Journal of Physics: Conference Series

PAPER • OPEN ACCESS

A Review Study on Cloud Computing Issues

To cite this article: Qusay Kanaan Kadhim *et al* 2018 *J. Phys.: Conf. Ser.* **1018** 012006

1st International Conference on Big Data and Cloud Computing (ICoBiC) 2017 IOP Publishing
IOP Conf. Series: Journal of Physics: Conf. Series **1018** (2018) 012006 doi:10.1088/1742-6596/1018/1/012006

A Review Study on Cloud Computing Issues

Qusay Kanaan Kadhim [1, 3], Robiah Yusof [1], Hamid Sadeq Mahdi[3], Sayed Samer Ali Al-shami [2], Siti Rahayu Selamat[1]

[1]Faculty of Information and Communication Technology,
[2] Institute of Technology Management and Entrepreneurship
[1, 2] Universiti Teknikal Malaysia Melaka (UTeM)
Hang Tuah Jaya, 76100 Durian Tunggal, Melaka, Malaysia.
[3] Department of Computer Science, Al Yarmouk University College, Iraq.
[3] University of Diyala / Basic Education College / Computer Science Department, Iraq.
corresponding author's Email: qusaykanaan@utem.edu.my

Abstract. Cloud computing is the most promising current implementation of utility computing in the business world, because it provides some key features over classic utility computing, such as elasticity to allow clients dynamically scale-up and scale-down the resources in execution time. Nevertheless, cloud computing is still in its premature stage and experiences lack of standardization. The security issues are the main challenges to cloud computing adoption. Thus, critical industries such as government organizations (ministries) are reluctant to trust cloud computing due to the fear of losing their sensitive data, as it resides on the cloud with no knowledge of data location and lack of transparency of Cloud Service Providers (CSPs) mechanisms used to secure their data and applications which have created a barrier against adopting this agile computing paradigm. This study aims to review and classify the issues that surround the implementation of cloud computing which a hot area that needs to be addressed by future research.

Keywords: Cloud, Computing, Security, Issues

1. Introduction

Growth in Cloud Computing ↓

Cloud computing becomes a promising networking for infrastructure pattern which can deploy large-scale application in a cost-effective method. It is defined as "applications delivered as services over the Internet and the hardware and systems software in the data centers that provide those services [1]. Recently, cloud computing has been widely adopted by the industry and organizations due to its usability and simple cloud of services oriented models. The number of cloud users can access the cloud of services to keep on the increasing daily and safe systems in the cloud computing environments. Cloud computing technology plays an important role in academic and industry organizations. The business processes are composed and implemented in the distributed loosely coupled environments and the composite of services which includes more services and thus the cloud of services will be connected by varies patterns and approaches. Cloud computing is providing organizations to use shared data storage and cloud resources. It is better than to develop with the own platforms. Further, cloud computing provides companies to have a data flexible, secure system, and cost-effective cloud infrastructure [2]. Additionally, the cloud computing can provide on demand dynamically scalable virtualized cloud resources via the web of internet. indeed, the cloud computing has not only changed the way of providing cloud services but influenced the way of application

Basically explains how Cloud Computing has come to be as popular as it is today

→ Talked about in my Research Speech

development, which helps companies to save IT resources during the lifecycle and shorten application development time [3].

Despite, the advantages of the cloud computing, it still surrounded by several issues that are associated with security management [4] includes lack of trust in data security and privacy by users, organizational inertia, loss of governance, and uncertain provider's compliance [5]. The security issue became extra complex under the cloud model as new scopes have arrived into the problem scope associated to the model data security [6], users' privacy [7], network security, platform and infrastructure issues. Recently, studies from various disciplines emphasized to the importance of cloud computing security management in all areas of application to mitigate those issues. The new version of cloud security management consists of the processes and methods that are useful to reduce cloud security issues. It also includes characteristics like on-demand service provision, virtualization and virtual data centers, and high flexibility access to data on cloud storage and release of service provision like storage, network, cloud applications, servers and its cloud services. Indeed, they present the conceptual model for cloud security that involved components such as data privacy, legally and standard, policy, compliance and regulatory issues of government organizations. Due to the fact that there are many types of security issues, this study reviewed many types of security issues. This study also classified the sub-issues of cloud computing under structured groups which helps future research to explore the related solution. To classify the security issues in the government, this study interviewed the managers of information technology department in 23 of Iraqi government departments and they classified the security issues under five main groups which are: - mobility and cloud government application security issues, cloud security services and application issues, cloud security data, cloud network security issues and issues for cloud security platform and infrastructure.

The study that is going to be discussed is only from sources based in Iraq

One of the possible errors for the study that is going to be discussed is the fact that it takes place in only one country

The intro basically talks about how cloud computing has come a long way since the start, but still has several issues that exist currently in cloud security.

2. Literature Review

2.1. Cloud Computing Issues

There are many cloud security issues appear in different type of technologies which include networks, databases, operating systems, virtualization, resource scheduling, transaction management, concurrency control and memory management are used in cloud computing [8]. The cloud computing is designed as computing utility. The majority of individuals and enterprises use to migrate their work into the cloud, workloads which became more heterogeneous because the cloud computing resources are much more heterogeneous as cloud providers constantly scale or update the clusters with new generations of machines. For example, the government organizations run applications and data transfer in their own the private cloud and then transform it to the public cloud. Nevertheless, there are many security issues exist in the cloud computing technology that threaten the data credibility and confidently. This emphasis to the important to design a cloud computing security with relevant standards and policies to protect the users. Despite, there are major efforts made to design effective cloud computing, the cloud service security is still facing new business and management problems arising from virtualization, multi-tenancy in the cloud, and data encryption technology, trusted cloud and cloud data sensitive confidentiality issues [9]. The government and non-government organizations are struggling to identify the features of cloud security issues and then build and prepare plan that can help them to make appropriate decision toward a successfully adoption of cloud computing technology projects for organizations. There are several key cloud security challenges within the cloud computing environment as shown in Figure , which include mobility and application security issues [10], cloud security services and applications issues, cloud security data issues [11], cloud network security issues [12] and cloud security platform and infrastructure issues [13]

There are many issues that appear in cloud security

The cloud must always be updated and constantly renewed as new issues are found.

Many organizations have a hard time finding and solving issues regarding cloud computing

2.1.1. Mobility and Cloud Government Application security issues

Despite the growth usage of mobile computing, exploiting its full potential advantages is difficult due to its inherent problems such as secure resource, frequent disconnections, and mobility [14].Data in the cloud is typically in a shared environment alongside with data from other customers. Encryption is critical to protect data sensitive confidentiality and privacy of the data while in transiting and in cloud storage. Legally, third-party cloud service providers (CSPs) and their customer organizations are

* Learned about this Cybersecurity class as well

Example: Verizon using AWS cloud, FISD using Google cloud, Different companies using Azure, etc.

distinct enterprises. If the CSP fails in its responsibilities, it could have legal liability implications for the CSP's customer organizations. In contrast, if a cloud customer organization fails in its responsibilities, it is less likely to be exposure to legal implications of the CSP [15]. There are some responsibilities for the organizations such as flexibility access issue of cloud providers, protect sensitive data in the cloud computing, understand legally and standards issues, software development life cycle management, portability and interoperability, and cloud platform reliability and latency. The policy is a foundational issue that is related to legal definition and organizational charter to facilitate and guide the establishment of the vision, missions, responsibilities, and authorities of major actors in cloud computing organization. Open Security Architecture (OSA) provides frameworks that are easily integrated into software applications for the security architecture community. Its patterns are based on schematics that show the data traffic flow control for secure cloud computing and particular implementation with policies implemented at each step for the cloud security issues. The cloud mobile applications can connect and request services hosted on a remote cloud computing by interfaces [18]. However, mobile Web services need to consider additional constraints other than standard Web services: frequent loss of connectivity, low computational resources, and low bandwidth [37]. In this section, the mobility and cloud government application security issues are discussed.

Table 1 Issues for Mobility and Cloud Government Application Security Based on Studies

Mobility and Cloud Government Application security issues	Related works
Lack of Standards, Legally, and Policy	[16]–[19].
Loss of Security Governance	[20]–[22].
Malicious Insider Threats in the Cloud Computing	[23], [24][25]
Cloud Computing Regulatory Requirements and Cloud Compliance	[22], [26]–[29].
Challenges	[30][31]
Cloud Computing Portability and Interoperability	
Biometric Security System for Cloud Computing Environment	[32][33][34][35]

2.1.2. Cloud Security Services and Application Issues

Service and application level issues relate to the security factors concerned with performance measurements of the cloud computing system, and the quality of service and cloud service level agreement [41]. For example, in what ways do mobile cloud computing systems ensure data of availability; what are the fault- tolerance (FT) mechanisms employed to ensure smooth execution and uninterrupted service [42]. Therefore, cloud of services able to extend dynamically to meet user on demand and requirements. Additionally, capabilities of the cloud can be rapidly and elastically increased to meet immediate demand and scaled down to release unused resources. However, monitoring should be done by the cloud service provider frequently to ensure that the cloud is secure and not infected by the malicious code that have been uploaded to the cloud by the hackers or attackers with the purpose of stealing sensitive data or even damaging the information of users. Further, Table 2 illustrates the issues for cloud security services and application based on studies bellow.

