Original Article

A Cloud based Database and its Features for Biological Data

Dr.D. Ponmary Pushpa Latha[1], S. Princy Suganthi Bai[2], T. Nancy Lydia[3], Dr.M. Vargheese[4]

[1]Associate Professor, Department of Digital Sciences, Karunya Institute of Technology and Sciences, Coimbatore. E-mail: ponmarymca@gmail.com

[2]Assistant Professor, Department of Computer Science, Sarah Tucker College, Tirunelveli. E-mail: vinodhprincy2003@gmail.com

[3]Assistant Professor, Department of Information Technology, Francis Xavier Engineering College. E-mail: nancylydia.thanarajt@gmail.com

[4]Professor, Department of Computer science and Engineering, PSN College of Engineering and Technology. E-mail: nellaivarghese@gmail.com

Abstract

Cloud computing provides the availability of computer resources. It is used for data storage and computing power. Large clouds have spread over many location and each location is considered as a data center. The main advantage of cloud computing is to reduce the costs that enlarge the business. Amazon Web Services provide IT infrastructure services. Databases and data management is one of the amazon services. It gives the relational database management system for the users of aws. Amazon S3 is simple storage service. In this research work, biological data is focused. Bioinformatics is one of the application of tools which is used for the computational and analysis to capture the data. It combines many fields such as computer science, mathematics, physics etc., Different user, policy creation, data manipulation for ligand data, bucket creation etc., features are developed for biological data. More focus has been given for the ligand data.

Keywords: Biological Data, Ligand Data, Platform as a Service.

DOI: 10.47750/pnr.2022.13.S03.008

INTRODUCTION

Cloud computing provides the availability of computer resources. It is used for data storage and computing power. Large clouds have spread over many location and each location is considered as a data center [1-4]. Every type of organizations are using cloud computing. This can be used for many process like data backup, disaster recovery etc., Even health care companies and the financial companies are using the cloud. There are different benefits are there in cloud computing. Agility, Elasticity, Cost saving and Deploy globally in minutes are the benefits of cloud computing. Easy access can be done because of agility in cloud computing. The resources can be scaled up or down and also it can be grown or shrinked based on the business needs. The cost can be saved. Because of the cloud, the application is deployed anywhere. The cloud is giving variety of services such as Infrastructure as a Service (IaaS), Platform as a Service (PaaS).

Amazon Web Service

The main advantage of cloud computing is to reduce the costs that enlarge the business. Amazon Web Services provide IT infrastructure services. They can give the results in faster way. AWS is easy to use. AWS cloud provider is much easier than other providers like google cloud platform etc., AWS provides good services. AWS uses different layer for security, so it provides security. Following are the different layers: Data protection, Identity and access management, infrastructure protection, threat detection and continuous monitoring, compliance and data privacy. Instead of using our own servers, user can use the aws servers. So it is cost effective [5-8].

Bioinformatics

Bioinformatics is one of the application of tools which is used for the computational and analysis of data to provide the best result. It combines many fields such as computer science, mathematics, physics etc., Biological Databases are having the collection of biological sciences, and it is gathered from literature, scientific experiments and computational analysis. It has the different types such as structure, sequence and other databases. Protein data bank is categorized under the structured databases (Fig 1).

Protein Data Bank

Protein Data Bank (PDB) is one of the database and it has three- dimensional structural data and it is used to store the

large moleculues (Fig 2).

Protein

Protein plays many critical roles in the body. They are doing the work of structure, function and regulating the body's tissues and organs (Fig 3).

