Serverless computing for data processing in open learning and research environments

Ihor A. Bezverbnyi[1], Mariya P. Shyshkina[2]

[1] *V. M. Glushkov Institute of Cybernetics of the National Academy of Sciences of Ukraine, 40 Academician Glushkov Ave., Kyiv, 03187, Ukraine*

[2] *Institute for Digitalisation of Education of the NAES of Ukraine, 9 M. Berlynskoho Str., Kyiv, 04060, Ukraine*

Abstract

Serverless computing is a paradigm that enables the execution of code without provisioning or managing servers. It offers benefits such as scalability, cost-efficiency, and ease of development for cloud-based applications. In this paper, we explore the potential of serverless computing for supporting data processing in open learning and research environments. We propose a concept of a hybrid serverless cloud, which combines different types of cloud services to provide access to various tools and resources for learners and researchers. We present a case study of wave files processing using a lambda function, which demonstrates the feasibility and effectiveness of our approach. We also discuss the challenges and opportunities of integrating serverless components within open systems of learning and research. Finally, we present a vision of a cloud-based open learning and research university environment that leverages serverless technologies to enhance the quality and accessibility of education and research.

Keywords

serverless computing, cloud computing, data processing, open learning, open research

1. Introduction

Cloud-based learning and research environments are emerging as a key paradigm for modernizing the educational process in higher education and fostering open science within the European Research Area [1, 2, 3]. Cloud technologies enable the creation of more convenient, flexible, and scalable systems for accessing electronic resources and services in learning and research activities, as well as facilitating collaboration, mobility, and overcoming geographical and temporal barriers [4, 5, 6, 7, 8, 9, 10]. This provides a basis for implementing the principles and technologies of open science for a wider range of users, such as creating and operating virtual research teams, improving scientific communication processes, accessing and sharing data in the research process, disseminating research results, and engaging with society [11]. Cloud computing tools and services form an information technology platform for the modern educational and scientific environment, becoming a network tool for shaping this environment

3L-Person 2022: VII International Workshop on Professional Retraining and Life-Long Learning using ICT: Person-oriented Approach, October 25, 2022, Kryvyi Rih (Virtual), Ukraine
ihorbezverbnyi@gmail.com (I. A. Bezverbnyi); marimodi@gmail.com (M. P. Shyshkina)
https://www.nas.gov.ua/EN/PersonalSite/Pages/default.aspx?PersonID=0000015986 (I. A. Bezverbnyi); https://iitlt.gov.ua/eng/structure/departments/cloud/detail.php?ID=269 (M. P. Shyshkina)
0000-0001-5569-2700 (I. A. Bezverbnyi); 0000-0001-5569-2700 (M. P. Shyshkina)

CEUR Workshop Proceedings (CEUR-WS.org)

[12]. Therefore, it is important to analyze the trends and challenges of integrating cloud data processing services into the activities of researchers and educational or research institutions.

2. The research results

2.1. The background issues

Cloud computing offers various models, such as IaaS, PaaS and SaaS, that can facilitate learning and research data processing. By abstracting resources and providing simple automation tools, modern cloud platforms simplify many routine tasks, such as installation, maintenance, backup, security, and more [5, 13]. Moreover, in the context of open science, open data and big data processing are essential. To meet the requirements of open science systems design, large amounts of data need to be available and accessible for joint processing by the community of scientists [5]. Therefore, cloud computing platforms can serve as a reasonable framework to support open learning and research processes, both in terms of managing and processing large amounts of data and making them available for collaborative use [5].

The computing capacity is crucial for processing and retrieving large amounts of data, which are needed at most stages of the research process, such as data collection, representation, visualization, analysis, interpretation and discussion. A possible way to save resources and provide flexible use of the cloud-based infrastructure is to use lambda functions within the serverless settings. This leads to the notion of Function-As-A-Service (FAAS) as a promising cloud-based model [14, 12, 15, 16].

