages in terms of agility, cost and scale. Automation in itself does not introduce additional risks due to virtualisation but the combination of automation with an agile infrastructure introduces considerable risk into the change control process:

1) The links between the change management and actions on the infrastructure can become obscure especially where several changes happen at once. Auditors would typically check that changes can be reconciled against a change request with an appropriate test plan; such checks become much harder.
2) The automation may allow policies that introduce flexibility into the data centre – for example, to add additional web servers during periods of heavy loading. There are configuration changes here that occur outside of

the normal change control framework and again it is important that changes can be reconciled to the policies and the conditions that allow the change.

3) Where there are a combination of manual and automated changed then understanding who did what changes can again be highly complex.

Both of which lead to the questions:

1) At what level of abstraction should there be “human” authorisation of changes?
2) What sort of analysis should be done on the events and state to provide assurance that control is being maintained?

Since desired state management introduces such a large change to the way IT operations are managed, it seems likely that there will be a mixture of automated actions together with manual interventions, increasing the level of complexity from an assurance perspective.

With automation it can actually be easier to provide assurance as long as audit data gathering and state checking are built into the automation system from the start. For example, every time a change control is executed via the automated process the automated system can easily record the steps taken and the changes performed as well as a final check of the system state. It is much easier to include auditing as part of the automation processes than build it later on top. This also removes the need for any additional agents that are currently being deployed to gather just assurance information. Section 5 includes further discussion on auditing of automated systems.

3.4 Discussion

The above discussion looks at some of the ways virtualisation can be used within the data centre and the impacts that it may have on risks within the data centre. Some of the changes lead to additional processes that auditors must check; often such checks will use the normal audit techniques of checking a sample of a process; collecting evidence and analysing it to establish the effectiveness of a process. For example, auditors need to check that the virtual infrastructure is well run and properly patched and all changes go through the appropriate change control process.

The added complexity for audit is in the introduction of an additional layer of indirection and it is not sufficient to audit this additional layer. Instead, auditors must ensure that the correct relationships between the virtual and physical world is being maintained. This means when conducting the standard data centre audit they need to look for evidence that the virtual resources have the appropriate logical controls over sharing and that these relate to physical controls. This implies that we need a detailed audit log of where virtual machines are running and to which virtual resources they are connected and how.

Now a customer has an abstract description of the infrastructure they require and ideally we would have an evidence trail showing that the infrastructure achieves this abstract description. In this case, assurance and audit is concerned with reconciling the actual systems to the abstraction provided to the management system.

4. Understanding the virtualization layer

The above discussion identifies the need to understand and provide assurance over the linkage between the virtual and physical layer. For example, this allows us to understand that the virtual machines are running on appropriate hardware in a physically secure boundary. Understanding how virtual resources are interconnected and connected to the underlying physical resources is also essential in establishing assurance on the workings of a virtual infrastructure. To understand such relationships requires an understanding of the underlying technologies and their lifecycles; with each technology working in a slightly different way and hence requiring different analysis.

This section gives an example of how such relationships can be understood and trust established by understanding the lifecycle events within the virtual infrastructure. Here we explore the lifecycles associated with creating and managing virtual machines within Xen [9] along with their linkages to virtual or physical resources. To understand that virtual machines are correctly formed and connected we need to retain an evidence trail around these life-cycle events and analyse them and correlate them with the appropriate change control records from the ITIL management systems [10].

The life-cycle events for virtual resources are concerned with:

– Creation/Destruction of the resources – For example, the creation of a virtual machine or a virtual disk or network.

- Binding/Releasing of the resource to/from a physical system – For example, starting a virtual machine on a given platform or migrating it a different platform or having a virtual disk on a given storage server.
- Connecting/Disconnecting virtual resources – For example, connecting a virtual machine to a virtual network and a virtual disk. Here we are concerned not only about the connection but also the form of the connection (e.g. if there is a policy to limit network access).
- Suspension of Virtual resources – That is where a resource has no given mapping to a physical resource providing its functionality. For example, a virtual machine may exist on disk with no physical platform to run the system.