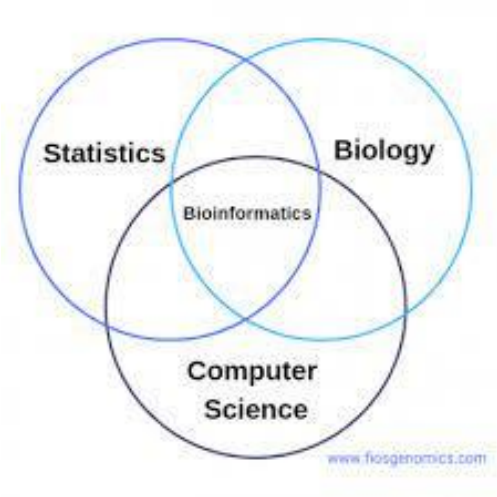

Fig. 1: Bioinformatics

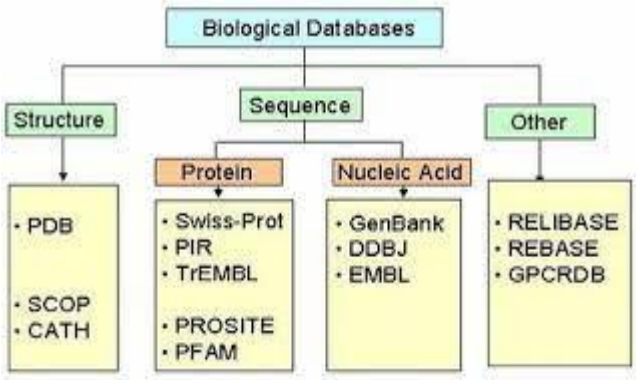

Fig. 2: Biological databases

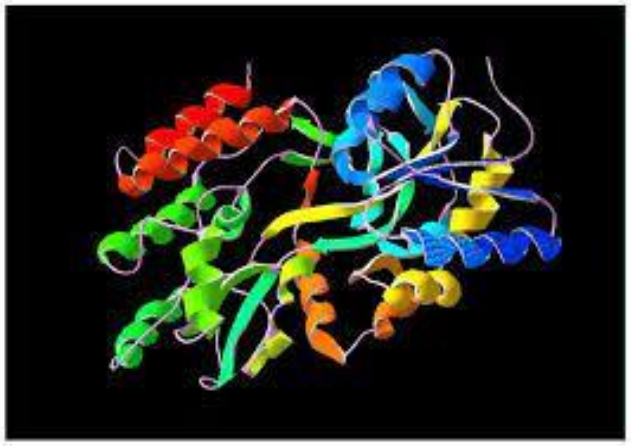

Fig. 3: IFQA - Protein

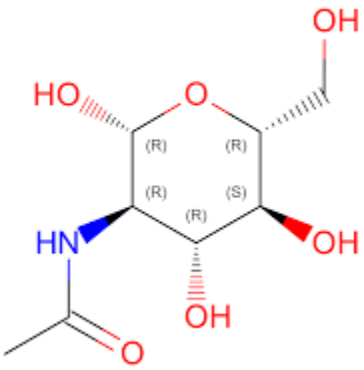

Fig. 4: NAG -Ligand

METHODOLOGY

A database is the resource which is used to store the data and it is used for easy access. The user can able to update the information. The ligand comprise of three sections: COMPOUND, REACTION and ENZYME. Amazon web service is the current emerging platform and that gives many service like infrastructure, platform and software as service. Databases and data management is one of the amazon services. It gives the relational database management system for the users of aws. Amazon S3 is simple storage service. It provides scalability, availability, security and performance of the data. There are many features of aws are needed for biological data. Some features has been incorporated in this research work. 1. Separate user can be created according to the different role. 2. Data can be stored and manipulated 3. Different policy can be defined 4. Buckets can be created to store the data. The data can be stored for disease specific compound or any structurally similar compounds. Here the front end is used as the client. Client is connected with backend through internet. Backend is designed with the following features: management, service, storage, application, security.

RESULT AND DISCUSSION

In cloud, biological data can be stored and it gives the different services and platform to the biological data. The data is secure, since it gives the user, policy and role to the user. Fig 5 provides the manipulation of files. In Fig 6, various compounds are displayed. In fig 7, Various user (Fig 10) can be created according to the hierarchy and the password can be changed (Fig 8). In fig 9, Buckets are created for data storage.