The applications and evaluation of serverless computing in different areas are among the current issues considered nowadays, for example for machine learning [17], network functions virtualization [18], geospatial architectures [19]. Casale et al. [20] propose a platform for decomposition and orchestration for serverless computing. Ortiz [21] present architecting serverless microservices on the cloud with AWS and also issues of instructors training to use these technologies. However, the area of educational application of serverless technologies to provide better use and implementation for learning and research within the university sector is poorly investigated and needs further research. There is a need to consider methodological issues and possible ways of serverless technology application within the open learning and research university environment.

The article aims to consider and evaluate a hybrid cloud-based serverless architecture as a possible open learning and research platform to support data processing and research collaboration. The main idea is that design and development of learning and research environment due to the proposed approach will result in more efficient use of the cloud-based resources, better access to learning and research data and collaboration support. The case study of the sound signal processing as a possible example of serverless approach application for learning and research is considered.

2.2. The conceptual basis

The paper introduces the main concepts and terms related to the design and development of university cloud-based learning and research environment (LRE), based on the principles of open science, open education, and cloud-oriented systems, as proposed by Bykov and Shyshkina [4].

The LRE of a higher education institution is defined as an environment that leverages the virtualized computer-technological infrastructure (corporate or hybrid-based) to support the content-technological and information-communication functions of learning and research activities [4].

Serverless technologies are adopted to build applications that require dynamic and unpredictable computing resources. The serverless hybrid cloud architecture enables the deployment of lambda-functions [22], which are cloud-based services that execute computing tasks on demand within the cloud-based infrastructure of a provider, without requiring the user to create and manage the server architecture.

2.3. The model and approach

Figure 1 illustrates the configuration of the serverless application architecture.

The proposed approach is to access lambda-functions through API Gateway, avoiding server management as lambda-functions return the values in static HTML format, which are stored and retrieved on S3-bucket, and can be further processed.

This approach allows the user to access specific electronic resources and computing capacities hosted on a hybrid serverless architecture from any device with an Internet connection.

The advantage of this approach is that it provides flexibility and scalability for learning or research processes that need computing resources for special purposes that may arise occasionally. For example, in the course of an experimental research, big data processing may be needed that require high computing power for a short time. It may be inefficient to maintain and manage a cloud server for these purposes. However, by using lambda-functions, the learner or researcher can access a server with powerful processing capabilities without deploying it every time as the function is needed. The necessary resources can be supplied more efficiently on demand.

2.4. Current developments and implementation

The cloud-based LRE was implemented at the Institute for Digitalisation of Education of the NAES of Ukraine as part of the research projects and pedagogical experiments conducted from 2012 to 2017. During this period, various cloud-based services were integrated into the research and educational process to support open education and open science [4].

In 2018, the V4+ Academic Research Consortium Integrating Databases, Robotics and Language Technologies was established, which aimed to address regional issues related to EU ICT research priorities. The consortium used the following cloud-based components for collaborative work:

- The BOX Cloud shared work-space – a cloud storage and transfer service that connected the researchers' computers and allowed them to share documents.
- The virtual machine with Windows 10 – a remote desktop that provided a common computing environment for the partners [5].

