Managing Forensic Recovery in the Cloud

George R. S. Weir
Department of Computer and Information Sciences
University of Strathclyde
Glasgow, UK
e-mail: george.weir@strath.ac.uk

Andreas Aßmuth and Nicholas Jäger
University of Applied Sciences
OTH Amberg-Weiden
Germany
e-mail: {a.assmuth,n.jaeger}@oth-aw.de

Abstract— As organisations move away from locally hosted computer services toward Cloud platforms, there is a corresponding need to ensure the forensic integrity of such instances. The primary reasons for concern are (i) the locus of responsibility, and (ii) the associated risk of legal sanction and financial penalty. Building upon previously proposed techniques for intrusion monitoring, we highlight the multi-level interpretation problem, propose enhanced monitoring of Cloud-based systems at diverse operational and data storage level as a basis for review of historical change across the hosted system and afford scope to identify any data impact from hostile action or 'friendly fire'.

Keywords-Cloud security; forensic readiness; message authentication codes; secret sharing.

I. INTRODUCTION

For many individuals, the primary use of Cloud computing is remote data storage. Presently, most major online Cloud service providers offer such storage. Apple users may engage iCloud as a supplement to local storage capacity and as an emergency backup for system configuration. Among similar service offerings we find Google Drive, Microsoft OneDrive and Amazon Drive.

Dropbox and its freemium business model, where users may register for a free account with a limited storage size and an option for more storage capacity and additional features for paid subscriptions, is also very popular. The broad appeal and immediate benefits from services of this type are apparent from the proliferation of such offerings, as underlined by the fact that many home broadband contracts include a measure of Cloud storage as standard. Thus, "BT Cloud is a free service for BT Broadband customers that allows you to securely back up, access and share your precious files and folders" [1]. Home broadband users will often rely on their remote storage and backup facility with little recognition that Cloud services are in operation.

Despite the apparent speed with which consumers have adopted Cloud-based services, there is recognition that security issues can arise in the Cloud setting just as in the context of locally hosted systems [2, 3, 4, and 5]. When occasional security issues are reported in the media, the greatest concern may be the availability and privacy of their data [e.g., 6].

II. NETWORK SECURITY RISKS

Addressing security risks is a familiar issue in the context of networked computing. In non-Cloud systems, the principal ingredients in management responses to security take three general forms:

- System hardening
- Software defences
- Data backup

Firstly, system hardening is an attempt to render known threats ineffective. This includes 'conventional' measures that reduce vulnerability, such as authentication, identity management and access control [7], as well as acting to disable unnecessary services, applying regular software updates (patches) and gauging of the relevance and associated risks from newly published exploits [8]. Modern Operating Systems have also been adapted to meet known cyber threats. Counter measures, like address space randomisation, mandatory access control or maybe sandboxing, are state of the art. In addition, advanced users might even build their own operating system and use selected kernel parameters to further harden their system. The second variety of response to address security issues is the application of software defences. This ranges from antivirus provision to firewalls and may also include some variety of intrusion detection, usually rule-based [9] or anomaly-based [10].

Any computing system may be described by a simple layer-based model as depicted in Fig. 1. Obviously, security on any higher layer strongly depends on access control mechanisms of lower layers. Even if users or service providers only aim for access control on a higher level to secure their application, these access control mechanisms in practice are more complex than those on lower layers. In addition, vulnerabilities or inadequate configuration on lower levels may lead to bypassing security measures on higher

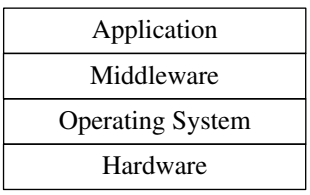

Figure 1: Layer-based model of a computing system

layers. Therefore, appropriate countermeasures are necessary on all layers.

A third security measure is the provision of regular data backup, as a means of ensuring that any system failure or intrusion does not result in irretrievable data loss.

III. CLOUD SECURITY RISKS

Perhaps unsurprisingly, Cloud configurations are subject to levels of security risk that go beyond those affecting conventional networked computer systems. In consequence, the security measures outlined above may not be sufficient in the Cloud setting. In elaborating this claim, the Cloud issues are best illustrated with reference to the differing Cloud service offerings [11]:

- Infrastructure as a Service (IaaS);
- Platform as a Service (PaaS);
- Software as a Service (SaaS).

