A SURVEY OF SOLUTIONS AND FUTURE DIRECTIONS FOR THE CHALLENGES OF IMPLEMENTING ERASURE CODES IN CLOUD AND FOG COMPUTING

Rana Al-Shami[1], Adnan. Al-Mutawkkil[2]
[1,2]Dept. of Information Technology
Faculty of Computer and Information Technology (FCIT), Sana'a University, Sana'a, Yemen

Abstract: More recently, computing storage systems have started implementing an erasure code approach in storing big data due to its many advantages such as reducing storage cost, providing fault tolerance, and reliability. However, many challenges and open issues associated with the use of a code erasure approach in cloud and fog computing environments have arisen. Different methods and techniques for solving these challenges have been presented by many recent studies, however, there are still shortcomings in comprehensive surveys. Therefore, this paper attempts to present a survey of the most important studies in the last five years and discuss the solutions and future trends in this field, in addition to providing a classification of these solutions according to the main research challenges and comparative analysis of the proposed solutions with respect to objectives, environments, tools, and codes used. The ultimate goal of this paper is to provide a summary for researchers for the development of more future work.

Keywords: Erasure Code; Storage Management; Big data; Cloud Computing; Fog Computing

I. INTRODUCTION

Despite the advancements in the field of information technology, the storage of huge data generated by Internet of things applications remains one of the greatest challenges facing modern storage systems [1]. The cloud computing model is one of the most important solutions that have been proposed for storing data, as a result of its many advantages such as high storage capacities, expandability, flexibility, and high capacity for data processing and analysis [2][3]. The cloud's handling of Internet devices' data has been associated with many challenges, such as centralization and delay. Fog computing has emerged as a networked environment extending from the edge of Internet devices to the cloud solving some of the challenges [4]. Cloud and fog computing storage systems use the Data Replication method to ensure data availability and provide fault tolerance. In the method of Data Replication, the cost of storage increases with the increase in the rapid growth of data [5]. Therefore, many cloud service providers have moved to using erasure code to reduce storage costs [6]. Erasure code is based on splitting data into blocks (parent data blocks and parity blocks) that are stored over a number of storage units. Many research focused on improving the performance and efficiency of erasure code and resolving some of the issues and challenges that have coincided with the implementation of erasure code in cloud computing and fog computing environments. However, to the best of our knowledge, there are no recent surveys that focus on these solutions. Hence, this paper reviews the most important research solutions published in the last five years and compares the proposed solutions on the basis of various parameters such as (Objective-environments, tools, and programming languages used - types of coding). Moreover, this paper also discusses open issues and future work to help researchers offer more solutions and future development. In this paper, there are six sections, Section I is an introduction; Section II presents background; Section III reviews the related surveys; Section IV presents the solutions and methods; Section V presents the results and finally, conclusion and future directions are presented in section VI.

.

II. BACKGROUND

A. *Erasure Code:*

Erasure code (EC) is an approach for encoding data and recovering lost data by decoding. Erasure code splits the data into parent K data blocks and parity data blocks that it generates by coding from the original data blocks [7]. The "stripe" is used to refer to storage units (n) consists on a group of equally sized blocks of data and blocks of parity for that data n= (k / m). Figure 1 illustrates the basic operations of an erasure code.
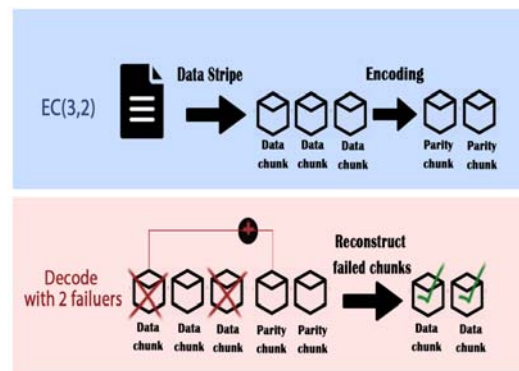

Figure 1.   The Basic Operations of An Erasure Code

The classical erasure codes are divided into two classes: Maximum Distance Separable codes (MDS) e.g (RS codes, X-code, H-Code, Cauchy-RS, code,Tip- code, ,etc…) and e.g (WEAVER ,Code-M ,V2-Code, etc…). MDS code is a linear code that takes into account storage efficiency.

Conversely, the non- MDS needs additional storage space [8].Erasure code was used on local disk storage systems such as Redundant Array of Independent Disks RAID-5, RAID-6, triple Disk Failure Tolerant arrays) to improve reliability and to provide fault tolerance to recover data lost during disk failure [9][10]. Researchers studied the application of erasure code in distributed storage systems to improve the performance and efficiency of systems [11]. Many of the leading companies such as Facebook's HDFS Amazon AWS, Microsoft Azure, and Google Cloud use erasure code approach for storing big data [12]. Erasure codes are proposed for various purposes, including decrease time for data recovery, encoding and decoding, reducing the bandwidth, enhancing data update and consistency of the data. Several open source libraries on erasure codes are available. Table I summarizes some of the most important libraries.