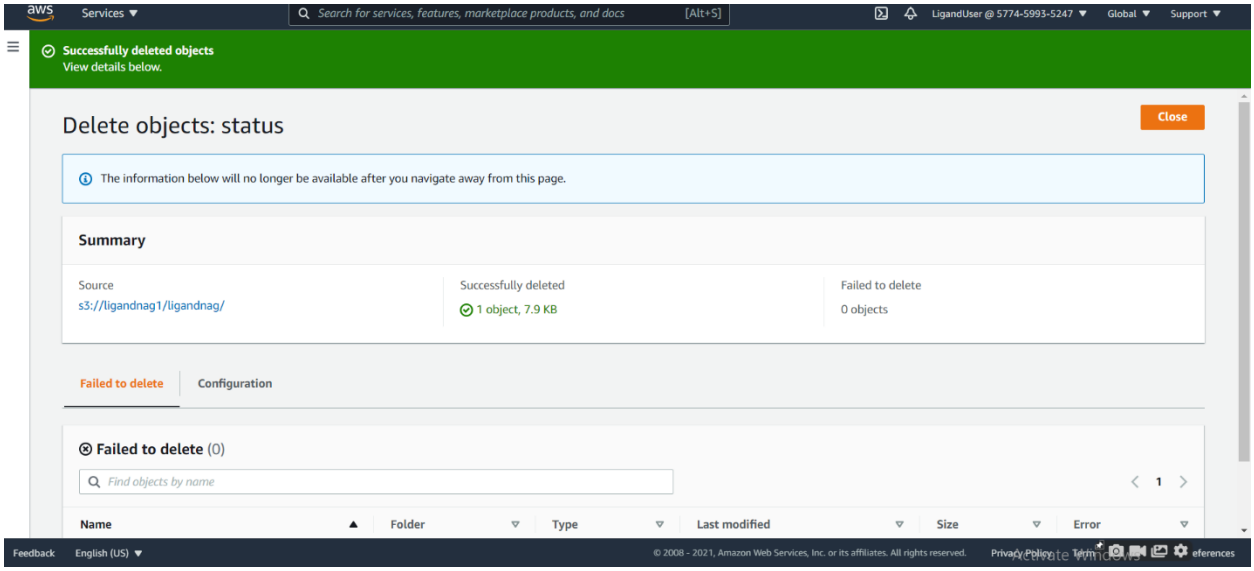

Fig. 5: Manipulation of files

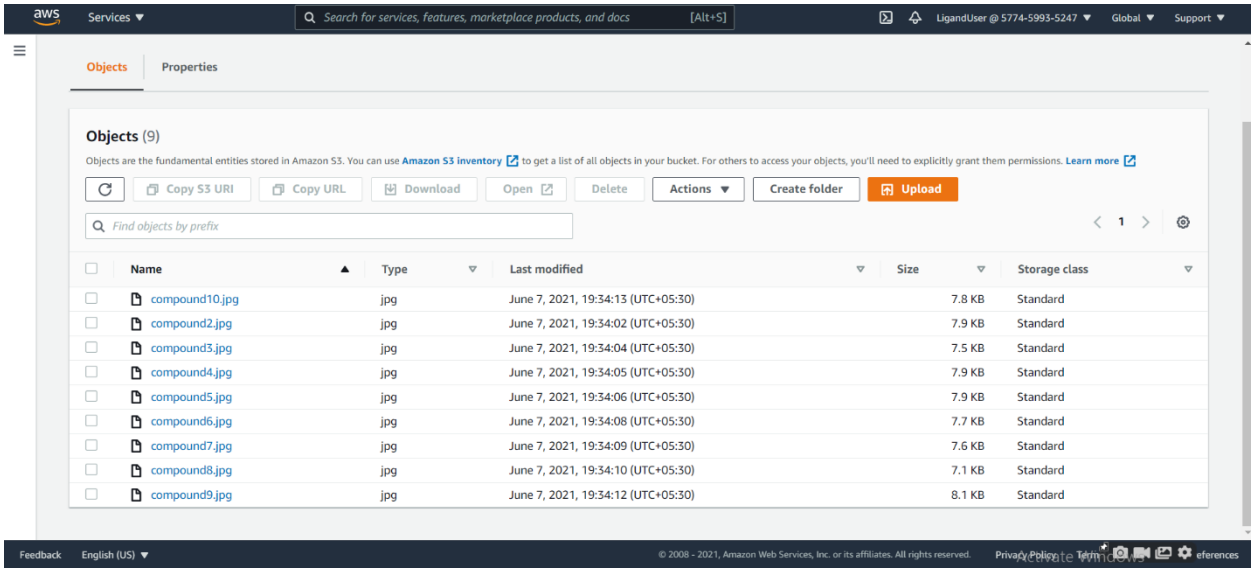

Fig. 6: Compound details

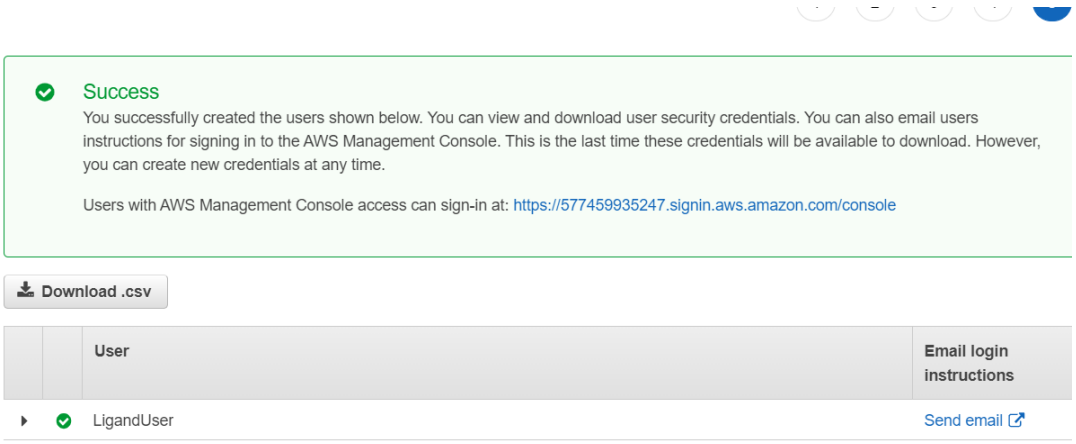

Fig. 7: User Creation

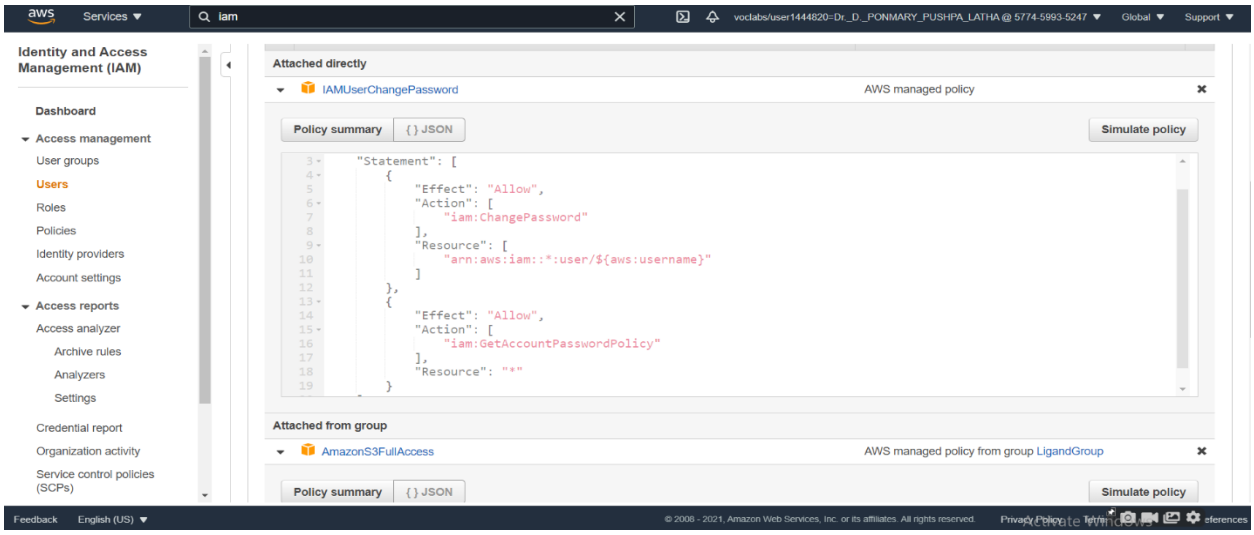

Fig. 8: Password Change

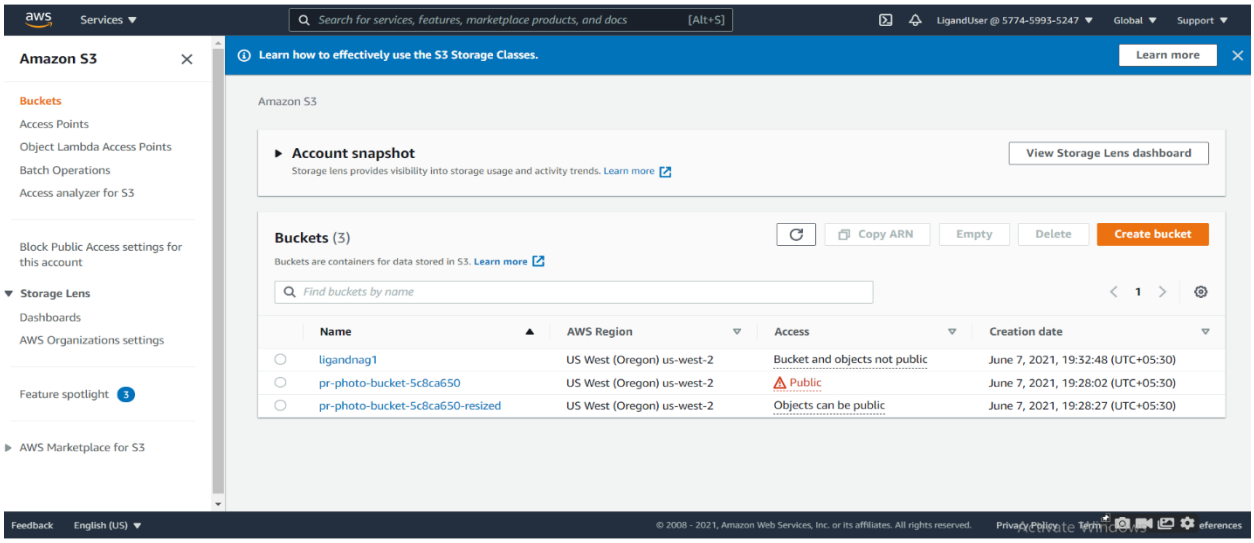

Fig. 9: S3 bucket creation for biological data

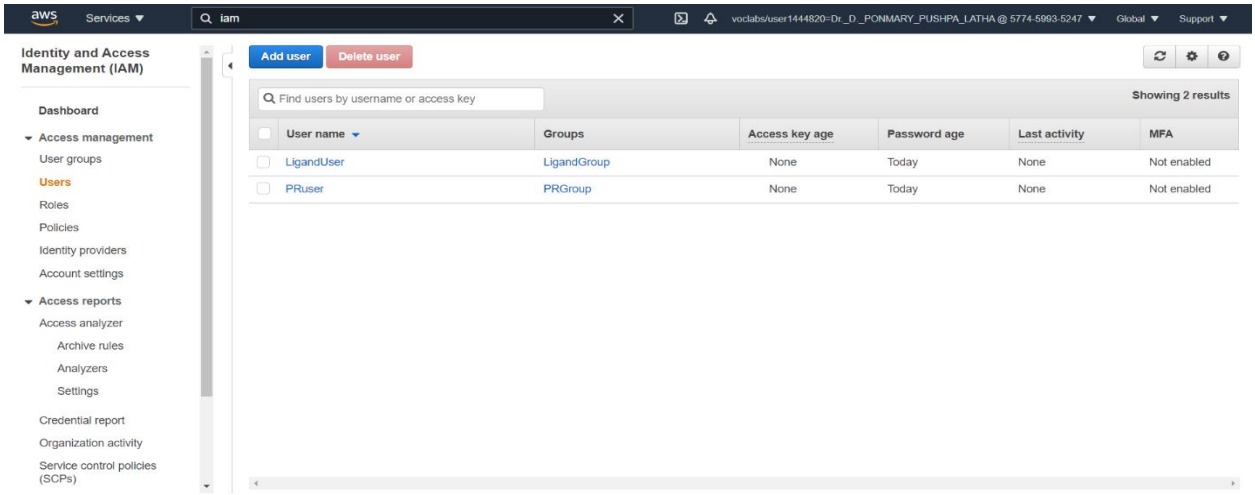

Fig. 10: Different user

Conclusion

Cloud provides the facility to access, store and retrieve the data in the web- enabled interface. At any place, the data is available. The data can be accessed with speed. If suppose the database is developed for a particular disease, when the user want to go with other disease. The database have the features of scalability and it has better