These models for Cloud service provision are helpfully elucidated by Gibson, et. al. [12], as follows:

- "IaaS provides users with a web-based service that can be used to create, destroy, and manage virtual machines and storage. It can be used to meter the use of resources over a period of time which in turn can be billed back to users at a negotiated rate. It alleviates the users of the responsibility of managing the physical and virtualized infrastructure, while still retaining control over the operating system, configuration, and software running on the virtual machines" [op. cit., p. 199].
- "Platform-as-a-Service providers offer access to APIs, programming languages and development middleware which allows subscribers to develop custom applications without installing or configuring the development environment" [op. cit., p. 200].
- "Software-as-a-Service gives subscribed or pay-per-use users access to software or services that reside in the cloud and not on the user's device" [op. cit., p. 202].

Clearly, our earlier noted approaches to system security are also applicable to Cloud-based systems. With an eye specifically on Cloud security, we can consider how each of these service offerings may be at risk and what precautions may be anticipated in response to these risks.

1. Infrastructure as a Service

This kind of service seems most prone to the types of exploit that one would expect with conventional networked computers, principally, because, in most cases, such virtual machines will be presented to the Internet as networked hosts. Here, the customer is deploying a virtual machine with associated Operating System and on-board software applications. This raises the prospect of vulnerabilities at network level, as well as application level issues, e.g., with Web systems and Database servers, Cross-Site Scripting (XSS) or SQL injections. Denial of service attacks are also a legitimate concern, especially since this kind of attack can achieve enormous bandwidths by using IoT devices for their purpose [13]. For these reasons, system hardening (especially, defending against known vulnerabilities) and software defences are appropriate for IaaS, including precautions such as anti-malware, firewalls, and Intrusion Detection Systems. Provision of these features may be the responsibility of the Cloud Service Provider (CSP), who determines what OS and defensive capabilities are made available. In some settings, the customer may be in a position to bolster the native defences on the virtual system provided by the CSP.

In similar vein, data backup is likely to be required by the IaaS customer. Indeed, the protection of customer data may jointly be the concern of the customer and the CSP. The former may enable off-Cloud backup, to avoid a single source of failure. While the CSP may also offer data backup to a separate Cloud data storage facility.

Despite reasonable expectation of such measures, there are indications that Cloud software infrastructure components are not always adequately secured from known vulnerabilities at the virtual machine level [14].

2. Platform as a Service

Computing facilities afforded to the customer of PaaS, are limited to the development of specific middleware or functional components. These services employ technologies such as Docker [15], Containers [16], DevOps [17] and AWS Lambda [18], in order to host customer-defined remote functionality. From a Cloud customer perspective, system hardening seems to be irrelevant in this context in relation to the host Operating System. On the other hand, any code developed for use on the Cloud platform must be protected from illicit operations, e.g., process hijacking, output redirection or the elevation of privileges.

Software defences of the variety outlined above seem less relevant to the PaaS context since the operations supported by the middleware are limited to specific data processing and do not afford full operating system access or modification. The primary concern should be the operational effectiveness and resilience of the customer-defined operations. Clearly, such services may also be impaired through illicit access, e.g., stealing authentication details in order to alter code on the host system. Managing this area of concern lies primarily in the hands of the Cloud customer, with the assumption that the CSP will prevent unauthorised access to customer account details.

3. Software as a Service

SaaS provides the Cloud customer with remote access to third-party data processing facilities via micro-services [19], or RESTful services [20]. Aside from network level attacks, such services should be protected from most other security concerns by having the host system hardened and equipped with suitable software defences. From the customer perspective, so long as their remote Cloud services operate

effectively, without interruption or data loss, there would seem to be little cause for concern. Of course, the risk of aberrant customer-side behaviour may arise through social engineering exploits or disgruntled employee actions.

This summary of security concerns affecting the three varieties of service has treated each Cloud model as an isolated networked computing facility. In reality, since the essence of Cloud provision is the virtualisation of services, our overview lacks one further important consideration, i.e., the possibility of service impairment as a result of activity at adjacent, upper or lower levels of the Cloud implementation.

