InterPARES Trust Project

Research Report

Title:	Ensuring trust in storage in Infrastructure-as-a-Service (IaaS) (EU08)
Status:	Draft (restricted)
Version:	1
Date submitted:	8 August 2015
Last reviewed:	8 August 2015
Author:	InterPARES Trust Project
Writer(s):	Hrvoje Stancic, Faculty of Humanities and Social Sciences, University of Zagreb Edvin Bursic, Financial Agency (FINA) and GRA, Faculty of Humanities and Social Sciences, University of Zagreb Adam Al-Hariri, GRA, Faculty of Humanities and Social Sciences, University of Zagreb
Research domain:	Infrastructure
URL:	

Document control

Version history			
Version	**Date**	**By**	**Version notes**
0.1	28 February 2015	All	Preliminary draft
0.2	14 May 2015	All	First draft
1.0	8 August 2015	Hrvoje Stancic	Final draft submitted for feedback and approval

Contents

INTRODUCTION

IntePARES Trust project approved the research on ensuring trust in storage in Infrastructure-as-a-Service (IaaS). Selected IaaS providers in Croatia were comparatively analysed in terms of availability of security policies. Identification and analysis of the security risks and minimising or eliminating them in the present or future IaaS implementations will ensure that the users can gain trust in storage in Infrastructure-as-a-Service.

Research timeline: 15 January to 15 July 2014.

This research involved two graduate research assistants – one PhD student and one graduate level student – with the goal to collect the required information related to the policies and standards addressing storage in IaaS as well as to produce the relevant comparative analysis.

Project results were disseminated (in chronological order):

1. **Workshop** of Croatian InterPARES Trust Team organized by project partner Digital Information-documentation Office of the Government of the Republic of Croatia, 28 March 2014, Zagreb, Croatia

2. Stančić, Hrvoje; Al-Hariri, Adam; Adžaga, Ivan. **Ensuring Trust in Storage in Infrastructure-as-a-Service – progress report**, InterPARES Trust – Joint European Team Research Workshop, Stockholm, 15-16 May, 2014

3. Stančić, Hrvoje; Al-Hariri, Adam; Buršić, Edvin, **Archival Approach to IaaS Cloud Services**, in: Hunjak, Tihomir; Lovrenčić, Sandra; Tomičić, Igor (Eds.), Central European Conference on Information and Intelligent Systems, Faculty of Organization and Informatics Varaždin University of Zagreb, Varaždin, 2014, pp. 216-222

4. Stančić, Hrvoje, **Report on the InterPARES Trust Project**, in: Babić, Silvija (Ed.), Dostupnost arhivskoga gradiva, Hrvatsko arhivističko društvo, Vinkovci, 2014, pp. 521-527 (published paper presented at the 47th Symposium of Croatian Archival Society, **Availability of archival material**, 22-24 October 2014, Vinkovci, Croatia)

5. InterPARES Trust **visibility event** - **Presentation of InterPARES Trust research results**, organised by project partner Faculty of Humanities and Social Sciences, University of Zagreb, Croatia, 21 November 2014

6. Stančić, Hrvoje, **Project InterPARES Trust – project activities**, 18th seminar Archives, Libraries, Museums – possibilities of cooperation in environment of global information infrastructure, 26-28 November 2014, Rovinj, Croatia

7. Stančić, Hrvoje, **Electronic trust** (interview), in: Römer János (ed.), Radio show *From the world of science*, Radio Sljeme, Croatia, aired on 19 March 2015 (duration: 20 min)

8. Stančić, Hrvoje. **Ensuring trust in storage in Infrastructure-as-a-Service – discussion on the findings with the Deputy Minister of Public Administration for e-Croatia**, Ministry of the Public Administration, 21 July 2015

RESEARCH

Research methodology

The research was divided in four stages: (1) Identification, (2) Data acquisition, (3) Analysis, and (4) Interpretation. The research was limited to the EU region with the focus on Croatia.

1. Identification

In the research Ensuring trust in storage in Infrastructure-as-a-Service (IaaS), the researchers looked for the minimum amount of information which would provide users' trust in the service and also position a service provider as a trusted service provider.

According to the US National Institute of Standards and Technology (NIST), cloud computing is: "A model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g. networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model is composed of five essential characteristics, three service models, and four deployment models."[1]

Project Records in the Cloud[2] identifies five essential characteristics of cloud solutions:

1. *On-demand self-service* allows users to access as many computing capabilities as they need
2. *Broad network access* allows users to access the cloud from any machine that has an Internet connection
3. *Resource pooling* allows the multi-tenant model supporting multiple users at the same time
4. *Rapid elasticity* allows users to change the amount of computing resources they need at any time
5. *Measured service* allows precise measuring of utilised resources in terms of storage, processing, bandwidth etc. These resources can be monitored, controlled and reported to the users, who are only charged for what they use by pay-as-you-go model. In most cases this approach reduces costs.