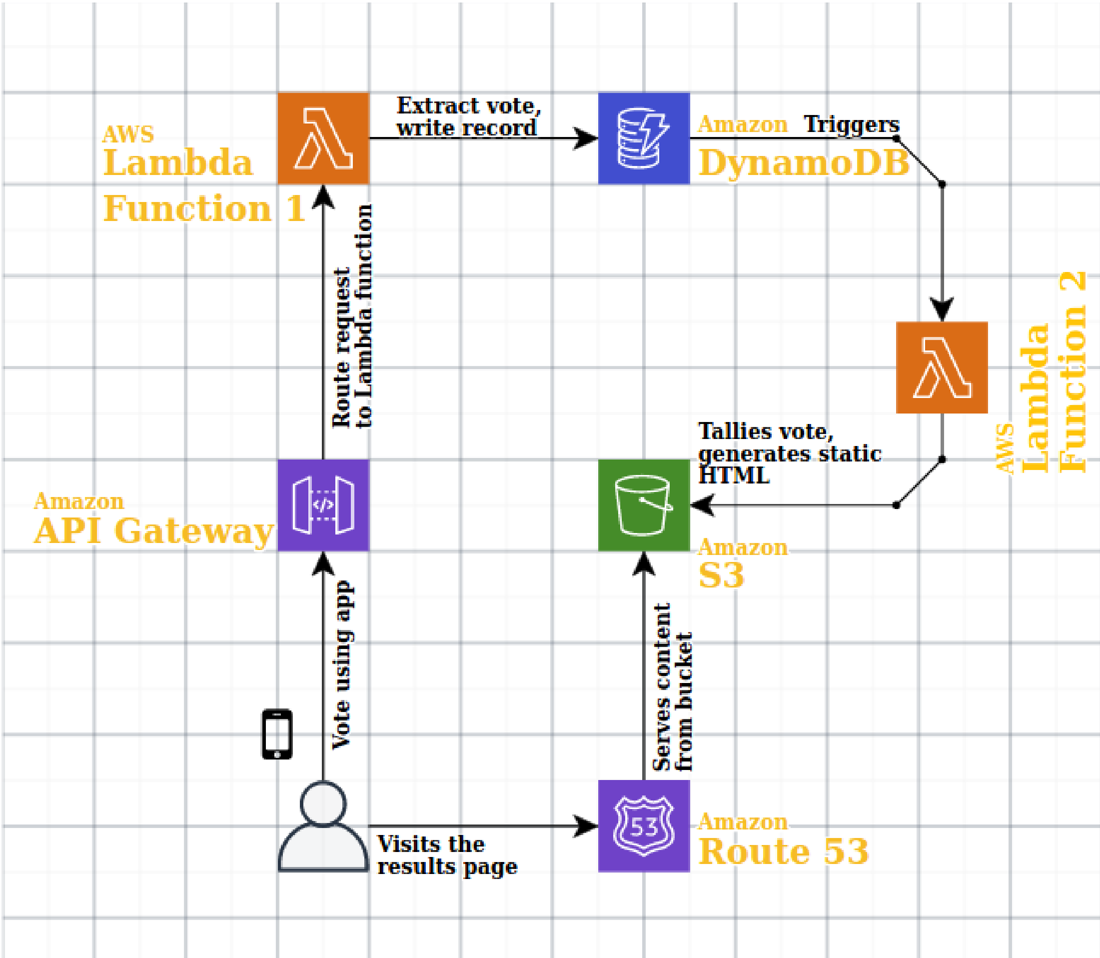

Figure 1: The serverless application architecture (retrieved from https://app.cloudcraft.co/) .

The cloud-based components that were developed and tested during this period were also applied in the learning process. The course "Cloud Computing Technologies" was designed and introduced in National University of Life and Environmental Sciences of Ukraine for training computer science bachelors. The students learned how to build cloud-based components on virtual machines using AWS and Azure platforms. The methodology of open learning and research platform implementation proved to be effective.

The next step of the research was the creation of the serverless hybrid cloud architecture to support collaborative research with Kyiv Glushkov Institute of Cybernetics of the NAS of Ukraine. The goal was to use lambda-functions for sound signal processing and analysis. Figure 2 shows an example of a sound signal oscillogram generated by a lambda-function.

The serverless environment was used for the following tasks:

