Available online www.ejaet.com

European Journal of Advances in Engineering and Technology, 2018, 5(12):1054-1058

Research Article ISSN: 2394 - 658X

Agile Research: Applying Industry best Practices to Develop Software for Academic Research

Aryyama Kumar Jana[1], Srija Saha[2]

[1]Electrical Engineering Department, Jadavpur University, Kolkata, India
[2]Electronics & Telecommunication Engineering Department, IIEST, Shibpur, Howrah, India
*janaaryyama@gmail.com, saha.srija01@gmail.com

ABSTRACT

Today cloud computing services have begun to be recognized as promising path for scientific research that might be both scalable as well as productive in the ever-changing academia. This paper explores how to use the state-of-the-art technology to improve the reproducibility of research code and increase the transparency of scientific discoveries. Specifically, it looks at how to use conventional computing resources in conjunction with newly emerging cloud-based offerings. To strengthen the reproducibility of the research code, this paper investigates the integration of cutting-edge cloud platforms such as Amazon Web Services (AWS). With a focus on test automation, continuous integration, and continuous deployment (CI/CD) and version management systems, we seek to build a trustworthy and scalable platform while staying within the bounds of the technologies at hand. Even though the use of cloud-based storage and computing services are still in its infancy in academic research especially in developing countries like India, our study examines the prospective advantages of using such technologies while considering the current technical constraints. Using the resources at our disposal, we emphasize the significance of careful record keeping and data management strategies to support the credibility of research investigation. This research paves the way for future developments in the nexus of cloud-based computing and scientific research by demonstrating the potential pragmatic advantages of incorporating modern technologies into research practices via case-studies and examples pertinent to the modern technological landscape.

Key words: Cloud Computing, Amazon Web Services (AWS), Scalability, Continuous integration, and Continuous deployment (CI/CD)

INTRODUCTION

Within the ever-changing field of academic research, the fundamental nature of exploration and discovery is shaped by the mutually beneficial relationship between scientific methods and technological advancements. At this critical moment, the scientific community needs to come face to face with the groundbreaking possibilities housed inside the rapidly expanding cloud computing solutions space, particularly within the vast ecosystem offered by Amazon Web Services (AWS).

The present academic community is enmeshed in a web of fundamental issues in the sophisticated and intricate embroidery of academic research, which includes the complication of computational scalability, collaborative efforts, and efficient data management. This paper not only comprehends but also proactively influences the possible impact of cloud computing services in ensuring the reproducibility of academic research, given the appealing potential that cloud based solutions offer. As a result, this research emerges as a pioneering endeavor that deliberately applies and modifies fundamental software development techniques inside the complex interactions of existing technologies.

The use of version management systems, automated software testing and continuous integration strategically forms the foundation of this investigation and has the potential to reshape the parameters of the research code reproducibility within the present scientific and temporal limitations. The story goes beyond the simple

integration of AWS and delves into the intricate coordination of careful documentation and tactical information management- a two-pronged approach that aims to improve the reproducibility of research code while also making a significant contribution to the larger debate about transparency and authenticity in the broad field of science.

Reaching into unchartered territory, this paper aims to tread both the way of understanding and active influence, attempting to shape the potential impact of cloud computing solutions in research code reproducibility. By means of a carefully planned and meticulously implemented sequence of real-world instances and case studies that precisely tailored to the current technological limitations and cultural context, this paper aims to not only clarify but also lay a groundwork for a fresh perspective of how the incorporation of cloud-based solutions can drastically alter the course of scholarly research.

This paper seeks to pave the way for a future in which seamless integrations of these cutting-edge tools becomes not only a possibility but also a necessity in pursuing the goals of scalable, efficient, and reproducible research investigation, thereby transforming the very core of academic research practices, by filling the apparent gap conventional computing infrastructure and emerging cloud computing platforms.

LITERATURE REVIEW

Within the constantly changing field of computer science, there has been a radical change in the way that information technology and science interact, with cloud computing services - most notably Amazon Web Services (AWS) - becoming a central focus of study. This literature review presents a thorough summary of the body of academic literature already in existence while illuminating important topics pertaining to the use of AWS and how it improves reproducibility and integrity.