There are be additional steps within the evidence trail involving supporting services such as the management system, patching, or network services that help in connecting systems.

Whilst there are a number of key life cycle events that need tracking within the audit trail the details of the events, what is collected and from where will vary according to the different virtualisation platform.

Here we provide an example of the stages that are involved in the creation of a virtual machine Xen virtual machine that is then linked to an iSCSI storage device. Here we assume that the there is a management system, for example based on SmartFrog [11], co-ordinating the creation of virtual machines.

Within this example we start by describing the start up of a virtual machine and how it is connected to virtual resources to illustrate the events that need to be collected and the different components from the virtual infrastructure that are involved. This demonstrates the types of sequences of events that need to be reconciled in order to validate the correct operations of the data centre.

Creating virtual machines

Bringing up a virtual machine involves an interaction with a number of processes with the XenD process co-coordinating the operations within Dom0.

1) The management system creates a configuration file that includes a description of the required devices such as networks and storage along with, in the case of Xen, a reference to the kernel file/image.
2) The management system then issues a create domain command to Dom0.
3) The XenD process on Dom0 starts to create the new virtual machine.
    a. A new VM is created with a communication channel between the VM and Dom0.

b. As the VM starts, it sends a message to Dom0 to see what devices it has.
c. Dom 0 uses the configuration file to return device information.
d. Dom 0 contacts each of the relevant backend drivers to create devices for the VM and a communication channel is created for the VM and driver to communicate. Driver scripts can be run at this point to ensure that policies are enforced at the device level.
e. Dom 0 gives the VM handles to all the devices.
f. The VM continues to boot.

4) The management system passes back a link to the machine to the customer.

The normal setup of Xen would place the management agent, the drivers and the configuration controller all within the Dom0 management domain. Drivers could run within separate domains (as supported by Xen) hence providing separations between key system elements.

Currently this data can only be logged/collected through mechanisms such as XenTrace or SysLog, which are under the control of Dom0. This creates an independence problem as in audit situations, we are asking the same subject to reconcile information with instructions it was given. This problem and some early research considerations are developed further in [12].

Other Events and assurance information

As well as collecting the basic lifecycle events there are other sets of events that are important in gaining assurance of the trustworthiness of the platform. For example:

- Events around the provision of other virtual resources. For example, events involving the lifecycle of virtual disks on an iSCSI server or the allocation of new VLAN ids.
- Attestations of platform and VM state. – TCG [13], mechanisms provide a way to measure aspects of a platform and the state into which it has booted. The measurements provide a base-line for understanding the lifecycle events; for example, this can help provide assurance that a VM boots on a trustworthy VMM layer. TCG's identity mechanism can also be used to identify the physical machine where a VM has booted and hence confirm its presence in a given data centre.
- Multiple VMs. Some virtual infrastructure designs may rely on combinations of virtual machines. For example, [14] demonstrate the use of a VM based on a

library OS to create a virtual TPM which would sit along side a client OS. Here we need to record the co-occurrence of these systems.
- VMM state. In some cases, it is useful to gain an insight into the state of the VMM, for example, to understand policies that are being enforced between VMs or that given (unenforced) connection preferences are maintained.
- The management state. Events, change control and desired state information provide the context by which all the lower level events can be understood and validated.

Audit Analysis

As discussed earlier audit analysis typically involves sampling a number of key processes for example reconciling a number of changes in the IT system with the change control records showing that they were appropriately approved. The audit process does not necessarily change with the addition of virtualization but where the additional flexibility is heavily used, the audit analysis will become much more complex. Typically, auditors will check a small sample of events but as the overall number increases the sample size should increase. As automation is used to control the systems or where event sequences can be modeled this enables us to analyze the raw events discussed in this section and aggregate them into composite events that can be understood and reconciled against the events within the management system.