Table I.   Open-Source Erasure Coding Libraries

libraries	language	Erasure Code
Liberasurecode.	C / Pluggable	liberasurecode_rs_vand-Reed-Solomon–Cauchy Reed-Solomon-ISA-L-SHSS-Flat XOR HD-libphazr
JERASURE	C/C++	Reed Solomon-Cauchy Reed Solomon-generator matrices RAID6 codes.
LUBY	C	Cauchy Reed Solomon -
SCHIFRA	C++ /VHDL	Reed Solomon
zfec	C API, Python API, and Haskell API.	several coding
Backblaze	java	Reed Solomon
Intel ISA-L	C/ Assembly	Reed Solomon
Gflib	C	Galois - Reed Solomon

B.   Cloud and Fog Computing

Cloud computing is an advanced paradigm to provide computing services over the Internet to users anytime and anywhere [13]. The services provided by cloud computing service providers such as (Amazon, IBM, Google, Microsoft, Yahoo, VMware, and Rackspace etc.) are divided into three main service categories: Infrastructure(IaaS) , Platform (PaaS) and (SaaS) Software [14]. The cloud enables users to obtain services such as storage, application deployment, and data processing without having to own the infrastructure or data centers, cost savings, reliability, elasticity, security, scalability, on-demand access, self-service and availability [15][16].Despite the advantages offered by cloud computing, there are many challenges related to managing, processing and analyzing IoT data and security issues[18][19]. Cisco proposed a fog computing model in 2014, which then was developed by an OpenFog Consortium [20]. Fog computing consists of a group of networked devices that extend from the edges of IoT devices to the cloud. Figure 2 illustrates the OpenFog architecture [21].
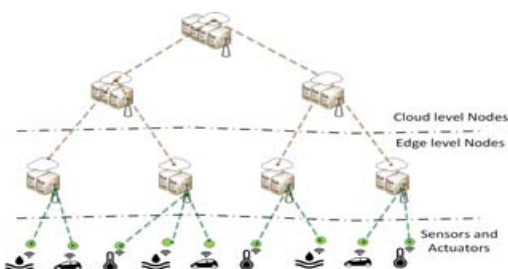
Figure 2.   The Openfog Architecture[21].

Fog computing succeeded in solving some cloud computing issues such as centralization and delay due to the spread of fog computing devices in several areas and their presence near the peripheral devices. Fog computing provides data storage services to reduce service latency, improve network bandwidth, and enhance security [22][23]. The data replication approach was used to increase data availability and fault tolerance however with the increasing amount of data associated with the increase in the number of IoT devices [24].Storage systems moved to use erasure code to reduce storage cost and reliability. Erasure code was used in local storage systems however, there are still many issues that need further study related to implementing erasure code in modern computing environments [25][26][27].

III.   RELATED SURVEYS

There are only two papers related to our work and both surveys didn't include recent research over the past five years. In the first paper, the authors focused on the challenges of implementing erasure codes in the cloud environment and mentioned many open problems and future work [28], while in the second paper the authors provided an overview of developments in coding mechanisms that were categorized into hierarchical codes, self-repairing codes, and renewable codes. They also discussed the metrics that should be taken into account when developing evaluation techniques [29].

IV.   SOLUTION AND METHODS

The methods and solutions for the challenges of implementing erasing codes in cloud and fog environments are classified according into three different categories as shown in Figure 3.
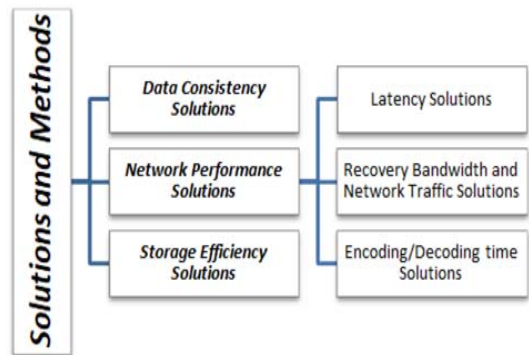
Figure 3.   Classification Of the Solutios

C.   Data Consistency Solutions

With the updating process of data blocks, codecs need to update the blocks of encoder and write new versions of the data for all blocks. Ensuring data consistency and reducing data update time is one of the most important challenges related to the data updating process, especially in large or distributed data. Therefore, some researchers turn to improve the data update process. In [30] authors suggested a hybrid protocol (RAPID) of locking and buffering mechanisms to speed up data updating. They used prediction-based heuristics to define and decrease the number of blocks updated .The results showed that the protocol improved data retrieval by 30% and predictive accuracy by 80%. In the same context author proposed a scheme (TUpdate) based on the a rack_aware tree construction technique consists of the data node as root and parity nodes as children to organize the

data connections and top-down data transmission to reduce update time. The proposed scheme reduces by 27% and 32% the total update time compared to both S-Update-D/P [31].