The conversation around the use of cloud computing has undergone a radical change, capturing a paradigmatic progression that goes beyond conventional computing. Early papers that highlight the revolutionary potential of cloud-based solutions including Armbrust et al., 2010 [1] and Mell and Grance, 2011 [2], make key contributions to this conversation. These foundational publications explore the complex world of cloud computing and provide a thorough grasp of its effects on practical applications and computing infrastructure.

A new age of understanding is ushered in by Armbrust et al., 2010 [1], who explains how cloud computing stimulates creativity by offering scalable and adaptable solutions. The authors discuss the possible advantages of this groundbreaking innovation, highlighting important features including resource scalability, affordability, and increased accessibility. This study turns into a pillar, providing the groundwork for rethinking research methodologies via the prism of cloud-based technologies.

Mell and Grance, 2011 [2] expand on the stories by exploring the complexities of the cloud computing potential. Their analysis systematically breaks down the salient characteristics of cloud-based solutions, highlighting the ways in which these technological advancements facilitate a more flexible and dynamic development environment. The focal point of this article is resource scalability, which highlights how cloud technologies provide developers with an unparalleled flexibility to expand computing resources in accordance with the needs of their projects.

The synergy of these avant-garde pieces calls on scholars to embrace a future in which the traditional limits of computer capacities are surpassed, while simultaneously highlighting the revolutionary potential of cloud-based technology. Armbrust et al., 2010 [1] and Mell and Grance, 2011 [2] offer strategic insights that pave a way for more extensive discussion about cloud-based solutions that can reform software development practices in general, indicating a new era in which affordability, scalability and accessibility come together to strategically alter the nature of software development.

The study by Armbrust et al., 2010 [1] highlights the exceptional scalability of cloud resources, which completely transforms the way that scientific computation is traditionally done. Essentially, it enables developers to quickly modify processing power and storage capacity to conform to the constantly changing requirements and complexity of modern scientific endeavors. Scalability is fundamental and goes beyond the finite constraints of conventional computer systems. Developers are freed from the bonds of pre-established computational limitations and no longer limited by predetermined processing capacity or storage restrictions. This improved flexibility allows resources to be better matched to the changing and sometimes erratic nature of requirements.

Mell and Grance, 2011 [2] have examined the financial benefits of cloud computing, highlighting affordability as the primary factor for its wide acceptance. The pay-as-you-go concept which has been included in cloud services has drastically changed how funding is obtained and has increased the availability of powerful computing resources with different budgetary limitations. Proper data handling and security are two essential factors to consider as developers traverse the digital world. Subashini and Kavitha's 2011 [3] work explores the pros and cons of cloud-based data handling systems, highlighting issues with data privacy, security, and authenticity.

METHODOLOGY

This section outlines the methodical approach used to incorporate Amazon Web Services (AWS) into the academic research framework putting particular emphasis on the techniques that can be implemented to improve the reproducibility and accessibility of scholarly research. The process is broken down into 3 main stages - computing infrastructure, programming and teamwork and data handling.

3.1. Computing Infrastructure

The infrastructure configuration, which forms the basis of the research project's technical framework, involves a rigorous procedure of selecting an array of AWS technologies to guarantee maximum efficiency, security, and scalability.

3.1.1 Provisioning AWS accounts with Enhanced Security

a. Multi Factor Authentication (MFA) should be incorporated during the set-up of AWS accounts in accordance with the cybersecurity best practices. MFA strengthens the AWS environments [4], security in general by acting as a strong deterrent against illicit access by demanding extra verification alongside password.

b. The configuration should closely follow the principle of least privilege, guaranteeing that all the system and users in AWS environment have access limited to what is required for their individual tasks. This strategy reduces the possibility of unauthorized access to critical resources and decreases the potential security flaws.

3.1.2 Setting up Resources for Best Performance

a. The research project's particular computational needs should be considered while designing the AWS architecture. To provide a scalable and adaptable computing environment, Amazon Elastic Compute Cloud (EC2) instances should be carefully provisioned [5] considering the CPU and memory specifications. Because of this dynamic provisioning, researchers can adjust their computational resources based on the project's changing requirements.

b. Utilizing Amazon's Simple Storage Service (S3), scalable object-based storage solution mitigates the storage issues. The configuration should make sure that the smooth retention and retrieval of data adheres to Amazon S3's intrinsic durability [6] and availability criteria. The foundation of data handling throughout the research study lifecycle is established by this design.