5. Discussion

Aggregating and Collecting Audit Information

Having trustable flexible data centre solutions requires an associated assurance solution, which only becomes practical where we have an automated audit collection and analysis framework.

The audit process first of all requires evidence about events to be available and in a flexible data centre; this means that we need to put the hooks into the utility management system (UMS) as well as the systems managing the virtual resources. For example, we need to capture and store events that the UMS creates as well as those happening at the virtualization layer; in storage devices; networks etc. Some such systems may create log files, which act as a good source of information others

such as management agents and device drivers may require instrumentation to be built in to the system.

The collection of data may also involve validating that the system is in a given state – for example, after establishing a bridge between a virtual network and a physical network a script can be run to report on the status of all such connections. Such state gathering evidence needs to be triggered, potentially a role for a local (platform) audit manager, [ref techcon].

The virtual platform (aka local audit manager) provides a place where we can place an audit staging post where information can be buffered; secured locally and even some simple analysis can be performed. Such a mechanism can be put in place using a simple library OS whose state can be attested to using TCG mechanism and therefore we can have some trust in its management of audit information and the separations from the overall management layer.

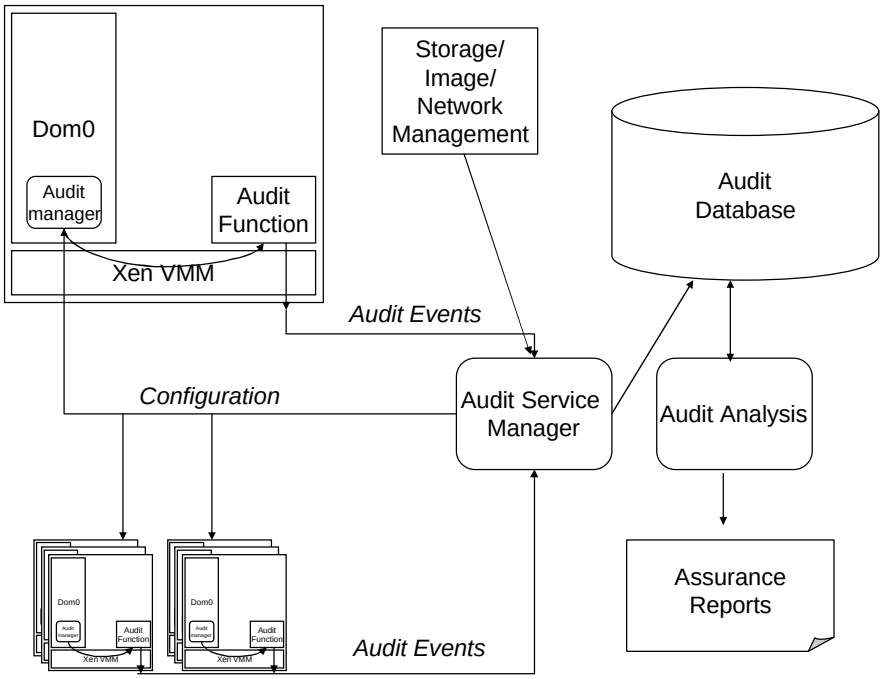

Fig 2 Secure Collection of data from virtual infrastructure

Figure 2 shows a basic framework for an automated collection system with hooks into the basic framework, audit domains on virtual machines to collect and forward audit event, an audit database and audit analysis and reporting systems.

Auditing automated systems

There is a tradeoff that can happen within audit around the need to check how fully automated systems are working. For example, if there is a fully automated workflow within a purchase system that only allows approved purchases from an approved supplier then audit may simply check that the system is functioning correctly rather than checking a sizable sample of transactions. An auditor would probably check any transactions going through an exception process.