Table 2. Issues for Cloud Security Services and Application Based on Studies

Cloud Security Services and Application Issues	Related works
Cloud Service Level Agreement (SLAs) and Quality of Service	[36]–[45].
Trusted for Cloud Services	[22], [43], [46]–[54].
Access Control in Cloud Computing Environment	[2], [16], [24], [43], [48], [52], [54].
Security of Cloud Interfaces and API	[23], [25], [27], [28], [55].
Availability of Cloud Data	[38], [48], [58]–[62].

Everything depends on reliability and security of the CSP

* Mobile Web Services differ from regular ones *

If CSP fails, companies dependent on those providers will fail as well. On the other hand if the company fails, it is less likely to be because of the Cloud Service Provider

Have to take into consideration during Original Work planning!

Two important factors are the quality of importance and the level agreement

Cloud services and capabilities change according to the situation of the service providers

Overall, the cloud service is primarily dependent upon the CSP's level of cloud service and their security.

Confidentiality and Integrity play a huge role in cloud computing just like they do in all of the other fields.

2.1.3. Cloud Security Data

Generally, the data in cloud computing belongs to different owners in the cloud computing resources which must be trusted. Therefore, unauthorized users should be forbidden from that data or information [52], [54], [59]. In trusted and cloud data sensitive confidentiality refers to original data that must keep in a password protected by data management systems with security guard services in the cloud computing environments. All cloud data are entered, stored and backed-up in a password-protected by the management of data in the cloud computing [63]. Besides, the cloud data storage is a model of data storage in which the integrity data is stored in logical pools. It allows cloud users to store their data in a remote server to get rid of expensive local storage and managing brand cost and then flexibility access data of interest anytime and anywhere [64].

Confidentiality refers to keeping information protected by passwords, encryption, etc. and integrity means staying true to not access others information on the cloud

Table 3: Issues for Cloud Security Data Based on Studies

Cloud Security Data Security	Related works
Cloud Data Privacy Security	[16], [17], [22], [25], [27], [31], [58], [65].
Data Protection in Cloud Computing Environments	[2], [54], [55], [47].
Cloud Data Confidentiality Issues	[48], [49], [52], [54].
Cloud Data Limitations and Segregation	[31], [62], [60]–[61].
Cloud Data Integrity	[16], [24], [27], [52], [54], [59].
Cloud Data Eavesdropping Attack and Leakage	[23], [33], [39], [55], [57].

* Learned in Cybersecurity class *

2.1.4. Cloud Network Security Issues

Cloud network security is one of the network security issues. Cloud Computing permits ever-present, convenient, on-demand network access to a shared pool of configurable networks that can be quickly provisioned and free with negligible management effort or service provider communication [67]. Due to the fact that the Cloud Computing embodies a comparatively new computing model, there is an important deal of vagueness about how security at network can be attained and how applications security is progressed to Cloud Computing [62]. The cloud network issues are the higher response time of nodes while performing data communication through co-operative caching [8].

Basically means that cloud computing allows for easy and accessible method for people to store and access data

Since cloud computing is relatively new, there are many ways to strengthen the security that are yet to be discovered

Table 4: Issues for Cloud Network Security Based on Studies

Cloud Network Security issues	Related works
Detection and Recovery	[2], [15], [31], [52], [55], [59].
Flow Control for Secure Cloud Computing	[2], [27], [43], [58].
Cloud Account or Cloud Service Hijacking	[19], [40], [63], [64] .
Cloud Network Traffic Analysis and control	[21], [25], [45], [65].
Bandwidth Cost in the Cloud	[26], [50], [51], [58]].
Distributed Denial of Service (DDoS) Attacks for the Cloud	[2], [15], [47], [66].

2.1.5. Cloud Security Platform and Infrastructure Issues

Security of the cloud infrastructure relies on trusted cloud and cryptography. In addition, no standard service contract exists that covers the ranges of cloud services available and the needs of different organizations. Beside that the cloud computing technology allow to design and implemented a real time alert system on top of the cloud infrastructure [44]. However, cloud computing offers an elastic infrastructure that Cloud Management Agents can use to obtain streaming resources that match the demand. Also, multi cloud providers support different platforms and offer constantly changing packages of capabilities. Further, infrastructure security is the basis of cloud computing security, mainly in the cloud for the upper layer of security services to provide security, infrastructure security by hardware and software security, can be a variety of intrusion defense, redundant backup of data, intrusion detection and prevention in network security. Cloud infrastructure security issues the risk

Cloud Management Agents can use the elastic infrastructure to obtain new resources that match the demand in the field

Multi Cloud providers offer constantly changing packages, which are more helpful

↳ Ending of the paragraph is kind of confusing

associated end-user's concern and is also the focus this work's research direction. Therefore, the centralized security solution for insecurity cloud is proposed and the scenarios of this system and methods after that constructed.

Table 5: Issues for Cloud Security Platform and Infrastructure Based on Studies

Cloud Security Platform and Infrastructure Issues	References
Cloud Platform Reliability and Latency	[31], [65], [71]–[68].
The multi-tenancy in the Cloud	[15], [24], [27], .
Scalability and Capability in the Cloud	

Since cloud computing involves so many different technologies, the cloud continues to become different

3. Cloud Security Issue Factors

Many cloud security issues can obtain from different technologies that including networks, databases, operating systems, virtualization, resource scheduling, transaction management, concurrency control and memory management are used in cloud computing [8]. Additionally, cloud computing is designed as computing as a utility. Customers rent computing resources in the cloud to complete their work. Then to ensure the quality of service (QoS) requirements defined by customers and guarantee the resource utilization in the cloud datacenters, effective resource management systems should be considered. Therefore, more individuals and enterprises migrating their work into the cloud, workloads in the cloud become more heterogeneous. Meanwhile, cloud computing resources are much more heterogeneous as cloud providers constantly scale or update the clusters with new generations of machines [69]. However, in reality government organizations run applications and data transfer in their own the private cloud and then transmute it to the public cloud. While there are many security issues exist in the cloud computing technology , cloud security should design relevant standards and policies as soon as possible [43].

Cloud Computing becomes heterogenous as more businesses and users store more data in the cloud.

As technology develops, cloud computing changes... use for original work!

Cloud computing is a new emerging technology, which every organization these days wants to adapt for its business for more profitability, interoperability, capability, and scalability. This network communication defined cloud computing, highlighted all the cloud service models likes public, private, hybrid and community cloud computing. In addition to information security risks under traditional IT architecture, cloud service security is still facing new business and management risk arising from virtualization, multi-tenancy in the cloud, and data encryption technology, trusted cloud and cloud data sensitive confidentiality issues [9]. However, cloud computing is changing the current IT delivery model for services. Benefits for business and IT include reduced costs, scalability, flexibility access to data on cloud storage, capacity utilization, higher efficiencies, performance, and mobility [40]. The enterprise will be able to identify the features of cloud security issues and then build and prepare plan that can help them to make appropriate decision toward a successfully adoption of cloud computing technology projects for organizations [60]. Generally, cloud computing's issues can bring negative effects on any companies or organizations, therefore an effective risk management is needed to balance the operational and financial cost as well as proactive actions to secure data, network, platform information systems and technologies [70]. According to [53], understanding the true potential of mobile cloud computing and identifying issues with mobile cloud security, privacy, feasibility and accessibility remain a major concern for both the customers and the enterprises [57]. Therefore, the presence of cloud security issues and challenges have critical influence on the success of cloud computing systems. Thus, it is critical to identify and classify control various issues during cloud computing environments by using controlling mitigation techniques those security issues, the success rate of cloud computing systems could be increased. Cloud security, cloud data privacy, feasibility and accessibility remain a major concern for both the users and the enterprises [56]. There are several key cloud security challenges within the cloud environment such as [28]: Key stores that must be protected in data storage, detection and recovery and in backup. Improper key storage may lead to encryption data. Flexibility accesses to key data storage have to be limited to the authorized personnel who require the individual keys. These keys ought to be under policies governing them [45].

Along with IT security risks, there are many management risks with cloud computing that are yet to be solved

Although issues in cloud security are developing more and more, essentially they are needed to provide success

Without problems in the cloud infrastructure, cloud architects don't know what to improve, so problems are basically necessary for growth.

4. Conclusion

Cloud computing is a new emerging technology, which every organization these days adapt it to facilitate the flexibility of their businesses in terms data storage, exchange, transform which enable them to upgrade their profitability, interoperability, capability, and scalability. This network communication defined cloud computing, highlighted all the cloud service models likes' public, private, hybrid and community cloud computing. The cloud computing is changing the current IT delivery model for services. Benefits for business and IT include reduced costs, scalability, and flexibility access to data on cloud storage, capacity utilization, higher efficiencies, performance, and mobility. Despite, the advantages of the cloud computing, it still surrounded by several issues that are associated with security management, includes lack of trust in data security and privacy by users, organizational inertia, loss of governance, and uncertain provider's compliance. The security issue became extra complex under the cloud model as new scopes have arrived into the problem scope associated to the model data security, users' privacy network security, and platform and infrastructure issues. This study was designed to highlight the cloud computing security issues. The finding of this study emphasises that there are five main issues associated with cloud computing implementation which are Mobility and Cloud Government Application security issues, Cloud Security Services and Application, Cloud Security data, cloud network security issues and cloud security platform and infrastructure issues. These issues form an open room for future research to fill up security issues gap through providing either technical approach or empirical model to mitigate these issues.

Acknowledgements

The authors would like to thank INSFORNET Research Group of Universiti Teknikal Malaysia Melaka for supporting this research

References

[1] M. Armbrust, A. Fox, R. Griffith, A. Joseph, R. Katz, A. Konwinski, G. Lee, D. Patterson, A. Rabkin, and I. Stoica, "A View of Cloud Computing," *Commun. ACM*, vol. 53, no. 4, pp. 50–58, 2010.