1. A Python-based web application was created using the Flask framework and tested on localhost.

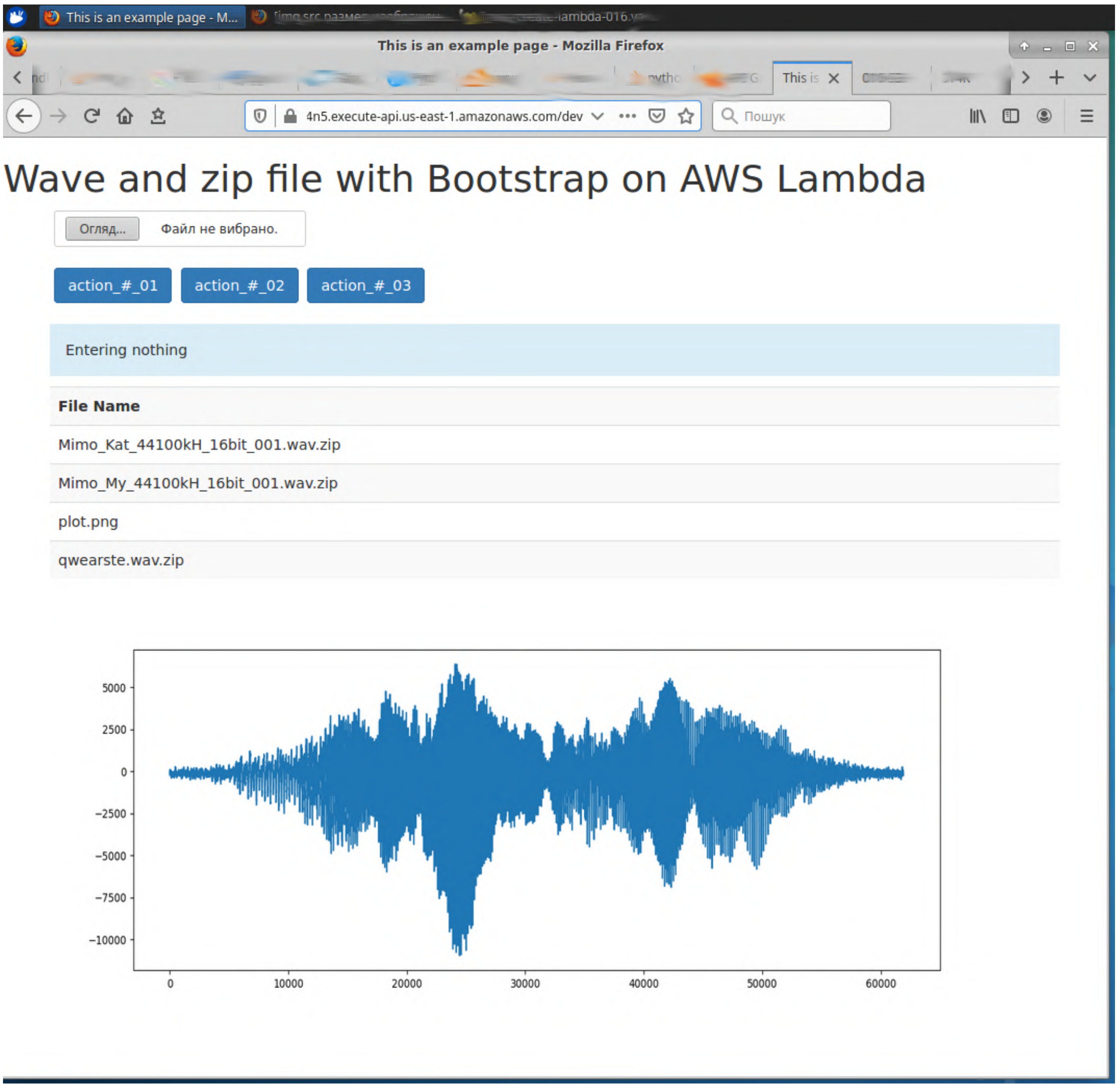

Figure 2: The result of the lambda-function processing the wave file.

2. A user account with necessary permissions was created in the AWS console to secure future applications. An S3 bucket and an EC2 server were also created in the AWS console. The working folder with the Python script (or another compatible language for AWS Lambda) was uploaded to S3.
3. To enable the processing, one or more layers with the required libraries were attached to the lambda-function. The libraries were installed in a virtual environment on the EC2 server. An additional layer was created from this environment. AWS Lambda also provides some freely distributable layers that can be used in future applications.
4. A YAML file was created using CloudFormation tool to specify the available resources for the application. The YAML file created a separate role for working with the future

application.

4.1 Using this role, a lambda-function was created, and its code was downloaded from the zip file created in S3.

4.2 Using this role, an API Gateway was created to allow calling the lambda-function from a browser.

5. The application was debugged and tested.

Using this sequence of steps, a hybrid environment with lambda-function was created and tested for sound signal processing.