The same argument could be applied to a fully automated data centre and hence, we could argue that it is not necessary to check the low-level events since they are a direct uncontrollable consequence of higher-level commands sent to the data centre. This could become the case where automation is mature, well understood and well controlled but current systems are normally run on top of large general-purpose operating systems allowing administrator access. The reliance on an automation system that spans both the main controller as well as agents on each of the management virtual machines on each physical machine makes it hard to establish that the system is functioning correctly in contrast to validating the functionality of a workflow within a well-controlled ERP system. The existence of admin accounts on each of the management domains is not the same as an exception process but gives those who can log in complete control over the system. As such, until the automation system is built from small well controlled components [14] there is a need to have an independent audit trail showing that the low level virtual machine configuration events reflect the intent within the management system.

It may be hard to encode effective exception processes within an automation system and as such, it becomes desirable to have audit analysis that can detect and extract certain behavior which could be of concern. This is effectively the difference between enforcing strong policies and having an audit trail identifying unusual but valid events representing additional risks that need further checking. For example, a data centre could have policies ensuring that developers are never allowed to log into a production system; however, where problems occur it is often necessary to allow developers limited access to production systems so that they can

fully understand the problem. In these cases, it is useful to ensure good audit rather than having strong policy enforcement mechanisms that can slow down or prevent necessary maintenance actions.

Trust in the audit trail

Audit provides an effective assurance mechanism in that it provides an independent check on what has happened based on analyzing a paper trail of evidence. From an assurance perspective, we need to develop trust in the evidence trail; both that the data represents what happened (not what someone wants us to think happened) and that the data has been kept in such a way that it remains trustable (i.e. it has not been changed since generation).

The first of these issues is probably the hardest to address. Given the range of sources of events, it is going to be hard to ground each and every event with a strong provenance. However, audit is concerned with reconciliation of events from a variety of sources thereby looking for consistency between the different sources. Having some events whose provenance we can justify can therefore help in grounding the overall audit trail. For example, if customer requests were to come in the form of signed e-mails this would help ground changes that are reported from the management system. Building trust from the bottom up the TCG based attestations from the VMM can help us justify trust in raw audit messages coming from this layer and help us determine the independence of the audit trail at this layer.

There are many mechanisms for securing audit trails from forward integrity schemes that provide light weight crypto that can be placed on each agent or staging post through to timestamped and chained messages [15,16,17,18,19] that allow the audit trails to be shared with interested parties. Such mechanisms use crypto to link and secure audit trails to demonstrate that individual events and event sequences have not changed but they also rely on independent components to manage the crypto keys and hence provide a degree of independence on the trail. For example, a time stamping service, sealing event sequences with time and a digital signature, only protects the audit trail for assurance purposes whilst those being audited cannot gain access to the signature key and later rework the evidence trail.

Different views on audit

Traditionally in an enterprise, an audit report goes to those being audited and the head of audit (reporting to the CFO who is ultimately responsible for compliance).

There is then a tracking process to ensure that issues raised are dealt with so that risks are appropriately mitigated, see [20]. The same could be true of the results of an automated audit framework for a virtual data centre however, the a flexibility brought by virtualization can bring dynamic IT resources than many rely on for a variety of business tasks and in outsourcing or IT service environments these relying parties may be different companies. As such, it can be useful to share views on the audit data showing the relying parties that the overall data centre is well run and that their own virtual resources are properly maintained. Providing such separation of different aspects of the audit programme will undoubtedly be one of the challenges that needs addressing within an overall audit framework.

Transparency on audit data may help engender trust with the various users of data centres but only when it can be presented in a form that is quick to view and easy for the recipient to understand. Trust will not be gained by simply providing access to large volumes of data but analysis and reporting systems that can help provide coherent views and traffic light summaries. It may be that it is not necessary to expose detailed audit trails to customer but rather provide them with a summary report and the ability to have the detailed trails audited (perhaps by independent auditors). This approach is in line with current practices such as the use of SAS 70 (Statement on Auditing Standards no 70) certification for data centres where independent auditors provide a certification for use by multiple customers. An ongoing challenge, especially in the context of outsourcing is how to provide continuous (and perhaps context specific/tailored) assurance that a data centre is well run. The growing use of virtualization makes this an interesting and challenging area for research.