[2] K. Jakimoski, "Security Techniques for Protecting Data in Cloud Computing," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 1, pp. 49–56, 2016.

[3] Z. Hong-lie, L. Xin, L. I. U. Yan-ju, and L. Cheng, "Research on Cloud Resource Section Method for the Multi-layer Ontology," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 1, pp. 193–200, 2016.

[4] A. N. Khan, M. L. Mat Kiah, S. U. Khan, and S. A. Madani, "Towards secure mobile cloud computing: A survey," *Futur. Gener. Comput. Syst.*, vol. 29, no. 5, pp. 1278–1299, 2013.

[5] A. M.-H. Kuo, "Opportunities and Challenges of Cloud Computing to Improve Health Care Services," *J. Med. Internet Res.*, vol. 13, no. 3, p. e67, 2011.

[6] L. M. Kaufman, "Data security in the world of cloud computing," *IEEE Secur. Priv.*, vol. 7, no. 4, pp. 61–64, 2009.

[7] M. B. Mollah, M. A. K. Azad, and A. Vasilakos, "Security and privacy challenges in mobile cloud computing: Survey and way ahead," *Journal of Network and Computer Applications*, vol. 84. pp. 38–54, 2017.

[8] D. Sarddar, P. Sen, and M. K. Sanyal, "Central Controller Framework for Mobile Cloud Computing," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 4, pp. 233–240, 2016.

[9] Z. Gao, Y. Li, H. Tang, and Z. Zhu, "Management Process Based Cloud Service," in *International Conference on Cyberspace Technology (CCT 2013)*, 2013, pp. 278–281.

[10] A. Botta, W. de Donato, V. Persico, and A. Pescapé., "Integration of cloud computing and

internet of things: a survey.," *Futur. Gener. Comput. Syst.*, vol. 56, p. 684–700, 2016.

[11] D. A. B. Fernandes, L. F. B. Soares, J. V. Gomes, M. M. Freire, and P. R. M. Inácio, "Security issues in cloud environments: a survey," *Int. J. Inf. Secur.*, vol. 13, no. 2, pp. 113–170, 2014.

[12] V. Inukollu, S. Arsi, and S. Ravuri, "Security Issues Associated With Big Data in Cloud Computing," *Int. J. Netw. Secur. Its Appl.*, vol. 6, no. 3, pp. 45–56, 2014.

[13] H. Rasheed, "Data and infrastructure security auditing in cloud computing environments," *Int. J. Inf. Manage.*, vol. 34, no. 3, pp. 364–368, 2014.

[14] N. Fernando, S. W. Loke, and W. Rahayu, "Mobile cloud computing: A survey," *Futur. Gener. Comput. Syst.*, vol. 29, no. 1, pp. 84–106, 2013.

[15] C. LLP, W. Chan, E. Leung, and H. Pili, "Enterprise Risk Management for Cloud Computing," 2012.

[16] A. Tuli, N. Hasteer, M. Sharma, and A. Bansal, "Exploring Challenges in Mobile Cloud Computing: An Overview," *Confluence 2013: The Next Generation Information Technology Summit (4th International Conference)*. p. 6, 2013.

[17] NSTAC, "NSTAC Report to the President on Cloud Computing," 2012.

[18] F. Al-anzi, S. Yadav, and J. Soni, "Cloud Computing: Security Model Comprising Governance, Risk Management and Compliance," in *2014 International Conference on Data Mining and Intelligent Computing (ICDMIC)*, 2014, pp. 1–6.

[19] R. Matt, "Cybersecurity and Cloud Computing in the Health Care and Energy Sectors: Perception and Reality of Risk Management," 2013.

[20] E. Takamura, C. Gomez-rosa, K. Mangum, and F. Wasiak, "MAVEN Information Security Governance , Risk Management , and Compliance (GRC): Lessons Learned," in *2014 IEEE Aerospace Conference*, 2014, pp. 1–12.

[21] J. Adjei, "Explaining The Role of Trust in Cloud Service Acquisition," in *2014 2nd IEEE International Conference on Mobile Cloud Computing, Services, and Engineering*, 2014, pp. 283–288.

[22] S.-T. Lai and F.-Y. Leu, "A Security Threats Measurement Model for Reducing Cloud Computing Security Risk," in *2015 9th International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing*, 2015, pp. 414–419.

[23] P. Anand, J. Ryoo, H. Kim, and E. Kim, "Threat Assessment in the Cloud Environment – A Quantitative Approach for Security Pattern Selection," in *IMCOM '16*, 2016, p. 8.

[24] E. Cayirci, "Modeling and Simulation as A Cloud Service: A Survey," in *Proceedings of the 2013 Winter Simulation Conference*, 2013, pp. 389–400.

[25] A. Michalas, N. Paladi, and C. Gehrmann, "Security Aspects of e-Health Systems Migration to the Cloud," in *2014 IEEE 16th International Conference on e-Health Networking, Applications and Services (Healthcom) Security*, 2014, pp. 212–218.

[26] P. Hazarika, V. Baliga, and S. Tolety, "The Mobile-Cloud Computing (MCC) Roadblocks," in *2014 Eleventh International Conference on Wireless and Optical Communications Networks (WOCN)*, 2014, pp. 1–5.

[27] M. Bamiah, S. Brohi, and S. Chuprat, "Cloud Implementation Security Challenges," in *Proceedings of 20121ntemational of Cloud Computing. Technologies, Applications & Management*, 2012, pp. 174–178.

[28] F. Al-Musawi, A. H. Al-Badi, and S. Ali, "A Road Map to Risk Management Framework for Successful Implementation of Cloud Computing in Oman," in *2015 International Conference on Intelligent Networking and Collaborative Systems*, 2015, pp. 417–422.

[29] J. K. Ganlea, K. Afriyie, and A. Y. Segbefia, "Microcredit: Empowerment and Disempowerment of Rural Women in Ghana," *World Dev.*, p. Pages 335–345, 2015.

[30] E. Aruna, A. Shri, and A. Lakkshmanan, "Security Concerns and Risk at Different Levels in Cloud Computing," in *2013 International Conference on Green Computing, Communication and Conservation of Energy (ICGCE)*, 2013, pp. 743–746.

[31] B. Shanthini and S. Swamynathan, "Genetic-based biometric security system for wireless sensor-based health care systems," in *Proceedings of the 2012 International Conference on Recent Advances in Computing and Software Systems, RACSS 2012*, 2012, pp. 180–184.

[32] G. Ahammed, R. Banu, and N. Fathima, "An Approach to Secure Communication in IoT (Internet of Things)," in *CONFERENCE ON INTERNET OF THINGS*, 2016, no. February, p. 315.

[33] C. Klein, "Cloudy Confidentiality : Clinical and Legal Implications of Cloud Computing in Health Care," *Anal. Comment.*, vol. 39, no. 4, pp. 571–578, 2011.

[34] A. Wadhawan and A. Bhatia, "Neural Network Based Intelligent Retrieval System for Verifying Dynamic Signatures," *Int. J. Adv. Sci. Technol.*, vol. 83, no. 2015, pp. 27–40, 2015.

[35] M. Alhomidi and M. Reed, "Security Risk Analysis as a Service," in *2013 8th International Conference for Internet Technology and Secured Transactions, ICITST 2013*, 2013, pp. 156–161.

[36] A. Khan, M. Fayaz, A. S. Shah, and F. Wahid, "Critical Analysis of Cloud Computing Software Development Process Models," *Int. J. Softw. Eng. Its Appl.*, vol. 10, no. 11, pp. 451–466, 2016.

[37] E. Arianyan, M. Ahmadi, and D. Maleki, "A Novel Taxonomy and Comparison Method for Ranking Cloud Computing Software Products," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 3, pp. 173–190, 2016.

[38] M. Carroll, A. Merwe, and P. Kotzé, "Secure Cloud Computing: Benefits, Risks and Controls," in *Information Security for South Africa -2011*, 2011, pp. 1–9.

[39] J. S. Sengar and R. Sharma, "Review : Ad-Hoc Cloud Architecture & Modern Cryptography," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 6, pp. 45–50, 2016.

[40] H. Rajaei and J. Wappelhorst, "Clouds & Grids: A Network and Simulation Perspective," in *Conference: 2011 Spring Simulation Multi-conference, SpringSim '11, Boston, MA, USA*, 2011, pp. 143–150.

[41] C. O. Rolim, F. L. Koch, C. B. Westphall, J. Werner, A. Fracalossi, and G. S. Salvador, "A Cloud Computing Solution for Patient ' s Data Collection in Health Care Institutions," no. ii, pp. 95–99, 2010.

[42] J. S. Sengar, "SURVEY : Reputation and Trust Management in VANETs," *Int. J. Grid Distrib. Comput.*, vol. 8, no. 4, pp. 301–306, 2015.

[43] S. Bouchenak, G. Gheorghe, G. Chockler, H. Chockler, and A. Shraer, "Verifying Cloud Services: Present and Future," *ACM SIGOPS Oper. Syst. Rev.*, vol. 27, no. 2, pp. 6–19, 2013.

[44] N. Sasikaladevi, "Trust Based Cloud Service Composition Framework," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 1, pp. 99–104, 2016.