3. Conclusion

The paper has presented the rationale and the methodology for introducing cloud technologies in the educational and research process of higher education institutions, as well as the design and implementation of the learning and research environment based on these technologies. The paper has demonstrated how cloud technologies can enhance the access to electronic educational resources, improve the efficiency of ICT infrastructure, and support open education and open science principles. The paper has also proposed a novel approach for using serverless technologies to provide cloud services for data processing, visualization and retrieval, which is a relevant and promising area of development and modernization of the university open learning and research environment.

The paper has reported the experience of developing and applying various cloud-based components for educational and scientific purposes based on the proposed architecture of the hybrid cloud-based environment with lambda-functions. The paper has shown how lambda-functions can enable flexible and scalable computing resources for learning or research tasks that require dynamic and unpredictable computing power.

This approach still needs further implementation and evaluation in different contexts and domains. Future work will focus on expanding the functionality and usability of the cloud-based components, as well as assessing their impact on learning outcomes and research quality.

References

[1] European Research Area (ERA) Roadmap 2015-2020, 2015. URL: https://era.gv.at/era/era-roadmap/european-era-roadmap-2015-2020/.

[2] M. Popel, S. V. Shokalyuk, M. Shyshkina, The Learning Technique of the SageMath-Cloud Use for Students Collaboration Support, in: V. Ermolayev, N. Bassiliades, H. Fill, V. Yakovyna, H. C. Mayr, V. S. Kharchenko, V. S. Peschanenko, M. Shyshkina, M. S. Nikitchenko, A. Spivakovsky (Eds.), Proceedings of the 13th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer, ICTERI 2017, Kyiv, Ukraine, May 15-18, 2017, volume 1844 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2017, pp. 327–339. URL: https://ceur-ws.org/Vol-1844/10000327.pdf.

[3] K. Vlasenko, O. Chumak, D. Bobyliev, I. Lovianova, I. Sitak, Development of an Online-Course Syllabus "Operations Research Oriented to Cloud Computing in the CoCalc System", in: A. Bollin, H. C. Mayr, A. Spivakovsky, M. V. Tkachuk, V. Yakovyna, A. Yerokhin, G. Zholtkevych (Eds.), Proceedings of the 16th International Conference on ICT in Education, Research and Industrial Applications. Integration, Harmonization and Knowledge Transfer. Volume I: Main Conference, Kharkiv, Ukraine, October 06-10, 2020, volume 2740 of *CEUR Workshop Proceedings*, CEUR-WS.org, 2020, pp. 278–291. URL: https://ceur-ws.org/Vol-2740/20200278.pdf.
[4] V. Y. Bykov, M. P. Shyshkina, The conceptual basis of the university cloud-based learning and research environment formation and development in view of the open science priorities, Information Technologies and Learning Tools 68 (2018) 1–19. URL: https://journal.iitta.gov.ua/index.php/itlt/article/view/2609. doi:10.33407/itlt.v68i6.2609.
[5] V. Bykov, D. Mikulowski, O. Moravcik, S. Svetsky, M. Shyshkina, The use of the cloud-based open learning and research platform for collaboration in virtual teams, Information Technologies and Learning Tools 76 (2020) 304–320. URL: https://journal.iitta.gov.ua/index.php/itlt/article/view/3706. doi:10.33407/itlt.v76i2.3706.
[6] V. Oleksiuk, O. Oleksiuk, The practice of developing the academic cloud using the Proxmox VE platform, Educational Technology Quarterly 2021 (2021) 605–616. doi:10.55056/etq.36.
[7] T. Vakaliuk, O. Spirin, O. Korotun, D. Antoniuk, M. Medvedieva, I. Novitska, The current level of competence of schoolteachers on how to use cloud technologies in the educational process during COVID-19, Educational Technology Quarterly 2022 (2022) 232–250. doi:10.55056/etq.32.
[8] V. Velychko, E. Fedorenko, N. Kaidan, V. Kaidan, Application of cloud computing in the process of professional training of physics teachers, Educational Technology Quarterly 2021 (2021) 662–672. doi:10.55056/etq.38.
[9] P. P. Nechypurenko, S. O. Semerikov, O. Y. Pokhliestova, Cloud technologies of augmented reality as a means of supporting educational and research activities in chemistry for 11th grade students, Educational Technology Quarterly 2023 (2023) 69–91. doi:10.55056/etq.44.
[10] R. Tarasenko, S. Amelina, S. Semerikov, L. Shen, Creating a cloud-based translator training environment using Memsource, Educational Technology Quarterly 2022 (2022) 203–215. doi:10.55056/etq.33.
[11] |foster, 2019. URL: https://www.fosteropenscience.eu/.
[12] C. Bargmann, Serverless & faas, 2018. URL: https://users.informatik.haw-hamburg.de/~ubicomp/projekte/master2018-gsem/Bargmann/folien.pdf.
[13] S. Svetsky, O. Moravcik, P. Tanuska, The Knowledge Management IT Support: From Personalized to Collaborative Approach, in: Proceedings of the 14th International Conference on Intellectual Capital, Knowledge Management & Organisational Learning, Academic Conferences and Publishing International Limited, 2017, pp. 253–260.
[14] M. Roberts, J. Chapin, What Is Serverless?, O'Reilly Media, Inc., 2017.
[15] E. Jonas, J. Schleier-Smith, V. Sreekanti, C.-C. Tsai, A. Khandelwal, Q. Pu, V. Shankar, J. M. Carreira, K. Krauth, N. Yadwadkar, J. Gonzalez, R. A. Popa, I. Stoica, D. A. Patterson, Cloud Programming Simplified: A Berkeley View on Serverless Computing, Technical