security of the environment. When the data is huge, it can be stored easily and the stored data can be scalable and the security of the data is also maintained.

References

Arokia Paul Rajan, R. Service request scheduling based on quantification principle using conjoint analysis and Z-score in cloud, (2018). International Journal of Electrical and Computer Engineering, 8 (2), pp. 1238-1246.

Barroso, L.A., Clidaras, J., Hölzle, U. The datacenter as a computer: An introduction to the design of warehouse-scale machines, second edition (2013) Synthesis Lectures on Computer Architecture, 24, pp. 1-156.

Van Eyk, E., Toader, L., Talluri, S., Versluis, L., Uta, A., Iosup, A. Serverless is More: From PaaS to Present Cloud Computing (2018) IEEE Internet Computing, 22 (5), art. no. 8481652, pp. 8-17.

Baldini, I., Castro, P., Chang, K., Cheng, P., Fink, S., Ishakian, V., Mitchell, N., (...), Suter, P. Serverless computing: Current trends and open problems, (2017) Research Advances in Cloud Computing, pp. 1-20.

Saha, A., Jindal, S., EMARS: Efficient Management and Allocation of Resources in Serverless, (2018) IEEE International Conference on Cloud Computing, CLOUD, 2018-July, art. no. 8457882, pp. 827-830.

Lynn, T., Rosati, P., Lejeune, A., Emeakaroha, V., A Preliminary Review of Enterprise Serverless Cloud Computing (Function-as-a-Service) Platforms, (2017) Proceedings of the International Conference on Cloud Computing Technology and Science, CloudCom, 2017-December, pp. 162-169.

McGrath, G., Brenner, P.R. Serverless Computing: Design, Implementation, and Performance (2017) Proceedings - IEEE 37th International Conference on Distributed Computing Systems Workshops, ICDCSW 2017, art. no. 7979855, pp. 405-410.

Castro, P., Ishakian, V., Muthusamy, V., Slominski, A. Serverless Programming (Function as a Service) (2017) Proceedings - International Conference on Distributed Computing Systems, art. no. 7980271, pp. 2658-2659.

H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T.N. Bhat, H. Weissig, I.N. Shindyalov, P.E. Bourne. (2000) The Protein Data Bank Nucleic Acids Research, 28: 235-242.

Ibrahim, S. (2022). Mathematical Modelling and Computational Analysis of Covid-19 Epidemic in Erbil Kurdistan Using Modified Lagrange Interpolating Polynomial. International Journal of Foundations of Computer Science, 1-17.