Clearly, any security aspects that affect the operational resilience of the underlying Cloud infrastructure is of direct concern to the CSP and can have a knock-on effect upon customer services. The underlying Cloud technology, i.e., the hardware and software configurations that provision our three Cloud models, may be subject to attack or deliberate manipulation in a fashion that impinges detrimentally upon the Cloud services supported by that particular hardware and software ensemble. This may be construed as a service attack 'from below'. The scope for such attacks are precisely the characteristic exploits that may affect any networked host (listed earlier).

Attacks 'from the side' are a growing concern in Cloud security. 'Side channel attacks', originate with co-hosted customers who manipulate the behaviour of their virtual system to influence the behaviour of the host system and thereby affect co-hosted customers. Several studies suggest that such 'co-tenancy', an essential feature of IaaS and PaaS, carries dangers. Thus, "Physical co-residency with other tenants poses a particular risk" [21], such as "cache-based side-channel attacks" [22], and "resource-freeing attacks (RFAs)" in which "the goal is to modify the work- load of a victim VM in a way that frees up resources for the attacker's VM" [23]. Most worrying are contexts where one customer's 'malicious' virtual machine seeks to extract information from another customer's virtual machine on the same Cloud platform [24]. Such risks to Cloud facilities are fundamental to their service provision.

A final attack vector that threatens some Cloud systems is 'from above'. In this case, poorly implemented virtual systems may afford scope for customers to 'break free' of their virtual system and access or directly affect the underlying Operating System or middleware/hypervisor. Clearly, it must be ensured that there is no information leakage from virtual machines and that attackers or malicious customers are not capable of breaking out of the virtual machine and gaining access to the host OS or the virtual machines of other customers [25].

The characteristics of these Cloud service offerings with associated security measures and the likely risk conditions are captured in Fig.2. The prospect of action from one Cloud user affecting another is described as intra-platform interference.

IV. DIGITAL FORENSIC READINESS

The numbers of cases of network intrusion and data breach are on the rise: "there is a massive increase in the records being compromised by external hacking – from roughly 49 million records in 2013 to 121 million and counting in 2015" [26].

Service model	Main features	Security Measures	Risks
Infrastructure (IaaS)	Virtual machines, Operating systems, Storage, Software applications	System hardening, Software defences, Data backup	Social engineering, Intrusion, Malware, Denial of service, Elevation of privileges, Intra-platform interference
Platform (PaaS)	APIs, Programming languages, Development middleware, (Containers, Dockers, AWS Lambda, DevOps)	System hardening, Software defences	Social engineering, Elevation of privileges, Intra-platform interference, Information leakage
Software (SaaS)	Remote applications, Micro-services, RESTful services	System hardening, Software defences	Social engineering, Intra-platform interference

Figure 2: Summary of features, security measures and risks

One positive effect of this growth in unauthorized data access is the raised awareness of digital forensics (DF) and a marked change in its perception from a solely post-event reactive investigative tool to a pro-active policy to establish intelligence capabilities in advance of any incidents. This change in role reflects the concept of digital forensic readiness. Thus, "Pro-active DF management must ensure that all business processes are structured in such a way that essential data and evidence will be retained to ensure successful DF investigations, should an incident occur" [27, p.18].

Naturally, this concept of digital forensic readiness is equally applicable to Cloud systems and novel techniques have been proposed to facilitate the data collection that this entails [28]. Yet, the Cloud context introduces particular problems with respect to forensic readiness.

V. CLOUD FORENSIC RECOVERY

Forensic readiness in the Cloud is complicated by the variety of contexts in which Cloud services are deployed and the diversity of software settings in which security risks may

arise. Forensic readiness must accommodate these complexities and, in turn, this suggests that a single infrastructure-based digital forensic readiness solution may be infeasible.

The primary reason for concern is the need to capture relevant data on system operation at the various operational levels of the Cloud system and any potential interaction across these levels. This means capturing program logs, system logs and user activity logs. In any end-customer Cloud facility, the data protected may not extend beyond any currently live information and data held in associated database systems. The ready recycle capability of Cloud services also has implications for the persistence of digital forensic evidence. An intrusion that steals data from a virtual machine and then seeks to reset that machine may well succeed in destroying evidence of the intrusion, thereby removing any forensic traceability on the nature and quantity of stolen data.