[45] V. Saranya, S. Ramya, R. Kumar, and T. Nalini, "Efficient and Parallel Data Processing and Resource Allocation in the Cloud by u sing Nephele ' s Data Processing Framework," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 3, pp. 33–40, 2016.

[46] M. Irfan, M. Usman, Y. Zhuang, and S. Fong, "A Critical Review of Security Threats in Cloud Computing," in *2015 3rd International Symposium on Computational and Business Intelligence (ISCBI)*, 2015, pp. 105–111.

[47] P. Senthil, N. Boopal, and R. Vanathi, "Improving the Security of Cloud Computing using Trusted Computing Technology," *Int. J. Mod. Eng. Res.*, vol. 2, no. 1, pp. 320–325, 2012.

[48] F. S. Al-anzi, A. A. Salman, and N. K. Jacob, "New Proposed Robust, Scalable and Secure Network Cloud Computing Storage Architecture," no. May, pp. 347–353, 2014.

[49] V. Akshaya and T. Purusothaman, "Business Intelligence as a Service in Analysis of Academic Courses," *Int. J. Appl. Eng. Res.*, vol. 11, no. 4, pp. 2458–2467, 2016.

[50] S. Kai, T. Shigemoto, T. Kito, S. Takemoto, and T. Kaji, "Development of Qualification of Security Status Suitable for Cloud Computing System," in *Proceedings of the 4th international workshop on Security measurements and metrics - MetriSec '12*, 2012, p. 17.

[51] M. M. A. Ghosh, R. R. Atallah, and S. S. A. Naser, "Secure Mobile Cloud Computing for Sensitive Data: Teacher Services for Palestinian Higher Education Institutions," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 2, pp. 17–22, 2016.

[52] K.-F. Ho, H. Hirai, Y.-H. Kuo, H. Meng, and K. Tsoi, "Indoor Air Monitoring Platform and Personal Health Reporting System: Big Data Analytics for Public Health Research," in *2015 IEEE International Congress on Big Data*, 2015, no. 2, pp. 309–312.

[53] A. Priya, T. Meena, and M. Devi, "Efficient Approach for Data Retrievability on Cloud Storage Systems," *IJSRSET*, vol. 2, no. 2, pp. 408–412, 2016.

[54] H. Suo, Z. Liu, J. Wan, and K. Zhou, "Security and Privacy in Mobile Cloud Computing," in *2013 9th International Wireless Communications and Mobile Computing Conference (IWCMC)*, 2013, pp. 655–659.

[55] R. Kumar and S. Rajalakshmi, "Mobile Cloud Computing: Standard Approach to Protecting and Securing of Mobile Cloud Ecosystems," in *Proceedings - 2013 International Conference on Computer Sciences and Applications, CSA 2013*, 2013, pp. 663–669.

[56] P. Srivastava, "Multiple Key Based Architecture to Secure Cloud Database," vol. 4, no. September, pp. 1–7, 2015.

[57] N. Ahmed and A. Abraham, "Modeling Security Risk Factors in a Cloud Computing Environment," *J. Inf. Assur. Secur.*, vol. 8, no. 2013, pp. 279–289, 2013.

[58] S. Mazur, E. Blasch, Y. Chen, and V. Skormin, "Mitigating Cloud Computing security risks using a self-monitoring defensive scheme," *Aerosp. Electron. Conf. (NAECON), Proc. 2011 IEEE Natl.*, pp. 39–45, 2011.

[59] P. Rohmeyer and T. Ben-zvi, "Managing Cloud Computing Risks in Financial Services Institutions," in *2015 Proceedings of PICMET '15: Management of the Technology Age*, 2015, pp. 519–526.

[60] K. Hashizume, D. Rosado, E. Fernández-Medina, and E. Fernandez, "An analysis of security issues for cloud computing," *J. Internet Serv. Appl.*, vol. 4, no. 5, pp. 1–13, 2013.

[61] D. G. Rosado, R. Gómez, D. Mellado, and E. Fernández-Medina, "Security Analysis in the Migration to

[62] Kashyap Rajesh and Sarika Sharma, "Security Challenges and Issues in Cloud Computing – The Way Ahead," *Int. J. Innov. Res. Adv. Eng.*, vol. 2, no. 9, pp. 32–35, 2015.

[63] B. S. Al-Attab and H. S. Fadewar, "Security Issues and Challenges in Cloud Computing," *Int. J. Emerg. Sci. Eng.*, vol. 2, no. 7, pp. 22–26, 2014.

[64] H. T. Dinh, C. Lee, D. Niyato, and P. Wang, "A Survey of Mobile Cloud Computing : Architecture , Applications , and Approaches," *Wirel. Commun. Mob. Comput.* -, no. Cc, pp. 1–38, 2013.

[65] M. Vermaat, S. Sebok, S. Freund, J. Campbell, and M. Frydenberg, *Discovering Computers 2016: Tools, Apps, Devices, and the Impact of Technology*. 2016.

[66] M. I. M. Hanifah, R. C. Omar, N. H. N. Khalid, A. Ismail, I. S. Mustapha, I. N. Z. Baharuddin, R. Roslan, and W. M. Z. Zalam, "Integrated Geo Hazard Management System in Cloud Computing Technology," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 160, p. 12081, 2016.

[67] A. A. Soofi and M. I. Khan, "Encryption Techniques for Cloud Data Confidentiality," *Int. J. Grid Distrib. Comput.*, vol. 7, no. 4, pp. 11–20, 2014.

[68] D. Tse, "Challenges on Privacy and Reliability in Cloud Computing Security," in *International Conference on Information Science, Electronics and Electrical Engineering (ISEEE), 2014*, 2014, pp. 1181–1187.

[69] L. Xu and J. Li, "Building Efficient Resource Management Systems in the Cloud: Opportunities and Challenges," *Int. J. Grid Distrib. Comput.*, vol. 9, no. 3, pp. 157–172, 2016.

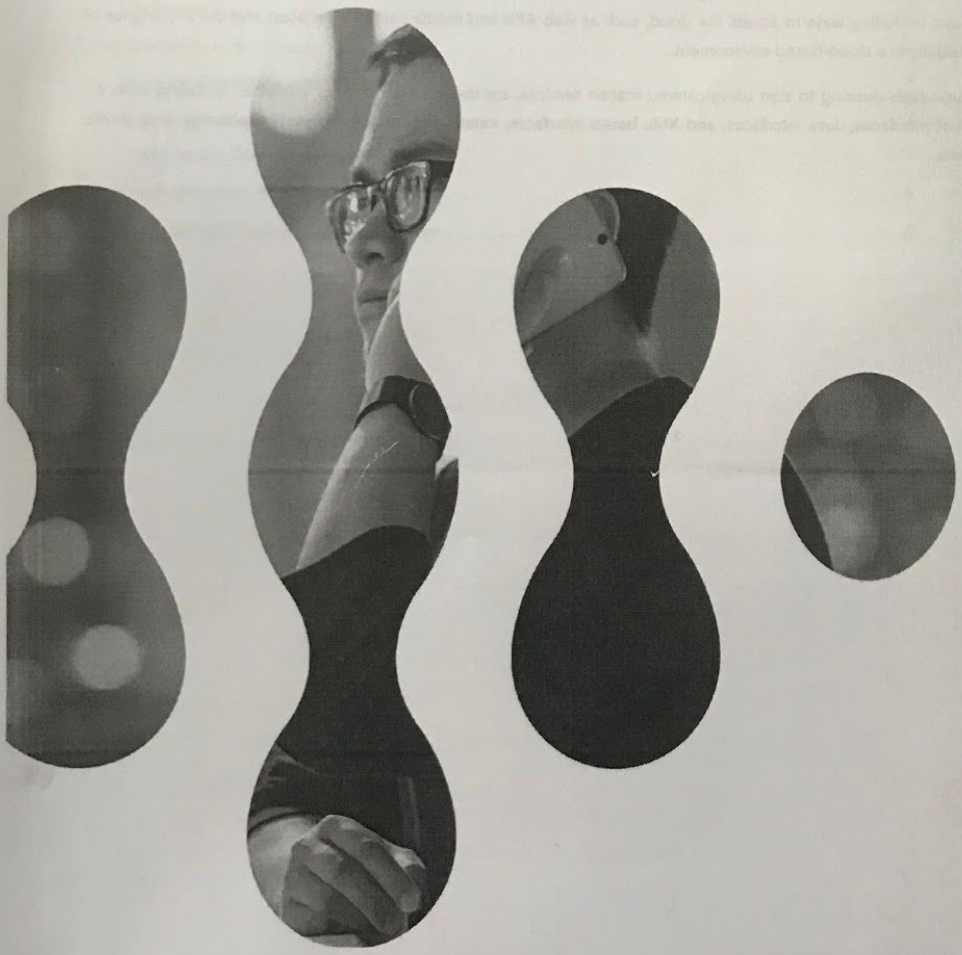
[70] B. Al-shargabi and O. Sabri, "A study of Adopting Cloud Computing from Enterprise Perspective using Delone and Mclean IS Success Model," *Int. J. Comput. Sci. Inf. Secur.*, vol. 14 S1, no. February, p. 5500, 2016.

[71] A. Khrisna and Harlili, "Risk Management Framework with COBIT 5 and Risk Management Framework for Cloud Computing Integration," in *2014 International Conference of Advanced Informatics: Concept, Theory and Application (ICAICTA) Risk*, 2014, pp. 103–108.

Dialogic.