Stancic, Rajh and Milosevic[3] differentiate between three service models as follows:

1. *Software as a Service (SaaS)* – ability to deliver applications from cloud-based physical infrastructure, accessible via various client software tools or devices. The

[1] Mell, Peter; Grance, Timothy. The NIST Definition of Cloud Computing. NIST Special Publication 800-145, National Institute of Standards and Technology, Gaithersburg, September 2011, p. 2, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf (August 1 2015).

[2] Duranti, L. *Records in the Cloud: Detailed Description*, http://www.recordsinthecloud.org/secure/documents, (April 8 2014).

[3] Stancic, H; Rajh, A; Milosevic, I. "Archiving-as-a-Service", Influence of Cloud Computing on the Archival Theory and Practice. In Duranti, L; Shaffer, E. (Eds.), *The Memory of the World in the Digital Age: Digitization and Preservation*, pp. 108-125, Vancouver, Canada, 2012.

user has no awareness or control of the underlying physical components or software configuration capabilities outside the delivered application.

2. *Platform as a Service (PaaS)* – ability to deliver complete environments (operating systems and required tools) for testing or development of external applications. The user, however, has no control over the configuration settings of the application-hosting environment.
3. *Infrastructure as a Service (IaaS)* – ability to deliver complete virtual data centres to the user who is then able to configure and deploy virtual machines and other relevant/corresponding virtual components according to their personalized requirements.

Regarding the four deployment models Stancic at al.[4] further state that cloud implementations include:

1. *Private cloud* where it is implied that the cloud infrastructure is built and provisioned for private use by a single organization. Private clouds in practice tend to be service-oriented with specific roles and requirements.
2. *Community cloud* where the physical infrastructure is implemented, administered, and operated by several organizations in a certain community of consumers from organizations that have shared goals and requirements.
3. *Public cloud* where the cloud infrastructure is intended for "rent" by the public users, as delegated by the provider usually for profit or other means of compensation for the provider.
4. *Hybrid cloud* which is the combination of two or more physical cloud infrastructures from different branches of the above listed deployment models that are physically separate but are connected via the means of mutual data and application portability or management hierarchies.

In InterPARES Trust's project terminology database the term "trust" is defined as "confidence of one party in another, based on alignment of value systems with respect to specific actions or benefits, and involving a relationship of voluntary vulnerability, dependence and reliance, based on risk assessment".[5] This means that the users of cloud services should have enough information on a particular service (e.g. in Terms of Service) in order to trust it, or the service level agreement (SLA) between users and cloud service provider (CSP) should equally protect interests of both parties involved.

After the initial research space defined we tried to define the questions that customers or clients would naturally ask before exploiting a (trusted) service. For example:

1. What should you consider when purchasing a Cloud service?
2. Is there enough information that could guarantee your trust in the service?

[4] Stancic, H; Rajh, A; Milosevic, I. "Archiving-as-a-Service", Influence of Cloud Computing on the Archival Theory and Practice. In Duranti, L; Shaffer, E. (Eds.), *The Memory of the World in the Digital Age: Digitization and Preservation*, pages 108-125, Vancouver, Canada, 2012.

[5] Project *InterPARES Trust: Trust and Digital Records in an Increasingly Networked Society*, http://interparestrust.org.

However, these are the questions that arise whenever a new technological sulutions surface at the market. To better understand the implication of these questions to the cloud solutions in particular we decided to create a questionnaire and survey the cloud service providers in search for answers. The questions were organized in 10 categories following the reasoning of the NIST expert team:

1. General information
2. Governance
3. Compliance
4. Trust
5. Architecture
6. Identity and access management
7. Software isolation
8. Data protection
9. Availability
10. Incident response.

The initial set consisted of 54 questions. Than, the partners have reviewed them and narrowed the set down to 36 questions. The questions were:

Category	**Questions**
1. General information	1. Which components are used in IaaS? 2. What types of services are offered in IaaS? 3. What technologies are being used? 4. What implications used technologies have on security and privacy of the system?
2. Governance	5. Is it possible for a client to monitor security of computing environment and data security? How? 6. What kind of security assures a client that his data is not mixed with another's? 7. What kind of security assures a client that there is no data shared with employees of different rank or/and not created by others? 8. What audit mechanisms and tools are used to determine how data is stored, protected and used to validate services, and to verify policy enforcement?
3. Compliance	9. Does the service comply with other countries' laws, regulations, standards and specifications for clients outsde the country of service? 10. How is the service secured against unauthorized access, use, disclosure, disruption, modification, or destruction of data? 11. What technical and physical safeguards does the service assure? 12. Does the service use subcontractors for any part of the used technology or offered service?