Conclusion

This paper has provided an analysis of the security and assurance consequences that virtualization will bring to the data centre. We have presented an analysis of additional risks that may be introduced and over which assurance should be given. We provided a more detailed view on some of the lifecycle events that need to be understood to provide adequate assurance over a virtual infrastructure. Given this background we have begin to describe, and suggest approaches to the requirements of an automated audit framework. Without clear audit strategies supported by technology we believe it will be hard for virtual data centres to be

adopted for critical business applications and this is particularly true for shared data centres.

References

[1] Public Company Accounting Oversight Board (PCAOB), see http://www.pcaobus.org

[2] ACL: Audit Analytics and Continuous Monitoring Solutions, see http://www.acl.com

[3] Open Trusted Computing, EU Funded Research Project, see http:/www.opentc.net

[4] Cabuk S, Dalton C, Ramasamy H, Schunter M, "Towards Automated Provisioning of Secure Virtualized Networks", Proceedings of the 14th ACM conference on Computer and communications security, Virginia 2007

[5] COBIT, Control Objectives for IT, see http://www.isaca.org

[6] COSO: (Discussion Document) Guidance on Monitoring Internal Control Systems, see http://www.coso.org

[7] "Continuous Control Monitoring: Enabling rapid response to control breakdowns", in research findings of Audit Director Roundtable 2004, http://www.audit.executiveboard.com/ADR/

[8] Kallahalla, M., Uysal, M., Swaminathan, R., Lowell, D. E., Wray, M., Christian, T., Edwards, N., Dalton, C. I., and Gittler, F. 2004. SoftUDC: A Software-Based Data Center for Utility Computing. Computer 37, 11 (Nov. 2004), 38-46. DOI= http://dx.doi.org/10.1109/MC.2004.221

[9] Barham, P., Dragovic, B., Fraser, K., Hand, S., Harris, T., Ho, A., Neugebauer, R., Pratt, I., and Warfield, A. 2003. "Xen and the art of virtualization", In Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles (Bolton Landing, NY, USA, October 19 - 22, 2003). SOSP '03. ACM Press, New York, NY, 164-177. DOI= http://doi.acm.org/10.1145/945445.945462

[10] ITIL v3, see http://www.itil.org.uk

[11] Sabharwal R, "Grid Infrastructure Deployment Using Smartfrog Technology", HP Knowledge Brief 2006

[12] Baldwin A, Dalton C, Parthipan L, Shiu S, "Trusted Audit for a Virtual Data Centre" abstract submitted to Tech Con 2008

[13] Pearson S., B. Balacheff, L. Chen, D. Plaquin and G. Proudler. Trusted Computing Platforms: TCPA technology in comtext. HP Books, Prentice Hall (2002)

[14] Anderson M, Moffie M, Dalton C, "Towards Trustworthy Virtualization Environments: Xen Library OS Security Service Infrastructure", HP Tech Reort 2007, HPL-2007-69

[11] Bellare, M., and Yee, B., Forward-Security in Private-Key Cryptography. Topics in Cryptology - CT-RSA 03, LNCS Vol. 2612Springer-Verlag, 2003

[12] Haber, S.A., W.S. Stornetta, How to timestamp a digital document. Journal of Cryptography 3(2):88-111 1991

[13] Merkle, R. C., “Protocols for Public Key Cryptography”, IEEE Symposium on Security and Privacy, pp 122-134, 1980

[18] Schneier, B., Kelsey, J., "Cryptographic Support for Secure Logs on Untrusted Machines," 7th USENIX Security Symposium Proceedings, USENIX Press, 1998.

[19] Baldwin, A. Shiu, S.:Enabling Shared Audit Data, In Proceedings of the 6th Information Security Conference Eds Boyd and Mao, Springer Verlag (2003)

[20] Baldwin A, Beres Y, Shiu S, “Using Assurance Models to Aid the Risk and Assurance Lifecycle”, HP Tech Report 2007, hpl-2007-48