White Paper

Introduction to Cloud Computing



*First few pages of article annotated on laptop

Introduction to Cloud Computing

White Paper

Executive Summary

A common understanding of "cloud computing" is continuously evolving, and the terminology and concepts used to define it often need clarifying. Press coverage can be vague or may not fully capture the extent of what cloud computing entails or represents, sometimes reporting how companies are making their solutions available in the "cloud" or how "cloud computing" is the way forward, but not examining the characteristics, models, and services involved in understanding what cloud computing is and what it can become.

This white paper introduces internet-based cloud computing, exploring the characteristics, service models, and deployment models in use today, as well as the benefits and challenges associated with cloud computing. Also discussed are the communications services in the cloud (including ways to access the cloud, such as web APIs and media control interfaces) and the importance of scalability and flexibility in a cloud-based environment.

Also noted for businesses desiring to start using communication services, are the interface choices available, including Web 2.0 APIs, media control interfaces, Java interfaces, and XML based interfaces, catering to a wide range of application and service creation developers.

Introduction to Cloud Computing

White Paper

Table of Contents

Introduction to Cloud Computing

White Paper

Introduction

This white paper is an introduction to the terms, characteristics, and services associated with internet-based computing, commonly referred to as cloud computing. Characteristics, such as infrastructure, provisioning, network access, and managed metering are presented.

The primary business service models being deployed (such as software, platform, and infrastructure as a service) and common deployment models employed by service providers and users to use and maintain the cloud services (such as the private, public, community, and hybrid clouds) are discussed.

Also introduced are the benefits and challenges associated with cloud computing, and for those seeking to use communications services in the cloud, briefly presented are different ways of determining the interfaces needed to use these communications services.

Cloud Computing

The term "cloud", as used in this white paper, appears to have its origins in network diagrams that represented the internet, or various parts of it, as schematic clouds. "Cloud computing" was coined for what happens when applications and services are moved into the internet "cloud." Cloud computing is not something that suddenly appeared overnight; in some form, it may trace back to a time when computer systems remotely time-shared computing resources and applications. More currently though, cloud computing refers to the many different types of services and applications being delivered in the internet cloud, and the fact that, in many cases, the devices used to access these services and applications do not require any special applications.

Many companies are delivering services from the cloud. Some notable examples include the following:

- **Google** — Has a private cloud that it uses for delivering Google Docs and many other services to its users, including email access, document applications, text translations, maps, web analytics, and much more.
- **Microsoft** — Has Microsoft® Office 365® online service that allows for content and business intelligence tools to be moved into the cloud, and Microsoft currently makes its office applications available in a cloud.
- **Salesforce.com** — Runs its application set for its customers in a cloud, and its Force.com and Vmforce.com products provide developers with platforms to build customized cloud services.

But, what is cloud computing? The following sections note cloud and cloud computing characteristics, services models, deployment models, benefits, and challenges.

Characteristics

Cloud computing has a variety of characteristics, with the main ones being:

- **Shared Infrastructure** — Uses a virtualized software model, enabling the sharing of physical services, storage, and networking capabilities. The cloud infrastructure, regardless of deployment model, seeks to make the most of the available infrastructure across a number of users.
- **Dynamic Provisioning** — Allows for the provision of services based on current demand requirements. This is done automatically using software automation, enabling the expansion and contraction of service capability, as needed. This dynamic scaling needs to be done while maintaining high levels of reliability and security.
- **Network Access** — Needs to be accessed across the internet from a broad range of devices such as PCs, laptops, and mobile devices, using standards-based APIs (for example, ones based on HTTP). Deployments of services in the cloud include everything from using business applications to the latest application on the newest smartphones.
- **Managed Metering** — Uses metering for managing and optimizing the service and to provide reporting and billing information. In this way, consumers are billed for services according to how much they have actually used during the billing period.

In short, cloud computing allows for the sharing and scalable deployment of services, as needed, from almost any location, and for which the customer can be billed based on actual usage.

Service Models

Once a cloud is established, how its cloud computing services are deployed in terms of business models can differ depending on requirements. The primary service models being deployed (see Figure 1) are commonly known as:

- **Software as a Service (SaaS)** — Consumers purchase the ability to access and use an application or service that is hosted in the cloud. A benchmark example of this is Salesforce.com, as discussed previously, where necessary information for the interaction between the consumer and the service is hosted as part of the service in the cloud.

Also, Microsoft has made a significant investment in this area, and as part of the cloud computing option for Microsoft® Office 365, its Office suite is available as a subscription through its cloud-based Online Services.

- **Platform as a Service (PaaS)** — Consumers purchase access to the platforms, enabling them to deploy their own software and applications in the cloud. The operating systems and network access are not managed by the consumer, and there might be constraints as to which applications can be deployed. Examples include Amazon Web Services (AWS), Rackspace and Microsoft Azure.
- **Infrastructure as a Service (IaaS)** — Consumers control and manage the systems in terms of the operating systems, applications, storage, and network connectivity, but do not themselves control the cloud infrastructure.

Also known are the various subsets of these models that may be related to a particular industry or market. Communications as a Service (CaaS) is one such subset model used to describe hosted IP telephony services. Along with the move to CaaS is a shift to more IP-centric communications and more SIP trunking deployments. With IP and SIP in place, it can be as easy to have the PBX in the cloud as it is to have it on the premise. In this context, CaaS could be seen as a subset of SaaS.

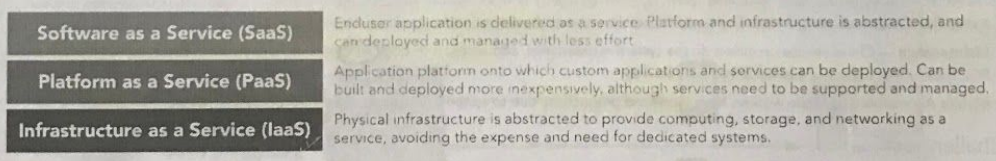


Figure 1. Service Model Types

Deployment Models

Deploying cloud computing can differ depending on requirements, and the following four deployment models have been identified, each with specific characteristics that support the needs of the services and users of the clouds in particular ways (see Figure 2).

- **Private Cloud** — The cloud infrastructure has been deployed, and is maintained and operated for a specific organization. The operation may be in-house or with a third party on the premises.
- **Community Cloud** — The cloud infrastructure is shared among a number of organizations with similar interests and requirements. This may help limit the capital expenditure costs for its establishment as the costs are shared among the organizations. The operation may be in-house or with a third party on the premises.
- **Public Cloud** — The cloud infrastructure is available to the public on a commercial basis by a cloud service provider. This enables a consumer to develop and deploy a service in the cloud with very little financial outlay compared to the capital expenditure requirements normally associated with other deployment options.
- **Hybrid Cloud** — The cloud infrastructure consists of a number of clouds of any type, but the clouds have the ability through their interfaces to allow data and/or applications to be moved from one cloud to another. This can be a combination of private and public clouds that support the requirement to retain some data in an organization, and also the need to offer services in the cloud.

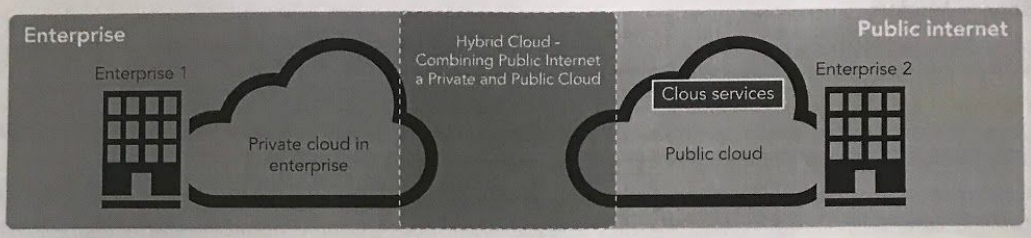


Figure 2. Public, Private, and Hybrid Cloud Deployment Example

Benefits

The following are some of the possible benefits for those who offer cloud computing-based services and applications:

- **Cost Savings** — Companies can reduce their capital expenditures and use operational expenditures for increasing their computing capabilities. This is a lower barrier to entry and also requires fewer in-house IT resources to provide system support.
- **Scalability/Flexibility** — Companies can start with a small deployment and grow to a large deployment fairly rapidly, and then scale back if necessary. Also, the flexibility of cloud computing allows companies to use extra resources at peak times, enabling them to satisfy consumer demands.
- **Reliability** — Services using multiple redundant sites can support business continuity and disaster recovery.
- **Maintenance** — Cloud service providers do the system maintenance, and access is through APIs that do not require application installations onto PCs, thus further reducing maintenance requirements.
- **Mobile Accessible** — Mobile workers have increased productivity due to systems accessible in an infrastructure available from anywhere.

Challenges

The following are some of the notable challenges associated with cloud computing, and although some of these may cause a slowdown when delivering more services in the cloud, most also can provide opportunities, if resolved with due care and attention in the planning stages.