Category	Questions
4. Trust	13. Is the service secured from denial of service attack? 14. Does the service secure ownership rights over data? 15. Does the service have any certificate relevant to the service? 16. What kind of risk management does the organization provide? 17. What kind of physical and logical security is assured for the virtual servers and applications?
5. Architecture	18. How is a hypervisor or virtual machine monitor secured? 19. How does the service secure virtual machine images from attack looking for proprietary code and data? 20. Does the service use image management process to govern the creation, storage, and use of virtual machine images? 21. How does the service secure from attacks on the client side? 22. How does the service secure from attacks on the server side? 23. Is the service using encrypted network exchange?
6. Identity and Access Management	24. How does the service protect ancillary data: details about the consumers' accounts, data about customer-related activity, data collected to meter and charge for consumption of resources, logs and audit trails, and other such metadata that are generated and accumulated within the environment, data of an organization's initiative (e.g., the activity level or projected growth of a startup company), metadata collected by the provider?
7. Software Isolation	25. How does the service prevent man-in-the-middle attacks? 26. Is the service secured from attacks on the server that target passwords?
8. Data Protection	27. What kind of encryption does the servie use to secure data stored in IaaS? 28. Have the service conducted deliberate attacks in order to test the system's protection? 29. What procedures are used for data sanitization upon termination of service, i.e. how does the service ensure that the data after deletion are not recoverable? 30. Where, geographically, are the data stored? 31. Where, geographically, is data backup stored?
9. Availability	32. In a situation of a lawful raid how is the service availability assured to the users not being lawfuly raided? 33. Is there a policy regarding user data availability in case of a bankruptcy or other facility loss and how is it defined?
10. Incident Response	34. Is there an incident response plan and how is it defined? 35. Does the service keep track of the data using which the scope of the incident, and assets affected can be determined? 36. Does the service keep a forensic copy of incident data for legal proceedings or as needed by the consumer? Or, does the servce give incident data to the consumers?

Next, each category will be breafly explained.

1. General information

Questions grouped under general information category should provide answers to some core questions about Infrastructure-as-a-Service implemented by the surveyed CSPs. Beside some basic information about the CSPs, we narrowed the research goals to the three specific areas being storage, service and networks[6]. We investigated what technologies are being used, and how the implemented storage, servers, networks and technology affect security and privacy.

2. Governance

Governance is the key factor in assuring security over data produced by a company. In this category we examined how can user verify integrity of data stored by a CSP and how can user keep track of computer environment security. We have also examined how CSPs ensure that the data from different users are not mixed. Finally, we questioned the usage of prescribed relevant procedures, rule books and internal policies.

3. Compliance

For a company considering IaaS it is important to be aware of the fact by which laws the CSP is governed by, where is geographically the data stored, and is any part of the service subcontracted. Along with those critical questions, we also examined what are technical and physical measures of protection which secure service from unauthorised access, usage, discovery, interruption, alteration and termination of data.

4. Trust

This category was thought of as to be the most important for the non-expert users that may read this document. It provides the fundamental questions we discovered to be the most important and the most interesting to users when chosing a trustworthy CSP. We wanted to know if any risk management systems were implemented, and what kind of physical and logical security were set up for virtual servers and applications. Another concern regarding trust in the service was connected with the issues of ownership of the data given to the custody of CSP and how are the data protected from employees' of the CSP. We also questioned the existence of any relevant certificates implemented, such as: ISO 27001:2005, ISO 9001:2008, TIA, EU or NATO-relevant certificates. Also, there is a matter of protection against various attacks such as DoS (Denial of Service) or DDoS (Distributed DoS) attacks, man-in-the-middle attacks and various server attacks.

5. Architecture

Since the hardware and software architecture used to deliver cloud services can vary significantly, the actual set-up can have repercutions to the security. Therefore we wanted to investigate what type of solutions CSPs have implemented. Regarding the possible attacks on architecture, we examined measures of protection against attack on hypervisor, virtual machine monitor, images, proprietary code, client (on user's computer) and server. We also

[6] As in *Understanding the Cloud Computing Stack: SaaS, PaaS, IaaS*, http://www.rackspace.com/knowledge_center/whitepaper/understanding-the-cloud-computing-stack-saas-paas-iaas, (April 4 2014).

wanted to find out whether the process of virtual machine images management is used or not.

6. Identity and Access Management

Data sensitivity and privacy of information have always been the area of concern for organizations. The identity proofing and authentication aspects of identity management entail the use, maintenance, and protection of PII (personally identifiable information) collected from users. The data collected by the provider of the purchased service include details about the accounts of consumers, data about customer-related activity, data collected to meter and charge for consumption of resources, logs and audit trails, and other metadata that are generated and accumulated within the environment, data of an organization's initiative (e.g., the activity level or projected growth of a startup company), as well as metadata collected by the CSP.

In the wrong hands, the loss of these data can be damaging to the clients' business. All of the above mentioned data – ancillary data – should be protected, and clients given the assurance of the protection.

7. Software Isolation

In order to achieve the flexibility of on-demand services, cloud service providers have to use high degrees of multi-tenancy over the large number of platforms. The multi-tenancy in IaaS cloud computing environment is typically done by multiplexing the execution of virtual machines from potentially different consumers on the same physical server.