Neither is it sufficient to provide each distinct operational layer of Cloud systems with its own comprehensive forensic readiness. At best, this condition will allow for forensic data recovery for that operational layer. But there is no one-size-fits-all solution that can capture all state, interaction and performance data such as would ensure full Cloud forensic recovery. In fact, this insight reveals a fundamental problem that may impact upon Cloud forensic readiness.

There are parallels here with issues in distributed systems and software architecture. Thus, "distributed software systems are harder to debug than centralized systems due to the increased complexity and truly concurrent activity that is possible in these systems" [29, p. 255]. Regardless of whether the Cloud setting is truly distributed in its realisation, its interconnected software functional layers represent a unique challenge when attempting to interpret the relationship between events or changes actioned at one functional level and the operational impact of such changes on other functional aspects of the services afforded by that Cloud.

When considering Cloud systems, from the perspective of software architecture there may be an assumption of 'a component- and message-based architectural style' in which there is 'a principle of limited visibility or substrate independence: a component within the hierarchy can only be aware of components "above" it and is completely unaware of components which reside "beneath" it' [30, p.825].

This multi-level interpretation problem is complicated by the fact that events considered anomalous at one level of service offering may arise through actions considered legitimate at a 'lower' level of software implementation. From the digital forensic readiness perspective, this underlines the requirement to go beyond capture of significant events across the Cloud service software and functional levels, since significance is an aspect that may cross the boundaries between such layers in the system as a whole. A hypothetical example may clarify this issue.

A CSP customer may contract access to specific functional components (e.g., a Web service). The operational characteristics of the service are under the control of the CSP and not the customer. An authorised employee of the CSP may modify the algorithmic process and thereby affect the outcome of any service use by the customer. While a change in operational behaviour of the service (i.e., an anomaly) may eventually be detected by the customer, there may be no anomalous activity evident at the level of CSP employee activity. The focus of subsequent forensic investigation may light initially on the nature of customer activity, since this is where the anomaly is apparent, but proper understanding of the issue requires that events across different functional levels of the Cloud system be apprehended.

An informative view on this issue may be borrowed from Granular Computing [31], which aims to develop computational models of complex systems, such as human intelligence. A key characteristic of this work is that it 'stresses multiple views and multiple levels of understanding in each view' [31, p.85]. Here, the emphasis is upon 'holistic, unified views, in contrast to isolated, fragmented views. To achieve this, we need to consider multiple hierarchies and multiple levels in each hierarchy' [op. cit., p.88].

Our proposal for adequate Cloud forensic readiness has two components. Firstly, is the requirement for data capture. This is the obvious need to record any data at each layer of Cloud facility that may have a role to play in subsequent digital forensic analysis. Secondly, the captured data must be stored off the system being monitored in a manner that both ensures the integrity of the logging and minimises the likelihood that the stored data can be compromised, either as a result of hostile action or 'friendly fire'.

To achieve adequate data capture, we require 'state information' and data management across differing levels of any Cloud service, from the lowest software level up to the most abstracted 'user facing' software component. On their own, such records will not be sufficient to fully capture the potential interplay of differing software levels. For this purpose, subsequent digital forensic analytics will be required in order to establish a multi-dimensional representation of event chronology. This means that timestamps from events and data captured at different software levels of abstraction will be correlated to determine how events across the Cloud system are related.

Our requirement for secure and resilient log storage can build upon default system logging that will be present within the Cloud implementation but this must be supplemented to achieve log reliability.

Instead of using centralised log servers, which of course are attractive targets and easy to spot for attackers, we propose a different approach. In order to prevent adversaries from manipulating log files to hide their tracks, we use chained Message Authentication Codes for each entry to the log file on each node. If state-of-the-art MACs are used, this makes it impossible to delete or manipulate text in the log files. Next, each node uses secret sharing techniques as proposed by Adi Shamir [32] to divide the log file into parts. These parts are then sent to random other nodes which store these log data. Even if an adversary succeeds in taking over some of the nodes, he will need a certain number of these fragments to reconstruct the log data. But since for each log entry different nodes are chosen randomly as stated before,

the attacker effectively needs to control the whole Cloud ecosystem to stay hidden. Further information on this solution can be found in our previous paper [33].