- **Security and Privacy** — Perhaps two of the more "hot button" issues surrounding cloud computing relate to storing and securing data, and monitoring the use of the cloud by the service providers. These issues are generally attributed to slowing the deployment of cloud services. These challenges can be addressed, for example, by storing the information internal to the organization, but allowing it to be used in the cloud. For this to occur, though, the security mechanisms between organization and the cloud need to be robust and a Hybrid cloud could support such a deployment.
- **Lack of Standards** — Clouds have documented interfaces; however, no standards are associated with these, and thus it is unlikely that most clouds will be interoperable. The Open Grid Forum is developing an Open Cloud Computing Interface to resolve this issue and the Open Cloud Consortium is working on cloud computing standards and practices. The findings of these groups will need to mature, but it is not known whether they will address the needs of the people deploying the services and the specific interfaces these services need. However, keeping up to date on the latest standards as they evolve will allow them to be leveraged, if applicable.
- **Continuously Evolving** — User requirements are continuously evolving, as are the requirements for interfaces, networking, and storage. This means that a "cloud," especially a public one, does not remain static and is also continuously evolving.
- **Compliance Concerns** — The Sarbanes-Oxley Act (SOX) in the US and Data Protection directives in the EU are just two among many compliance issues affecting cloud computing, based on the type of data and application for which the cloud is being used. The EU has a legislative backing for data protection across all member states, but in the US data protection is different and can vary from state to state. As with security and privacy mentioned previously, these typically result in Hybrid cloud deployment with one cloud storing the data internal to the organization.

Communications in the Cloud

For service developers, making services available in the cloud depends on the type of service and the device(s) being used to access it. The process may be as simple as a user clicking on the required web page, or could involve an application using an API accessing the services in the cloud.

Telcos are starting to use clouds to release their own services and those developed by others, but using Telco infrastructure and data. The expectation is that the Telco's communications infrastructure provides a revenue generating opportunity

Using the Communications Services

When in the cloud, communications services can extend their capabilities, or stand alone as service offerings, or provide new interactivity capabilities to current services.

Cloud-based communications services enable businesses to embed communications capabilities into business applications, such as Enterprise Resource Planning (ERP) and Customer Relationship Management (CRM) systems. For "on the move" business people, these can be accessed through a smartphone, supporting increased productivity while away from the office.

These services are over and above the support of service deployments of VoIP systems, collaboration systems, and conferencing systems for both voice and video. They can be accessed from any location and linked into current services to extend their capabilities, as well as stand alone as service offerings.

In terms of social networking, using cloud-based communications provides click-to-call capabilities from social networking sites, access to Instant Messaging systems and video communications, broadening the interlinking of people within the social circle.

Accessing through Web APIs

Accessing communications capabilities in a cloud-based environment is achieved through APIs, primarily Web 2.0 RESTful APIs, allowing application development outside the cloud to take advantage of the communication infrastructure within it (see Figure 3).

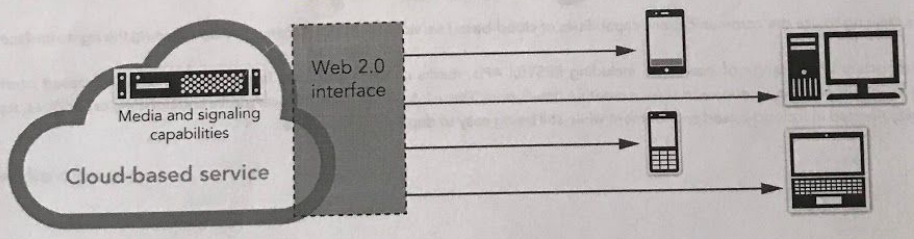


Figure 3. Web 2.0 Interfaces to the Cloud

These APIs open up a range of communications possibilities for cloud-based services, only limited by the media and signaling capabilities within the cloud. Today's media services allow for communications and management of voice and video across a complex range of codecs and transport types. By using the Web APIs, these complexities can be simplified and the media can be delivered to the remote device more easily. APIs also enable communication of other services, providing new opportunities and helping to drive Average Revenue per User (ARPU) and attachment rates, especially for Telcos.

Media Server Control Interfaces

When building communications capabilities into the "core of the cloud," where they will be accessed by another service, the Web 2.0 APIs can be used, as well as a combination of SIP or VoiceXML and the standard media controlling APIs such as MSML, MSCML, and JSR309. The combinations provide different capability sets, but with MediaCTRL being developed in the Internet Engineering Task Force (IETF) it is expected that MediaCTRL will supersede MSML and MSCML and have an upsurge in availability and more developments after it is ratified. JSR309 is a notable choice for those seeking Java development, as it provides the Java interface to media control.

Figure 4 is an example of accessing services in the cloud through Web 2.0 and media control interface APIs.

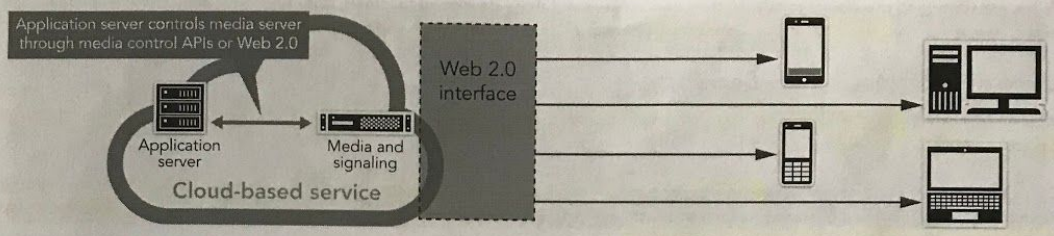


Figure 4. Accessing the Communications Capabilities from Within the Cloud

Whether businesses are deploying communications services for access from outside of or within the cloud, the environment is one that supports the speedy development and rollout of these capabilities.

Communications Scalability

To deliver on the scalability requirements for cloud-based deployments, the communications software should be capable of running in virtual environments. This allows for easily increasing and decreasing session densities based on the needs at the time, while keeping the physical resource requirement on servers to a minimum.

Getting Started with Communications Services

Businesses desiring to use the communications capabilities of cloud-based services will stand to benefit by determining the right interfaces.

Dialogic supports a broad range of interfaces, including RESTful APIs, media control interfaces, Java interfaces, and XML-based interfaces, catering to a wide range of application and service creation developers. These interfaces, available over media and signaling capabilities, support the scalability needed in a cloud-based environment while still being easy to deploy and administer.

www.dialogic.com

For a list of Dialogic locations and offices, please visit: https://www.dialogic.com/contact.

10/17 12023-02

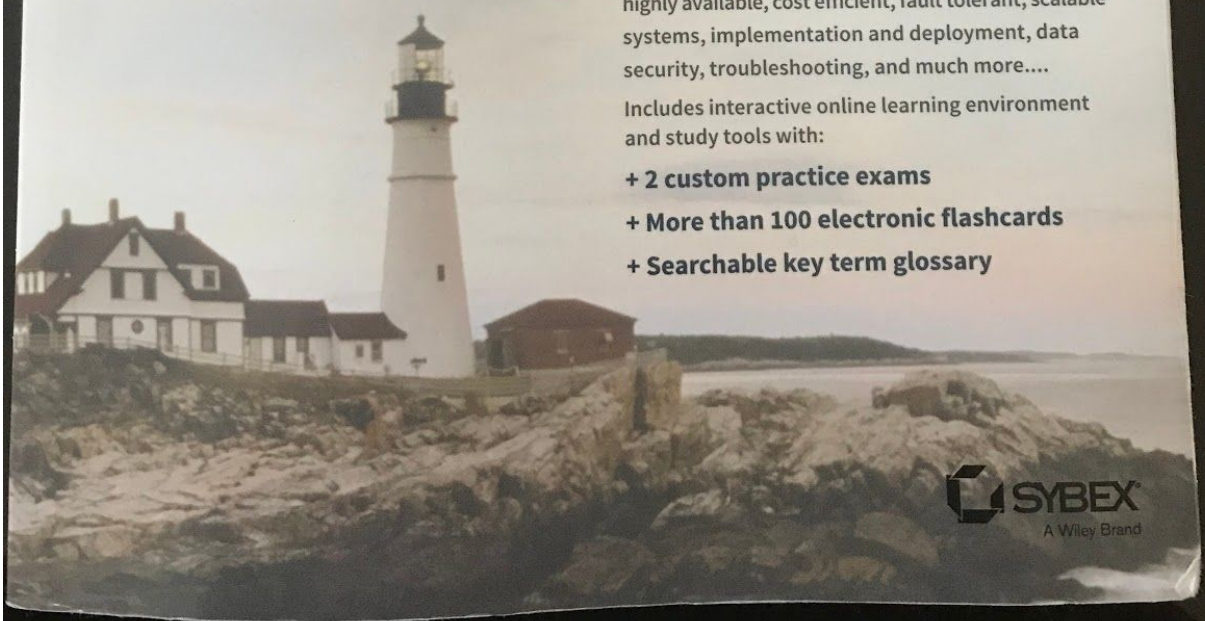
amazon web services

Joe Baron, Hisham Baz, Tim Bixler, Biff Gaut,
Kevin E. Kelly, Sean Senior, John Stamper

AWS® Certified Solutions Architect

OFFICIAL STUDY GUIDE

ASSOCIATE EXAM



Covers exam objectives, including designing highly available, cost efficient, fault tolerant, scalable systems, implementation and deployment, data security, troubleshooting, and much more....

Includes interactive online learning environment and study tools with:

+ 2 custom practice exams

+ More than 100 electronic flashcards

+ Searchable key term glossary

SYBEX
A Wiley Brand

implement continuous monitoring and automation of controls to minimize exposure to security risks. Services like AWS Config Rules, Amazon Inspector, and AWS Trusted Advisor continually monitor for compliance or vulnerabilities giving you a clear overview of which IT resources are or are not in compliance. With AWS Config Rules, you will also know if some component was out of compliance even for a brief period of time, making both point-in-time and period-in-time audits very effective. You can implement extensive logging for your applications using Amazon CloudWatch Logs and for the actual AWS API calls by enabling AWS CloudTrail. AWS CloudTrail is a web service that records API calls to supported AWS Cloud services in your AWS account and creates a log file. AWS CloudTrail logs are stored in an immutable manner to an Amazon S3 bucket of your choice. These logs can then be automatically processed either to notify or even take action on your behalf, protecting your organization from non-compliance. You can use AWS Lambda, Amazon Elastic MapReduce (Amazon EMR), Amazon Elasticsearch Service, or third-party tools from the AWS Marketplace to scan logs to detect things like unused permissions, overuse of privileged accounts, usage of keys, anomalous logins, policy violations, and system abuse.