Report UCB/EECS-2019-3, University of California, Berkeley, 2019. URL: https://www2.eecs.berkeley.edu/Pubs/TechRpts/2019/EECS-2019-3.pdf.

[16] E. van Eyk, L. Toader, S. Talluri, L. Versluis, A. Uță, A. Iosup, Serverless is More: From PaaS to Present Cloud Computing, IEEE Internet Computing 22 (2018) 8–17. URL: https://michael.tsikerdekis.com/downloads/10.1109.MIC.2017.265102442.pdf#page=9. doi:10.1109/MIC.2018.053681358.

[17] M. S. Kurz, Distributed double machine learning with a serverless architecture, in: Companion of the ACM/SPEC International Conference on Performance Engineering, ICPE '21, Association for Computing Machinery, New York, NY, USA, 2021, p. 27–33. doi:10.1145/3447545.3451181.

[18] P. Aditya, I. E. Akkus, A. Beck, R. Chen, V. Hilt, I. Rimac, K. Satzke, M. Stein, Will Serverless Computing Revolutionize NFV?, Proceedings of the IEEE 107 (2019) 667–678. URL: https://www.ruichuan.org/papers/serverless-ieee19.pdf. doi:10.1109/JPROC.2019.2898101.

[19] S. Bebortta, S. K. Das, M. Kandpal, R. K. Barik, H. Dubey, Geospatial serverless computing: Architectures, tools and future directions, ISPRS International Journal of Geo-Information 9 (2020) 311. URL: https://www.mdpi.com/2220-9964/9/5/311. doi:10.3390/ijgi9050311.

[20] G. Casale, M. Artač, W.-J. van den Heuvel, A. van Hoorn, P. Jakovits, F. Leymann, M. Long, V. Papanikolaou, D. Presenza, A. Russo, S. N. Srirama, D. A. Tamburri, M. Wurster, L. Zhu, RADON: rational decomposition and orchestration for serverless computing, SICS Software-Intensive Cyber-Physical Systems 35 (2020) 77–87. doi:10.1007/s00450-019-00413-w.

[21] A. Ortiz, Architecting Serverless Microservices on the Cloud with AWS, in: Proceedings of the 50th ACM Technical Symposium on Computer Science Education, SIGCSE '19, Association for Computing Machinery, New York, NY, USA, 2019, p. 1240. doi:10.1145/3287324.3287533.

[22] Building applications with serverless architectures, 2019. URL: https://aws.amazon.com/lambda/serverless-architectures-learn-more/.