Multi-tenancy in virtual machine-based cloud infrastructure, together with the way physical resources are shared between the guest virtual machines, give raise to new sources of threat. In the man-in-the-middle attack, the intruder uses a program that appears to be the server to the client and at the same time it appears to be the client to the server. During the attack, and with no knowledge or suspicion, the client is giving the attacker his/her password on a silver pladder. The attack may be used simply to gain access to the message, or enable the attacker to modify the message before retransmitting it. So, if the administrative control of guest virtual machines is obtained then the man-in-the-middle attack can be used to modify the code used for authentication. Once the code is modified, the attacker has access to clients' data.

8. Data Protection

Data protection should be applied to data-at-rest, data-in-transit and data-in-use. The data-at-rest can easily be protected by an encryption mechanism, while the data-in-transit are much harder to encrypt. A fully homomorphic encryption scheme allows data to be processed without being decrypted.[7]

[7] "The aim of homomorphic cryptography is to ensure privacy of data in communication, storage or in use by processes with mechanisms similar to conventional cryptography, but with added capabilities of computing over encrypted data, searching an encrypted data, etc. Homomorphism is a property by which a problem in one algebraic system can be converted to a problem in another algebraic system, be solved and the solution later can also be translated back effectively. Thus, homomorphism makes secure delegation of computation to a third party possible. (...) Fully Homomorphic Encryption combines security with usability. It can help preserve

Making sure that the data stays computationaly accurate is important as well as the preservation of its integrity. The data remanence – the residual representation of data that has been at some point nominally erased or removed – should also be addressed. For specific information about how the data security should be achieved, CSPs should refer to the National Institute of Standards and Technology's (NIST) Special Publication, 800-88, Guidelines for Media Sanitization.[8]

For the data protection techinques one should consider the recognized standards of encryption, e.g. the NIST's Federal Information Processing Standards (FIPS).[9] The encryption key length should be considered, too. The key lengths used should minimally be 112-bit for the Triple DES (Data Encryption Standard) and minimally 128-bit for AES (Advanced Encryption Standard). Higher AES key sizes (192-bit, 256-bit) would also be appropriate, but performance may be slower.[10]

In a public cloud computing environment, the data from one consumer are physically collocated (e.g., in an IaaS data store) with other customers' data, which can complicate things. There are many examples of researchers, obtaining used drives from online auctions and other sources, recovering large amounts of sensitive information. With the proper skills and equipment, it is also possible to recover data from the failed drives if they are not disposed of properly. More detailed information about sanitization is given in the NIST's Guidelines for Media Sanitization.

It is also important to keep in mind that the provider of a service can be located in one country and that its data center(s) can be located in another, thanks to the versatility of the Internet. Therefore, it is important to understand the possible implications this situation has on the data being stored and backed-up for the disaster recovery purposes. The information on geographic location(s) where the data are stored and backed-up can result in increase or decrease of trust regarding the security of organization's data. It is also important to understand the laws and regulations of the country where the data are being stored because it could be relevant in case of a legal action or in case of loss and recovery of data.

9. Availability

In the year 2012, Megaupload servers were raided. And all assets were seized. The government insisted that all user data, even the legitimate data, should be destroyed. More

---

customer privacy while outsourcing various kinds of computation to the cloud, besides storage." Sharma, Iti. Fully Homomorphic Encryption Scheme with Symmetric Keys, Dissertation, Department of Computer Science & Engineering, University College of Engineering, Rajasthan Technical University, Kota, August 2013, pp. 3, 15, http://arxiv.org/ftp/arxiv/papers/1310/1310.2452.pdf (August 1 2015)

[8] Kissel, Richard; Regenscheid, Andrew; Scholl, Matthew; Stine, Kevin, Guidelines for Media Sanitization, NIST Special Publication 800-88, Revision 1, December 2014, http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf (August 1 2015)

[9] Federal Information Processing Standards Publications (FIPS PUBS), http://csrc.nist.gov/publications/PubsFIPS.html (August 1 2015)

[10] Barker, Elaine; Barker, William; Burr, William; Polk, William; Smid, Miles, Recommendation for Key Management – Part 1: General (Revision 3), NIST Special Publication 800-57, National Institute of Standards and Technology, Gaithersburg, July 2012, http://csrc.nist.gov/publications/nistpubs/800-57/sp800-57_part1_rev3_general.pdf (August 1 2015)

than 50 million users were at risk of loosing their data. More interesting is that New Zealand arrested the founder based on U.S. accusation. In need of a solution to the piracy problem something needed to be done.

Of course, many things can happen to a business as a consequence of decisions made by the service provider. The provider can have facility damage or loss due to many either natural disasters or human-influenced errors. Or, it can run out of business, go bankrupt, or have other kind of financial difficulties. However, this should not be concern of the consumers, so that is why we investigated the existence of a policy that can be enforced if such occasion occurs. On the other hand, every organization should have a contingency plan. If an organization relies on IaaS for data storage and processing, it needs to include in their contingency plan solutions to prolonged and permanent system disruptions, especially with mission critical operations untill the restoration of the service.