While AWS provides an excellent service management layer around infrastructure or platform services, organizations are still responsible for protecting the confidentiality, integrity, and availability of their data in the cloud. AWS provides a range of security services and architectural concepts that organizations can use to manage security of their assets and data in the cloud.

Think Parallel

The cloud makes *parallelization* effortless. Whether it is requesting data from the cloud, storing data to the cloud, or processing data in the cloud, as a Solutions Architect you need to internalize the concept of parallelization when designing architectures in the cloud. It is advisable not only to implement parallelization wherever possible, but also to automate it because the cloud allows you to create a repeatable process very easily.

When it comes to accessing (retrieving and storing) data, the cloud is designed to handle massively parallel operations. In order to achieve maximum performance and throughput, you should leverage request parallelization. Multi-threading your requests by using multiple concurrent threads will store or fetch the data faster than requesting it sequentially. Hence, a general best practice for developing cloud applications is to design the processes for leveraging multi-threading.

When it comes to processing or executing requests in the cloud, it becomes even more important to leverage parallelization. A general best practice, in the case of a web application, is to distribute the incoming requests across multiple asynchronous web servers using a load balancer. In the case of a batch processing application, you can leverage a master node with multiple slave worker nodes that processes tasks in parallel (as in distributed processing frameworks like Hadoop).

Reduce Privileged Access

Another common source of security risk is the use of service accounts. In a traditional environment, service accounts would often be assigned long-term credentials stored in a configuration file. On AWS, you can instead use IAM roles to grant permissions to applications running on Amazon EC2 instances through the use of temporary security tokens. Those credentials are automatically distributed and rotated. For mobile applications, the use of Amazon Cognito allows client devices to get controlled access to AWS resources via temporary tokens. For AWS Management Console users, you can similarly provide federated access through temporary tokens instead of creating IAM users in your AWS account. In that way, an employee who leaves your organization and is removed from your organization's identity directory will also lose access to your AWS account.

Best Practice

Follow the standard security practice of granting least privilege—that is, granting only the permissions required to perform a task—to IAM users, groups, roles, and policies.

Security as Code

Traditional security frameworks, regulations, and organizational policies define security requirements related to things such as firewall rules, network access controls, internal/external subnets, and operating system hardening. You can implement these in an AWS environment as well, but you now have the opportunity to capture them all in a script that defines a "Golden Environment." This means that you can create an AWS CloudFormation script that captures and reliably deploys your security policies. Security best practices can now be reused among multiple projects and become part of your continuous integration pipeline. You can perform security testing as part of your release cycle and automatically discover application gaps and drift from your security policies.

Additionally, for greater control and security, AWS CloudFormation templates can be imported as "products" into AWS Service Catalog. This enables centralized management of resources to support consistent governance, security, and compliance requirements while enabling users to deploy quickly only the approved IT services they need. You apply IAM permissions to control who can view and modify your products, and you define constraints to restrict the ways that specific AWS resources can be deployed for a product.

Real-Time Auditing

Testing and auditing your environment is key to moving fast while staying safe. Traditional approaches that involve periodic (and often manual or sample-based) checks are not sufficient, especially in agile environments where change is constant. On AWS, you can

your infrastructure. With the ability to spin up temporary environments, security testing can now become part of your continuous delivery pipeline. Solutions Architects can leverage a plethora of native AWS security and *encryption* features that can help achieve higher levels of data protection and compliance at every layer of cloud architectures.

Best Practice

Inventory your data, prioritize it by value, and apply the appropriate level of encryption for the data in transit and at rest.

Most of the security tools and techniques with which you might already be familiar in a traditional IT infrastructure can be used in the cloud. At the same time, AWS allows you to improve your security in a variety of ways. AWS is a platform that allows you to formalize the design of security controls in the platform itself. It simplifies system use for administrators and those running IT and makes your environment much easier to audit in a continuous manner.

Use AWS Features for Defense in Depth

AWS provides a wealth of features that help Solutions Architects build *defense in depth*. Starting at the network level, you can build an Amazon Virtual Private Cloud (Amazon VPC) topology that isolates parts of the infrastructure through the use of subnets, security groups, and routing controls. Services like AWS Web Application Firewall (AWS WAF) can help protect your web applications from SQL injection and other vulnerabilities in your application code. For access control, you can use AWS Identity and Access Management (IAM) to define a granular set of policies and assign them to users, groups, and AWS resources. Finally, the AWS platform offers a breadth of options for protecting data with encryption, whether the data is in transit or at rest.

Understanding the security features offered by AWS is important for the exam, and it is covered in detail in Chapter 12, "Security on AWS."

Offload Security Responsibility to AWS

AWS operates under a shared responsibility model, where AWS is responsible for the security of the underlying cloud infrastructure, and you are responsible for securing the workloads you deploy on AWS. This way, you can reduce the scope of your responsibility and focus on your core competencies through the use of AWS managed services. For example, when you use managed services such as Amazon RDS, Amazon ElastiCache, Amazon CloudSearch, and others, security patches become the responsibility of AWS. This not only reduces operational overhead for your team, but it could also reduce your exposure to vulnerabilities.

TABLE 14.1 Storage Scenarios and AWS Storage Options

Sample Scenario	Storage Option
Your web application needs large-scale storage capacity and performance. *-or-* You need cloud storage with high data durability to support backup and active archives for disaster recovery.	Amazon S3
You require cloud storage for data archiving and long-term backup.	Amazon Glacier
You require a content delivery network to deliver entire websites, including dynamic, static, streaming, and interactive content using a global network of edge locations.	Amazon CloudFront
You require a fast and flexible NoSQL database with a flexible data model and reliable performance.	Amazon DynamoDB
You need reliable block storage to run mission-critical applications such as Oracle, SAP, Microsoft Exchange, and Microsoft SharePoint.	Amazon EBS
You need a highly available, scalable, and secure MySQL database without the time-consuming administrative tasks.	Amazon RDS
You need a fast, powerful, fully-managed, petabyte-scale data warehouse to support business analytics of your e-commerce application.	Amazon Redshift
You need a Redis cluster to store session information for your web application.	Amazon ElastiCache
You need a common file system for your application that is shared between more than one Amazon EC2 instance.	Amazon Elastic File System (Amazon EFS)

Let's return to our sample web application architecture and show how different storage options can be leveraged to optimize cost and architecture. We can start by moving any static assets from our web instances to Amazon S3, and then serve those objects via Amazon CloudFront. These static assets would include all of the images, videos, CSS, JavaScript, and any other heavy static content that is currently delivered via the web instances. By serving these files via an Amazon S3 origin with global caching and distribution via Amazon CloudFront, the load will be reduced on the web instances and allow the web tier footprint to be reduced. Figure 14.4 shows the updated architecture for our sample web application.

Chapter 3, "Amazon Elastic Compute Cloud (Amazon EC2) and Amazon Elastic Block Store (Amazon EBS)." Let your instances ask a question at boot: "Who am I and what is my role?" Every instance should have a role to play in the environment (such as database server, application server, or slave server in the case of a web application). Roles may be applied during launch and can instruct the AMI on the steps to take after it has booted. On boot, an instance should grab the necessary resources (for example, code, scripts, or configuration) based on the role and "attach" itself to a cluster to serve its function.

Benefits of bootstrapping your instances include:

- Recreate environments (for example, development, staging, production) with few clicks and minimal effort.
- Maintain more control over your abstract, cloud-based resources.
- Reduce human-induced deployment errors.
- Create a self-healing and self-discoverable environment that is more resilient to hardware failure.

Designing intelligent elastic cloud architectures, where infrastructure runs only when you need it, is an art. As a Solutions Architect, elasticity should be one of the fundamental design requirements when defining your architectures. Here are some questions to keep in mind when designing cloud architectures:

- What components or layers in my application architecture can become elastic?
- What will it take to make that component elastic?
- What will be the impact of implementing elasticity to my overall system architecture?

Leverage Different Storage Options

AWS offers a broad range of storage choices for backup, archiving, and disaster recovery, as well as block, file, and object storage to suit a plethora of use cases. For example, services like Amazon Elastic Block Storage (Amazon EBS), Amazon S3, Amazon RDS, and Amazon CloudFront provide a wide range of choices to meet different storage needs. It is important from a cost, performance, and functional aspect to leverage different storage options available in AWS for different types of datasets.

One Size Does Not Fit All

Your workload and use case should dictate what storage option to leverage in AWS. No one storage option is suitable for all situations. Table 14.1 provides a list of some storage scenarios and which AWS storage option you should consider to meet the identified need. This table is not meant to be an all-encompassing capture of scenarios, but an example guide.

Consider only storing a unique session identifier in a HTTP cookie and storing more detailed user session information server-side. Most programming platforms provide a native session management mechanism that works this way; however, these management mechanisms often store the session information locally by default. This would result in a stateful architecture. A common solution to this problem is to store user session information in a database. Amazon DynamoDB is a great choice due to its scalability, high availability, and durability characteristics. For many platforms, there are open source, drop-in replacement libraries that allow you to store native sessions in Amazon DynamoDB.