3.1 3 Using Amazon Key Management Service to Strengthen Cybersecurity

a. Virtual Private Cloud: VPCs' to be put in place to strengthen the security [7] of the research environment to the use of networking technologies, resources may be isolated giving the researchers a private and safe area within the AWS cloud. It also makes it easier to define unique IP address ranges, manage incoming and outgoing traffic and improve segmentation in networks.

b. AWS KMS and encryption protocols: Strict security measures should be put into effect by setting encryption protocols. The AWS Key Management System (KMS) [8] is essential to manage the cryptographic keys, guaranteeing security of data while data is in transit or at rest. The research endeavor maintains the security and confidentiality of sensitive data by using industry standard encryption techniques.

c. Identity and Access Management (IAM): To improve the identification and access restrictions, AWS Identity and Access Management (IAM) [9] should be carefully set up. This involves defining user roles, access restrictions and granular permissions. A precise access control technique that adheres to the least privilege concept is made possible by IAM advancements.

In addition to meeting the research projects immediate compute and storage needs and requirements, this extensive infrastructure design that includes cutting edge AWS capabilities also establishes a valid basis for

availability, scalability, reliability, security and effective resource management over the course of the research project.

3.2. Programming and Teamwork

3.2.1 Using Amazon CodeCommit for Version control

The research code's version control system can be established using the Git repositories hosted on AWS CodeCommit. To monitor changes over time and cooperate easily, researchers can make use of Git's distributed version control abilities, enabled by AWS CodeCommit [10]. The research team can manage their development code in a centra=l secure repository by utilizing AWS CodeCommit's inherent interaction with other services offered by AWS.

3.2.2 Using AWS CodeBuild, CodePipeline and CodeDeploy for CI/CD

The research workflow can be enhanced by integrating AWS CodeBuild, CodePipeline and CodeDeploy [11] to improve the effectiveness and reliability of the code used for the research project. The automated creation and testing of the research codebase can be done using AWS CodeBuild, which makes sure that every piece of code is thoroughly tested before being integrated. CodePipeline manages the movement of code from deploying the source code, coordinating the whole Continuous Integration and Continuous Deployment (CI/CD) pipeline. AWS CodeDeploy makes the software release process seamless and reduces downtime by automating the deployment process further.

3.2.3 Collaboration in Coding Using AWS Cloud9

The cloud-based IDE (Integrated Development Environment) - AWS Cloud9 [12] can be used to help researchers collaborate in real time. The research team can have their own live, engaging, and collaborative virtual space by having concurrent editing and debugging issues. Moreover, interactions and collaborations outside of coding can be made easier. Researchers can easily collaborate with their team using videoconferencing, discuss project milestones, and exchange documentation can improve team efficiency.

3.2.4 Code Evaluation and Review using AWS CodeStar

To maintain the research team's collaborative spirit, AWS CodeStar [12] can be used to expedite the code review process. This service makes it simpler for teammates to look at, commend on and go through code changes by offering an integrated user interface for collaboration in coding. AWS CodeStar enables a systematic and effective code review process, which improves the overall quality of the research code.

3.2.5 Information Exchange using AWS Chime

AWS chime, a communication service by AWS [13], can be used by researchers for collaborative discussions and knowledge sharing in real time by using audio-visual means and instant messaging, resulting in a dynamic and engaging virtual workspace. The interaction and information sharing aspects of research are further improved by smooth integration of AWS Chime with other collaborative tools.

Through the strategic integration of AWS Codebuild, CodeStar, CodeDeploy, CodePipeline, Cloud9, and Chime the research team can create a collaborative environment that extends beyond the conventional coding limits in addition to meeting the technological requirements, this holistic approach of developing code by collaborating with teammates sets the foundation of cutting-edge cooperative research activities.

3.3. Data Handling

3.3.1 Amazon S3 and Amazon Glacier for Data storage and recovery

With its robust and scalable architecture, Amazon S3 [14] is the leader in data management. It can be used as the primary repository. There is provision of versioning in S3 buckets which can greatly improve the repeatability of the analysis while also enabling the researchers to keep track of the modification. In addition, long term storage can be implemented with AWS Glacier [14] which complies with economical archiving procedures and guarantees that the data is preserved for future use.