10. Incident response

Incident response involves an organized method for dealing with the consequences of an attack against the security of a computer system. The cloud provider's role is vital in performing incident response activities, including incident verification, attack analysis, containment, data collection and preservation, problem remediation, and service restoration. Each layer in a cloud application stack generates event logs. These data are accessible and under the control of the cloud provider. The plan should cover the restoration of the service ASAP and determination of the scope of the incident and assets affected. Also, the incident response plan sholud cover the situations when a repair of the security breach is needed. The cloud service provider should make sure it does not happen again.

Conclusion

The set of 36 questions divided into 10 categories was considered as sufficient to provide enough information on the IaaS model of a cloud service in order for the users to consider the service as responsible, reliable, accurate, secure, transparent and trustworthy as well as that it considers privacy issues, duties to remember (i.e. digital preservation), and the right to be forgotten (i.e. safe deletion).

2. Data acquisition

In the second stage of the research, the developed questionnaire was used to gather information on the IaaS cloud services offered by 10 national CSPs in Croatia. The online questionnaire was realized and sent to the 10 identified CSPs. Although the response rate was 30%, which might seem a lot but in reality that ment that 3 out of 10 CSPs fully answered the questionnaire, we decided to continue to analyze the gathered data. The reasons why the other CSPs did not answer the questionnaire remain unknown.

3. Analysis

This stage followed the data acquisition stage and was, during certain period of time, overlapping with it. The researchers comparatively analysed the returned finlled out questionnaires. Following this stage, the results were interpreted and are shown in the next section.

4. Interpretation

During the last stage, the results of the analysis were interpreted and the recommendations were formulated. These are shown grouped by the research categories.

1. General information

The respondents were highly positioned individuals in organizations such as CEOs and assistant directors. This part of the questionnaire was developed in order to gather information on some core questions about Infrastructure-as-a-Service implemented by the surveyed CSPs. Therefore, we consider that respondents were qualified enough to answer complex questions about service they provide. Considering the services being offered in the cloud we found that the respondents mostly provide all of the standard service models – IaaS, SaaS and PaaS. Also we discovered that services are provided to both legal and natural persons. Considering components used in IaaS mostly there were redundant storage and server solutions from verified hardware companies, which are pretty standard and common for CSPs. Types of services provided in IaaS are usually simple, e.g. rent of infrastructure, but one of the respondening CSPs stands out with bundle of free additional services such as virtual server, console access, preinstalled images, daily back-up, traffic monitor, secure monitoring of network traffic, twice per year safety scanning, possibility to expand resources and help-desk service. Considering technologies used, virtualization was the only one. All of the respondents mentioned that implemented hardware and software solutions provide total separation of clients and reservation of guaranteed resources.

2. Governance

The usual answer on the question if clients can verify integrity of data stored with CSP and keep track of computer environment security was that clients have complete autonomy and responsibility over data in their own virtual server. Considering the efforts that data from different clients are not mixed, the respondents answered that with virtualization

technology and logical and physical separation of resources they assure the mixing of clients' data is avoided. CSPs claim that they follow international standards, e.g. ISO 27001.

3. Compliance

Considering legal questions, all of the respondents claim to be working according to the current Croatian and international legal regulations. Geographically the data are stored in Croatia, while one provider also offers storage with their partners outside of Croatia, but only on request. Physically, facilities are under 24/7 protection by security guards. Technically, facilities are secured by several independent systems, and the server space is constantly being monitored. None of the respondents use subcontractors.

4. Trust

Two of three respondents refer to ISO 27001:2005 in relation to risk management. The remaining one did not implement ISO 27001. Regarding the physical and logical security of virtual servers and applications the research team received almost the same answer as it was for physical and technical protection, while only one respondent mentioned separation on disc level and different layers of network infrastructure. Ownership of the data is regulated by contract, except with one CSP which provides free service, and where ownership of the data is regulated by a rule book. All of the CSPs have employment contract clause about data confidentiality. Considering possible relevant certificates, one of the respondents excels with NATO and EU certificates. The surveyed CSPs usually provide fair protection against various attacks. One provider does not have protection against denial of service (DoS) attack, while others have.

5. Architecture

Regarding the possible attacks on architecture, we examined protection against attack on hypervisor. The surveyed CSPs generally limit and monitor access to hypervisors, while one is also monitoring integrity of the used systems. Images and proprietary code are protected by logical unit number (LUN) where each username have its own dedicated disc space, or by a controlled interface for images. None of the CSPs offer any client-side protection by default. One of them offers its own solution by request. Two out of the three surveyed CSPs use virtual machine image management. Finally, the respondents provided limited answers about usage of encryption, types of encryption and protocols used on SSL and TLS level. Of course, this information can be considered as business-critical, and it is understandable that the surveyed CSPs were reluctant to provide more information. However, they did say that they mostly use encrypted network exchange, and symmetrical and asymmetrical type of encryption.