VI. Conclusions

As organisations move increasingly away from locally hosted computer services toward Cloud-platforms, there is a corresponding need to ensure the forensic integrity of such instances. The primary reasons for concern are (i) the locus of responsibility, and (ii) the associated risk of legal sanction and financial penalty. In the first place, while Cloud service providers (CSPs) are responsible for the availability and robustness of their commercial offerings, they will not be responsible for the management of such services by their customers, nor for the data security associated with customer-level use of the Cloud services. Responsibility for these aspects resides with the CSP's customers, whose data processing and data management are built upon the purchased Cloud services. In the second place, legislative demands on data protection, such as the forthcoming EU General Data Protection Regulation, will require companies to notify all breaches within 72 hours of discovery, or face significant financial penalty.

These concerns can be addressed and the business risk mitigated through development of forensic readiness in customer-level Cloud systems. We have argued that this requires a range of logging and data capture facilities across the Cloud system software infrastructure that maintain the possibility of tracking activity at different levels of software abstraction (the multi-level interpretation problem). Our second proposition is that such digital forensic readiness must be combined with techniques to ensure that logged data is incorruptible and robust. We have previously proposed techniques for intrusion monitoring that ensure log data credibility and provide robust decentralised log storage and recovery [33] for post-hack scenarios.

References

[1] http://bt.custhelp.com/app/answers/detail/a_id/41948/~/what-is-bt-cloud%3F [retrieved: December, 2017]

[2] Nanavati, M., Colp, P., Aiello, B., & Warfield, A. (2014). Cloud security: A gathering storm. Communications of the ACM, 57(5), 70-79.

[3] Tirumala, S. S., Sathu, H., & Naidu, V. (2015, December). Analysis and prevention of account hijacking based incidents in cloud environment. In Information Technology (ICIT), 2015 International Conference on (pp. 124-129). IEEE.

[4] Zhou, M., Zhang, R., Xie, W., Qian, W., & Zhou, A. (2010, November). Security and privacy in cloud computing: A survey. In Semantics Knowledge and Grid (SKG), 2010 Sixth International Conference on (pp. 105-112). IEEE.

[5] Chen, Y., Paxson, V., & Katz, R. H. (2010). What's new about cloud computing security. University of California, Berkeley Report No. UCB/EECS-2010-5 January, 20(2010), 2010-5.

[6] BBC News, 2014. http://www.bbc.co.uk/news/technology-29076899

[7] Takabi, H., Joshi, J. B., & Ahn, G. J. (2010). Security and privacy challenges in cloud computing environments. IEEE Security & Privacy, 8(6), 24-31.

[8] Carroll, M., Van Der Merwe, A., & Kotze, P. (2011, August). Secure cloud computing: Benefits, risks and controls. In Information Security South Africa (ISSA), 2011 (pp. 1-9). IEEE.

[9] K. Ilgun, R. A. Kemmerer and P. A. Porras, "State transition analysis: A rule-based intrusion detection approach", IEEE transactions on software engineering, vol. 21, no. 3, pp. 181-199, 1995.

[10] P. Garcia-Teodoro, J. Diaz-Verdejo, G. Maciá-Fernández and E. Vázquez, "Anomaly-based network intrusion detection: Techniques, systems and challenges", Computers and Security, vol. 28, no. 1, pp. 18-28, 2009.

[11] P. Mell and T. Grance, "The NIST definition of cloud computing", NIST, 2011. Available from http://faculty.winthrop.edu/domanm/csci411/Handouts/NIST.pdf, [retrieved: February, 2017].

[12] Gibson, J., Rondeau, R., Eveleigh, D., & Tan, Q. (2012, November). Benefits and challenges of three cloud computing service models. In Computational Aspects of Social Networks (CASoN), 2012 Fourth International Conference on (pp. 198-205). IEEE.

[13] Sweetland Edwards, H., How Web Cams Helped Bring Down the Internet, Briefly, Time Magazine, 25th October 2016, http://time.com/4542600/internet-outage-web-cams-hackers/ [retrieved: December, 2017]

[14] Zhang, S., Zhang, X., & Ou, X. (2014, June). After we knew it: empirical study and modeling of cost-effectiveness of exploiting prevalent known vulnerabilities across iaas cloud. In Proceedings of the 9th ACM symposium on Information, computer and communications security (pp. 317-328). ACM.

[15] Dhakate, S., & Godbole, A. (2015, December). Distributed cloud monitoring using Docker as next generation container virtualization technology. In India Conference (INDICON), 2015 Annual IEEE (pp. 1-5). IEEE.