Stateful Components

Inevitably, there will be layers of your architecture that you won't turn into stateless components. First, by definition, databases are stateful. In addition, many legacy applications were designed to run on a single server by relying on local compute resources. Other use cases might require client devices to maintain a connection to a specific server for prolonged periods of time. For example, real-time multiplayer gaming must offer multiple players a consistent view of the game world with very low latency. This is much simpler to achieve in a non-distributed implementation where participants are connected to the same server.

Deployment Automation

Whether you are deploying a new environment for testing or increasing capacity of an existing system to cope with extra load, you will not want to set up new resources manually with their configuration and code. It is important that you make this an automated and repeatable process that avoids long lead times and is not prone to human error. Automating the deployment process and streamlining the configuration and build process is key to implementing elasticity. This will ensure that the system can scale without any human intervention.

Automate Your Infrastructure

One of the most important benefits of using a cloud environment is the ability to use the cloud's Application Program Interfaces (APIs) to automate your deployment process. It is recommended that you take the time to create an automated deployment process early on during the migration process and not wait until the end. Creating an automated and repeatable deployment process will help reduce errors and facilitate an efficient and scalable update process.

Bootstrap Your Instances

When you launch an AWS resource like an Amazon EC2 instance, you start with a default configuration. You can then execute automated bootstrapping actions as described in

support growth in users, traffic, or data size with no drop in performance. These architectures should provide scale in a linear manner, where adding extra resources results in at least a proportional increase in ability to serve additional system load. The growth in resources should introduce economies of scale, and cost should follow the same dimension that generates business value out of that system. While cloud computing provides virtually unlimited on-demand capacity, system architectures need to be able to take advantage of those resources seamlessly. There are generally two ways to scale an IT architecture: vertically and horizontally.

Scaling Vertically

Vertical scaling takes place through an increase in the specifications of an individual resource (for example, upgrading a server with a larger hard drive, more memory, or a faster CPU). On Amazon Elastic Compute Cloud (Amazon EC2), this can easily be achieved by stopping an instance and resizing it to an instance type that has more RAM, CPU, I/O, or networking capabilities. Vertical scaling will eventually hit a limit, and it is not always a cost-efficient or highly available approach. Even so, it is very easy to implement and can be sufficient for many use cases, especially in the short term.

Scaling Horizontally

Horizontal scaling takes place through an increase in the number of resources (for example, adding more hard drives to a storage array or adding more servers to support an application). This is a great way to build Internet-scale applications that leverage the elasticity of cloud computing. Not all architectures are designed to distribute their workload to multiple resources, and it is important to understand system characteristics that can affect a system's ability to scale horizontally. One key characteristic is the impact of stateless and stateful architectures.

Stateless Applications

When users or services interact with an application, they will often perform a series of interactions that form a session. A *stateless application* needs no knowledge of the previous interactions and stores no session information. A stateless application can scale horizontally, because any request can be serviced by any of the available system compute resources. Because no session data needs to be shared between system resources, compute resources can be added as needed. When excess capacity is no longer required, any individual resource can be safely terminated. Those resources do not need to be aware of the presence of their peers; all that is required is a way to distribute the workload to them.

Let's assume that the web application we used in the previous section is a stateless application with unpredictable demand. In order for our web instances to meet the peaks and valleys associated with our demand profile, we need to scale elastically. A great way to introduce elasticity and horizontal scaling is by leveraging Auto Scaling for web instances. An Auto Scaling group can automatically add Amazon EC2 instances to an application in response to heavy traffic and remove them when traffic slows. Figure 14.3 shows our web application architecture after the introduction of an Auto Scaling group.

- If the single web server fails, the system fails.
- If the single database fails, the system fails.
- If the Availability Zone (AZ) fails, the system fails.

Bottom line, there are too many eggs in one basket.

Now let's walk through transforming this simple application into a more resilient architecture. To begin, we are going to address the single points of failure in the current architecture. Single points of failure can be removed by introducing *redundancy*, which is having multiple resources for the same task. Redundancy can be implemented in either standby or active mode.

In standby redundancy when a resource fails, functionality is recovered on a secondary resource using a process called *failover*. The failover will typically require some time before it is completed, and during that period the resource remains unavailable. The secondary resource can either be launched automatically only when needed (to reduce cost), or it can be already running idle (to accelerate failover and minimize disruption). Standby redundancy is often used for stateful components such as relational databases.

In active redundancy, requests are distributed to multiple redundant compute resources, and when one of them fails, the rest can simply absorb a larger share of the workload. Compared to standby redundancy, it can achieve better utilization and affect a smaller population when there is a failure.

To address the redundancy issues, we will add another web instance and add a standby instance for Amazon Relational Database Service (Amazon RDS) to provide high availability and automatic failover. The key is that we are going to add the new resources in another AZ. An AZ consists of one or more discrete data centers. AZs within a region provide inexpensive, low-latency network connectivity to other AZs in the same region. This allows our application to replicate data across data centers in a synchronous manner so that failover can be automated and be transparent for the users.

Additionally, we are going to implement active redundancy by swapping out the Elastic IP Address (EIP) on our web instance with an Elastic Load Balancer (ELB). The ELB allows inbound requests to be distributed between the web instances. Not only will the ELB help with distributing load between multiple instances, it will also stop sending traffic to the affected web node if an instance fails its health checks. Figure 14.2 shows the updated architecture with redundancy for the web application.

This *Multi-AZ* architecture helps to ensure that the application is isolated from failures in a single Availability Zone. In fact, many of the higher level services on AWS are inherently designed according to the Multi-AZ principle. For example, Amazon Simple Storage Service (Amazon S3) and Amazon DynamoDB ensure that data is redundantly stored across multiple facilities.

Introduction

For several years, software architects have created and implemented patterns and best practices to build highly scalable applications. Whether migrating existing applications to the cloud or building new applications on the cloud, these concepts are even more important because of ever-growing datasets, unpredictable traffic patterns, and the demand for faster response times.

Migrating applications to AWS, even without significant changes, provides organizations with the benefits of a secured and cost-efficient infrastructure. To make the most of the elasticity and agility possible with cloud computing, however, Solutions Architects need to evolve their architectures to take full advantage of AWS capabilities.

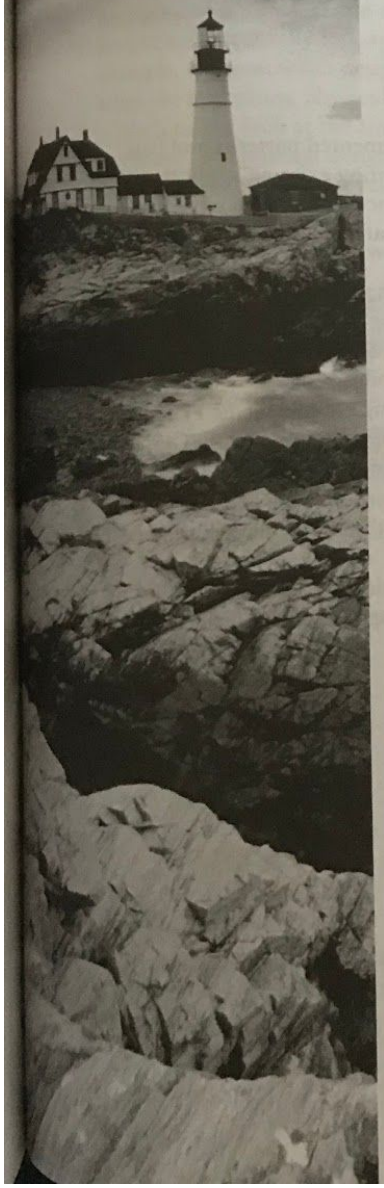
For new applications, AWS customers have been discovering cloud-specific IT architecture patterns that drive even more efficiency and scalability for their solutions. Those new architectures can support anything from real-time analytics of Internet-scale data to applications with unpredictable traffic from thousands of connected *Internet of Things (IoT)* or mobile devices. This leaves endless possibilities for applications architected using AWS best practices.

This chapter highlights the tenets of architecture best practices to consider whether you are migrating existing applications to AWS or designing new applications for the cloud. These tenets include:

- Design for failure and nothing will fail.
- Implement elasticity.
- Leverage different storage options.
- Build security in every layer.
- Think parallel.
- Loose coupling sets you free.
- Don't fear constraints.

Understanding the services covered in this book in the context of these practices is key to succeeding on the exam.

Chapter 14



Architecture Best Practices

THE AWS CERTIFIED SOLUTIONS ARCHITECT ASSOCIATE EXAM OBJECTIVES COVERED IN THIS CHAPTER MAY INCLUDE, BUT ARE NOT LIMITED TO, THE FOLLOWING:

Domain 1.0: Designing highly available, cost-efficient, fault-tolerant, and scalable systems

✓ **1.1 Identify and recognize cloud architecture considerations, such as fundamental components and effective designs.**

Content may include the following:

- How to design cloud services
- Planning and design
- Familiarity with:
  - Best practices for AWS architecture
  - Hybrid IT architectures (e.g., AWS Direct Connect, AWS Storage Gateway, Amazon Virtual Private Cloud [Amazon VPC], AWS Directory Service)
  - Elasticity and scalability (e.g., Auto Scaling, Amazon Simple Queue Service [Amazon SQS], Elastic Load Balancing, Amazon CloudFront)