6. Identity and Access Management

The data sensitivity and privacy of information have always been an area of concern for organizations, i.e. in this context – clients. Specifically, we have asked CSPs how they protect ancillary data that clients produce within their institutions. All of the respondents use some kind of data isolation, either by using a stand-alone system or by limiting access to administrators only.

7. Software Isolation

Considering the man-in-the-middle attack, all of the respondents use some sort of protection such as complex cryptographic algorithms, personal PKI, or protection through SSL. The provider-side server attacks are avoided by using intrusion prevention systems (IPS).

8. Data Protection

Regarding the data protection, the research team examined protection of data-in-transit and data-at-rest. It was found that one provider does not use encryption while others use high end disc systems which provide several types of encryption. Also, most of the respondents conduct deliberate attacks on their systems in order to test the overall security. All providers have a mechanism to retrieve accidentally or intentionally deleted data. While one CSP have this covered in the service level agreement (SLA), others provide up to 90 days to the clients to retrieve such data. Regarding the data sanitization, CSPs usually mention this as the clients' responsibility. Several deletion methods are used on the provider-side, such as zero fill or virtual volume deletion. One provider specifically explained the procedure as follows: “Overwrite all addressable locations with a character, its complement, then a random character and verify.” Considering the geographical placement of storage where the data are stored, all CSPs store the data at the territory of the Republic of Croatia, though one provider states it can store data abroad if needed.

9. Availability

The research team surveyed CSPs about their assurance that the data will always be available to the clients. The questions covered CSPs reactions to a lawful raid. Usually, CSPs are not responsible for the data stored by their clients and, as mentioned before, they do not even have access to it. In case of a court order requiring a client's hardware and/or data all providers are obliged to give in either hardware, which is removed from the data storage facility in extreme cases, or a copy of the clinet's volume. The latter solution is considered as more practial and more elegant solution still providing all needed information. However, if a physical server is seized, CSPs would reallocate resources to maintain service at highest possible level for other clients. In case that CSPs run out of business, go bankrupt, or have other financial difficulties they claim that they would most certainly give the latest copy of the data to the clients. One provider mentioned that some elements of this scenatio are covered by SLA, specifically if in case when clients decided to use dedicated infrastructure. In that case clients would be offered to repurchase the infrastructure, and the price would be set by the national tax administration.

10. Incident response

In case of a security incident all providers can track compromised data and undertake corrective actions to reduce possible damage. Also, all of the respondents record forensic data, and one of the CSPs is preserving them. All of the surveyed CSPs provide access to the forensic data to their clients.

CONCLUSIONS AND RECOMMENDATIONS

This research covered the selected cloud service providers (CSP) offering Infrastructure-as-a-Service (IaaS) and storing clients' data at the teritory of the Republic of Croatia. Ten of them were identified while only three were successfully surveyed. The research team discussed the survey success rate and decided to proceede with the analysis of the results of 3 out of 10 CSPs (30%) in spite of the fact that the sample could be questioned as representative. However, the survey success rate in itself also provides information. It might indicate 1) the level of motivation of CSPs in Croatia to deal with detailed surveys on their internal business processes and procedures. It might show that 2) CSPs did not want to answer the questions because they either did not want to share the information or that would have to admit their weaknesses, or that 3) the questionnaire was too complicated to be answered by one person (although this is not excuse because it was not required, and the CSPs could have organised their staff to gather relevant answers).

The comaprative analysis covered ten categories: 1) General information, 2) Governance, 3) Compliance, 4) Trust, 5) Architecture, 6) Identity and Access Management, 7) Software Isolation, 8) Data Protection, 9) Availability, and 10) Incident Response.

Overall, the results show how the CSPs percive the concept of trust in their service. Some CSPs use disclaimers saying that the sole responsibility of the clients' data is on the clients' side, and some understand that special care needs to be provided to the clients' data and internal business processes and procedures in order to become a trusted CSP.

Next, the selected results of the survey are excerpted.

In the section 1) General information, the comparative analysis showed that virtualisation is dominantely used, and that the CSPs are able to separate clients' data from eachother while guaranteeing resources (service elasticity).

In the section 2) Governance, CSPs indicated that the clients are responsible for the integrity of their data and that they have autonomy and responsibility over the data stored in the allocated virtual server.

In the section 3) Compliance, it was confirmed that primary and secondary storage locations are at the Croatian teritory, i.e. within the reach of the legal authorities of the Republic of Croatia. It was interesting to find out that none of the CSPs used subcontractors. This is important in terms of diminishing and/or mitigating the business-related (operational) risks.

In the section 4) Trust, it was discovered that CSPs relate to ISO 27001 for risk management, and that one also obtained NATO and EU certificates. Physical and logical security of (virtual) servers and applications are implemented. However, one provider does not have protection against DoS attack. Users looking for a trusted CSP should, among otherthings, check the level of security a CSP has implemented.