3.3.2 Amazon S3 versioning for backup and recovery

Amazon S3 versioning [15] can be used for automating backups. The availability of previous data versions is guaranteed by this innovative strategy which is essential for reproducibility and serves as a backup in the event of unanticipated data issues. A robust and recoverable data management approach is possible by using Amazon S3 versioning.

In addition to meeting modern requirements, the data management techniques emphasized by Amazon S3, IAM, Glacier, Data Exchange, and proactive backup plan, builds a strong basis for safe, scalable, and repeatable research data handling.

CONCLUSION

The use of Amazon Web Services (AWS) in academic research can revolutionize the research world by increasing productivity, transparency, and collaboration. The advancements of research procedures can be significantly influenced by AWS technology, which ranges from cutting edge data management techniques to creative code development techniques. AWS CodeBuild, CodeStar, CodeCommit, Cloud9, Chime can be deployed which not only expedites the development process but also promotes a collaborative coding environment. The combination of collaborative coding approaches, version controlling, and continuous integration and continuous deployment establishes a strong basis for reproducible and effective research investigations. Similarly to this, AmazonS3 has become the mainstay of data management for robust and scalable storage matched by AWS Glacier for long term perseverance. Maintaining confidentiality and compliance with regulations, AWS IAM and Data Exchange enables safe data exchange both internally and externally. The future focused strategy for backup and recovery makes use of Amazon S3 versioning, strengthening the reliability of research data. Upon closer examination of the approaches and procedures discussed in this paper, it is evident that AWS not only fulfills the technological requirements of the moment, but also establishes a model for the advancement of research methodologies. The ongoing integration of AWS technology in research study is a testament to the heritage of this era. This exploration of the AWS enabled environment can improve the research workflows and add to the continuing conversation about academic best practices. Essentially the forward-thinking data backup procedures, safe data management practices and collaborative coding environments introduced by AWS should become essential parts of the modern research toolset.

REFERENCES

[1] Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., ... & Zaharia, M. (2010). A view of cloud computing. Communications of the ACM, 53(4), 50-58.

[2] Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

[3] Subashini, S., & Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing. Journal of network and computer applications, 34(1), 1-11.

[4] Rajani, S., Ghorpade, V., & Dhange, M. (2016). Multi-factor authentication as a service for cloud data security. Int J Comput Sci Eng, 4, 43-46.

[5] Hazelhurst, S. (2008, October). Scientific computing using virtual high-performance computing: a case study using the Amazon elastic computing cloud. In Proceedings of the 2008 annual research conference of the South African Institute of Computer Scientists and Information Technologists on IT research in developing countries: riding the wave of technology (pp. 94-103).

[6] Hashimoto, N. (2015). Amazon S3 Cookbook. Packt Publishing Ltd.

[7] Morad, S. (2014). Amazon virtual private cloud connectivity options. Amazon Web Services Documentation https://media. amazonwebservices. com/AWSAmazonVPCConnectivityOptions. pdf.

[8] Amazon, E. C. (2015). Amazon web services. Available in: http://aws. amazon. com/es/ec2/(November 2012), 39.

[9] Anand, A. (2017). Managing Infrastructure in Amazon using EC2, CloudWatch, EBS, IAM and CloudFront. Int. J. Eng. Res, 6(03), 373-378.

[10] Dalbhanjan, P. (2015). Overview of deployment options on aws. Amazon Whitepapers.

[11] Felsen, N. (2017). Effective DevOps with AWS. Packt Publishing Ltd.

[12] Introducing AWS Cloud9 - AWS. 30 Nov 2017, https://aws.amazon.com/about-aws/whats-new/2017/11/introducing-aws-cloud9/.

[13] Announcing Amazon Chime: Frustration-Free Online Meetings with Exceptional Audio and Video Quality - AWS. Amazon Web Services, Inc., 13 Feb. 2017, https://aws.amazon.com/about-aws/whats-new/2017/02/announcing-amazon-chime-frustration-free-online-meetings-with-exceptional-audio-and-video-quality/.

[14] Han, Y. (2015). Cloud storage for digital preservation: optimal uses of Amazon S3 and Glacier. Library Hi Tech, 33(2), 261-271.

[15] Gulabani, S. (2015). Amazon S3 Essentials. Packt Publishing Ltd.