[16] Pahl, C., & Lee, B. (2015, August). Containers and clusters for edge cloud architectures--a technology review. In Future Internet of Things and Cloud (FiCloud), 2015 3rd International Conference on (pp. 379-386). IEEE.

[17] Balalaie, A., Heydarnoori, A., & Jamshidi, P. (2016). Microservices architecture enables DevOps: migration to a cloud-native architecture. IEEE Software, 33(3), 42-52.

[18] Villamizar, M., Garces, O., Ochoa, L., Castro, H., Salamanca, L., Verano, M., ... & Lang, M. (2016, May). Infrastructure cost comparison of running web applications in the cloud using AWS lambda and monolithic and microservice architectures. In Cluster, Cloud and Grid Computing (CCGrid), 2016 16th IEEE/ACM International Symposium on (pp. 179-182). IEEE.

[19] Namiot, D., & Sneps-Sneppe, M. (2014). On micro-services architecture. International Journal of Open Information Technologies, 2(9), 24-27.

[20] Han, H., Kim, S., Jung, H., Yeom, H. Y., Yoon, C., Park, J., & Lee, Y. (2009, September). A RESTful approach to the management of cloud infrastructure. In Cloud Computing, 2009. CLOUD'09. IEEE International Conference on (pp. 139-142). IEEE.

[21] Zhang, Y., Juels, A., Oprea, A., & Reiter, M. K. (2011, May). Homealone: Co-residency detection in the cloud via side-channel analysis. In Security and Privacy (SP), 2011 IEEE Symposium on (pp. 313-328). IEEE.

[22] Zhang, Y., Juels, A., Reiter, M. K., & Ristenpart, T. (2014, November). Cross-tenant side-channel attacks in PaaS clouds. In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security (pp. 990-1003). ACM.

[23] Varadarajan, V., Kooburat, T., Farley, B., Ristenpart, T., & Swift, M. M. (2012, October). Resource-freeing attacks:

improve your cloud performance (at your neighbor's expense). In Proceedings of the 2012 ACM conference on Computer and communications security (pp. 281-292). ACM.

[24] Zhang, Y., Juels, A., Reiter, M. K., & Ristenpart, T. (2012, October). Cross-VM side channels and their use to extract private keys. In Proceedings of the 2012 ACM conference on Computer and communications security (pp. 305-316). ACM.

[25] Vateva-Gurova, T., Suri, N. and Mendelson, A., The Impact of Hypervisor Scheduling on Compromising Virtualized Environments, 2015 IEEE International Conference on Computer and Information Technology; Ubiquitous Computing and Communications; Dependable, Autonomic and Secure Computing; Pervasive Intelligence and Computing (CIT/IUCC/DASC/PICOM), 2015 (pp. 1910-1917).

[26] Security Week, 2015, Data Breaches by the Numbers, http://www.securityweek.com/data-breaches-numbers [retrieved: December, 2017]

[27] Grobler, C. P., & Louwrens, C. P. (2007, May). Digital forensic readiness as a component of information security best practice. In IFIP International Information Security Conference(pp. 13-24). Springer, Boston, MA.

[28] Kebande, V. R., & Venter, H. S. (2014). A Cloud Forensic Readiness Model Using a Botnet as a Service. In The International Conference on Digital Security and Forensics (DigitalSec2014) (pp. 23-32). The Society of Digital Information and Wireless Communication.

[29] Bates, P. C., & Wileden, J. C. (1983). High-level debugging of distributed systems: The behavioral abstraction approach. Journal of Systems and Software, 3(4), 255-264.

[30] Medvidovic, N., Taylor, R. N., & Whitehead Jr, E. J. (1996). Formal modeling of software architectures at multiple levels of abstraction. ejw, 714, 824-2776.

[31] Yao, Y. (2005, July). Perspectives of granular computing. In Granular Computing, 2005 IEEE International Conference on (Vol. 1, pp. 85-90). IEEE.

[32] Shamir, A., How to share a secret, Communications of the ACM, vol. 22, no. 11, pp. 612-613, 1979.

[33] Weir, G.R.S. and Aßmuth, A., Strategies for Intrusion Monitoring in Cloud Services, Cloud Computing 2017, The Eighth International Conference on Cloud Computing, GRIDs, and Virtualization, IARIA, Athens, Greece, February 2017.