Regarding 5) Architecture, the surveyed CSPs limit and monitor access to hypervisors. Virtual machine image management is used by two out of three CSPs. Limited answers were provided on usage of encryption, types of encryption and protocols used on SSL and TLS level. However, CSPs did say that they mostly use encrypted network exchange, and symmetrical and asymmetrical type of encryption.

In relation to 6) Identity and Access Management, CSPs use data isolation methods to protect clients' ancillary data.

In the section 7) Software Isolation, it was noted that complex cryptographic algorithms, personal PKI, or protection through SSL were the metods used to prevent man-in-the-middle attacks.

A positive practice was noted regarding 8) Data Protection, where the surveyed CSPs confirmed that they conduct deliberate self-attacks in order to test the overall security of their systems and find potential weak spots. CSPs do have mechanisms to retrieve accidentally deleted data, but interestingly they claim that the clients are responsible for the safe deletion or data sanitization. However, the research team thinks otherwise – CSPs should at least provide mechanisms for safe deletion and guarantee media sanitization if a client deletes the data according to the retention schedule or transfers its data to another CSP. This is certanly the area for improvement.

Regarding 9) Availability, if a court order is issued requiring seizure of a client's data CSPs would give either hardware where the data is stored or a copy of the clinet's volume. The letter is considered as more practical solution since the effect is the same. All CSPs claim that other clients' data would remain available all the time. If a provider goes out of business, CSPs claim that the clients would be offered to repurchase the infrastructure at the price set by the national tax administration. Although this sounds reassuring, the research team considers this as an idealistic scenario which might or might not prove possible in all cases.

Finally, regarding 10) Incident Response, CSPs claim that they record and (one of them) preserve forensic data, and that the clients are granted access to the part concerning their data.

Taking all this into account the research team believes that the trust between the clients and CSPs should be based on providing enough information by the CSPs and the possibility of the clinets to negotiate the needed functionalities. CSPs should also demonstrate their operational sustainability and conformance to the relevant standards. Therefore, the trust in CSPs offering IaaS should be looked upon as a combined socio-technical set of requirements, roles, rules, policies, procedures, best practicies, responsibilities, and responsible governance.

The research team also believes that the developed questionnaire, transformed into a check-list (see Appendix A), can on one side provide guidance for the users looking for a cloud service or deciding between several of them, and on the other side function as guidelines for the cloud service providers on what information about the service they should put online.

References

1. Barker, Elaine; Barker, William; Burr, William; Polk, William; Smid, Miles, Recommendation for Key Management – Part 1: General (Revision 3), NIST Special Publication 800-57, National Institute of Standards and Technology, Gaithersburg, July 2012, http://csrc.nist.gov/publications/nistpubs/800-57/sp800-57_part1_rev3_general.pdf (August 1 2015)
2. Dawoud, Wesam; Takouna, Ibrahim; Meinel, Christoph, Infrastructure as a Service Security: Challenges and Solutions, Informatics and Systems (INFOS), 2010, http://www.researchgate.net/publication/224136774_Infrastructure_as_a_service_security_Challenges_and_solutions (February 3 2014)
3. Duranti, L. Records in the Cloud: Detailed Description, http://www.recordsinthecloud.org/secure/documents (April 8 2014)
4. Federal Information Processing Standards Publications (FIPS PUBS), http://csrc.nist.gov/publications/PubsFIPS.html (August 1 2015)
5. Information Supplement: PCI Data Security Standard (PCI DSS) Cloud Computing Guidelines v. 2.0, Cloud Special Interest Group, PCI Security Standards Council, February 2013, https://www.pcisecuritystandards.org/pdfs/PCI_DSS_v2_Cloud_Guidelines.pdf (March 3 2014)
6. Jansen, Wayne; Grance, Timothy, Guidelines on Security and Privacy in Public Cloud Computing, Special Publication 800-144, National Institute of Standards and Technology, Gaithersburg, December 2011, http://csrc.nist.gov/publications/nistpubs/800-144/SP800-144.pdf (January 20 2014)
7. Kissel, Richard; Regenscheid, Andrew; Scholl, Matthew; Stine, Kevin, Guidelines for Media Sanitization, NIST Special Publication 800-88, Revision 1, December 2014, http://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.800-88r1.pdf (August 1 2015)
8. Mell, Peter; Grance, Timothy. The NIST Definition of Cloud Computing. NIST Special Publication 800-145, National Institute of Standards and Technology, Gaithersburg, September 2011, http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf (August 1 2015)
9. Project InterPARES Trust: Trust and Digital Records in an Increasingly Networked Society, http://interparestrust.org
10. Reynolds, Ed; Greenway, Mateen, Minimize the risk of your cloud-based services, White paper, HP Enterprise Security Services, May 2012, http://h20195.www2.hp.com/V2/GetPDF.aspx%2F4AA4-0150ENW.pdf (February 25 2014)
11. Sharma, Iti. Fully Homomorphic Encryption Scheme with Symmetric Keys, Dissertation, Department of Computer Science & Engineering, University College of Engineering, Rajasthan Technical University, Kota, August 2013, http://arxiv.org/ftp/arxiv/papers/1310/1310.2452.pdf (August 1 2015)
12. Stancic, H; Rajh, A; Milosevic, I. “Archiving-as-a-Service”, Influence of Cloud Computing on the Archival Theory and Practice. In Duranti, L; Shaffer, E. (Eds.), The Memory of the World in the Digital Age: Digitization and Preservation, pp. 108-125, Vancouver, Canada, 2012

13. Understanding the Cloud Computing Stack: SaaS, PaaS, IaaS, http://www.rackspace.com/knowledge_center/whitepaper/understanding-the-cloud-computing-stack-saas-paas-iaas (April 4 2014)

Appendix A – IaaS Checklist

This checklist is based on the questionnaire used during the collection of data for analysis of the Croatian cloud service providers (CSP) offering Infrastructure-as-a-Service (IaaS). The checklist consists of 36 questions divided into 10 categories:

1. General information (4 questions),
2. Governance (4 questions),
3. Compliance (4 questions),
4. Trust (5 questions),
5. Architecture (6 question),
6. Identity and Access Management (1 question),
7. Software Isolation (2 questions),
8. Data Protection (5 questions),
9. Availability (2 questions),
10. Incident Response (3 questions).

This checklist can be used by records managers and archivists when assessing a CSP offering IaaS as well as by CSPs as guidelines for providing online information about the service.

IaaS Checklist					
Question		**Y***	**N**	**?****	**Answer / additional info*****
1. General information					
1.	Which components are used in IaaS?				
2.	What types of services are offered in IaaS?				
3.	What technologies are being used?				
4.	What implications used technologies have on security and privacy of the system?				
2. Governance					
5.	Is it possible for a client to monitor security of computing environment and data security? How?				
6.	What kind of security assures a client that his data is not mixed with another's?				
7.	What kind of security assures a client that there is no data shared with employees of different rank or/and not created by others?				
8.	What audit mechanisms and tools are used to determine how data is stored, protected and used to validate services, and to verify policy enforcement?				
3. Compliance					
9.	Does the service comply with other countries' laws, regulations, standards and specifications for clients outsde the country of service?				
10.	How is the service secured against unauthorized access, use, disclosure, disruption, modification, or destruction of data?				
11.	What technical and physical safeguards does the service assure?				
12.	Does the service use subcontractors for any part of the used technology or offered service?				

* The questions which are not simple "Yes/No" questions, i.e. require elaborated answer, have the "Y / N / ?" fields shaded.

** The "?" column indicates a situation where no information is available or the question is not applicable to your situation.

*** The "Answer / additional info" column can be used in situations where either a question is not a "Yes/No" type of question or a simple "Yes/No" answer can be supplemented with useful information.

4. Trust					
13.	Is the service secured from denial of service attack?				
14.	Does the service secure ownership rights over data?				
15.	Does the service have any certificate relevant to the service?				
16.	What kind of risk management does the organization provide?				
17.	What kind of physical and logical security is assured for the virtual servers and applications?				
5. Architecture					
18.	How is a hypervisor or virtual machine monitor secured?				
19.	How does the service secure virtual machine images from attack looking for proprietary code and data?				
20.	Does the service use image management process to govern the creation, storage, and use of virtual machine images?				
21.	How does the service secure from attacks on the client side?				
22.	How does the service secure from attacks on the server side?				
23.	Is the service using encrypted network exchange?				
6. Identity and Access Management					
24.	How does the service protect ancillary data: - details about the consumers' accounts, - data about customer-related activity, - data collected to meter and charge for consumption of resources, - logs and audit trails, and other such metadata that are generated and accumulated within the environment, - data of an organization’s initiative (e.g., the activity level or projected growth of a startup company), - metadata collected by the provider?				
7. Software Isolation					
25.	How does the service prevent man-in-the-middle attacks?				
26.	Is the service secured from attacks on the server that target passwords?				

8. Data Protection					
27.	What kind of encryption does the servie use to secure data stored in IaaS?				
28.	Have the service conducted deliberate attacks in order to test the system's protection?				
29.	What procedures are used for data sanitization upon termination of service, i.e. how does the service ensure that the data after deletion are not recoverable?				
30.	Where, geographically, are the data stored?				
31.	Where, geographically, is data backup stored?				
9. Availability					
32.	In a situation of a lawful raid how is the service availability assured to the users not being lawfuly raided?				
33.	Is there a policy regarding user data availability in case of a bankruptcy or other facility loss and how is it defined?				
10. Incident Response					
34.	Is there an incident response plan and how is it defined?				
35.	Does the service keep track of the data using which the scope of the incident, and assets affected can be determined?				
36.	Does the service keep a forensic copy of incident data for legal proceedings or as needed by the consumer? Or, does the servce give incident data to the consumers?				