Table II. Objectives and Properties of the Proposed Solutions to Improve the Data Consistency

Papers	Objectives	Erasure Code	Platform	Tools/ Programming Language / Dataset
[30]	Data Retrieval Update Time Prediction Accuracy	RS (10, 8)	Cloud	Python
[31]	Update Time	CRS(14,10)(13,9),(12,8), (11,7), (14, 10)	Cloud	HDFS-Raid Openstack

D. Network Performance Solutions

1) Latency Solutions: In [32] authors Presented a proactive scheme (ProCode) based on predicting drive failure by monitoring the drive to determine the redundancy factor for the data or blocks expected to be lost and Implement the erasure code on the rest of the data .Compared with 3-way replication storage, the presented scheme reduced storage costs and the number of data loss events reduced read latency by 63% and recovery time by 78%.In [33] Authors introduced a system (EC-Store) and proposed a dynamic strategy for developing data based on access patterns and data traffic improved latency by 40% compared to baseline erasure code. Reduced latency by 20% compared to replication. In [34] authors formulated a problem to reduce the average response and iterative algorithm to solve the problem, they also proposed a framework to reduce the average latency of erasure-coded. Furthermore the proposed framework reduced the mean time. In [35] authors presented the ProactiveCach technology, which is based on a proactive approach to copy device objects that are expected to fail to cache memory. The proposed method increases 37 per cent throughput and reduced reading latency by up to 38 per cent. In [36] authors proposed a framework for developing mean and tail latency in erasure-coded storage systems with content (cache) storage at the edge and they also proposed a temporary storage policy called time-to-live (TTL) using probability server selection techniques to define bounds for mean-tail latency as well as an optimization problem formula and an optimization algorithm to solve it.

2) Recovery Bandwidth and Network Traffic Solutions : In [37] authors presented a hybrid code (AZ-Code) that combines MSR code to generate local parity mapping and LRC to generate the equivalence profile for multiple availability regions. When one node fails, the proposed code reduces the recovery time by78% compared with RS code and 34.81% compared with LRC when Double Node Failure, the recovery time is reduced 58.77% compared with RS code and 45.34% compared with LRC codes. In [38] authors proposed a distributed storage scheme using the Information-Centric Fog Network mechanism to retrieve blocks of data encoded by Binary Reed-Solomon to reduce cost storage. The mechanism is used in the data searching

and indexing processes to make the process more efficient both in time and in network bandwidth resources. In [39]authors suggested multi-level framework Zebra that encodes data into multiple layers on demand to reduce hot data retrieval cost and reduce the cost of cold data storage and low network transfer in the migration. The framework saves the average reconstruction overhead by 53.7% and 54.9% for the top 15% and reduces the network transfer such that only 2.2%.In [40] authors suggested improving the performance of parallel recovery by applying three different methods: chunk allocation, asynchronous recovery method and Contention Avoidance. The proposed method improved the recovery performance by 194%. In [41]authors used a framework similar to the work of MapReduce (EC-Shuffle) and focused on transferring parity chunks only while performing shuffle operations via two erasure cipher (forward encryption and back cipher. The results showed 50% drop in network traffic and achieved up to 38% performance improvement than FTI. Authors [42] worked to improve retrieval time by the Parallel Enumeration Method and the Parallel Instruction Method to compute the approximate optimal generator matrix and online data positioning scheme in a Parallel Heuristic method to reduce the network cost of repair. Compared with ACL, DFC, Pyramid and Xorbas, NEC can reduce the repair cost and repair time by 26.8%–40% and 13.9%-37.1% and HRepair further reduces the cost of repair time. In [43] authors proposed a hybrid method of redundancy and erasure code based on a proactive prediction approach to retrieve data from expected failure devices. The proposed approach applies slow recovery of cold data and copying of hot data to memory .Their results showed improvement in average recovery bandwidth.

3) Encoding/Decoding time solutions: In [44] authors proposed the Carousel coding method to allow data to be read in parallel from multiple servers without the need for decryption. Compared to RS code, their results showed the file retrieval time was improved by 29.0% of the RMS and reduced decoding time by 75.4%. In [45] authors improved a code based on XOR code using two methodologies (parity block repetition and rack awareness. They proposed a heuristic method for optimizing block placement with optimized code to speed up the repair process and recover lost data contributed to improved recovery performance for RS codes, XORBAS and Multi-XOR.

Table III. Objectives and Properties of the Proposed Solutions to Improve Network Performance

Papers	Objectives	Erasure Code	Platform	Tools/ Programming Language/ Dataset
[32]	Read Latency Recovery Time	RS(10,4),(6,3),(6,2),(10,2)	Cloud	HDFS-RAID Facebook
[33]	Latency	RS(2,2)	Cloud	C++ Apache Thrift . Jerasure 2.0 YCSB-E. Workload & Wikipedia Workloads Benchmarks

[34]	Mean Latency	MDS EC(7, 4)	Cloud	Tahoe Amazon S3
[35]	Throughput Reading Latency	RS(14,10,(9, 6),(6, 4)	Cloud	Ceph Storage Cluster
[36]	Reading Latency	EC(7,4)	Cloud	Real Cloud Environment Openstack Apache Jmeter Apache Server
[37]	Recovery Time Recovery Bandwidth	AZ Code	Cloud	Hadoop System Intel ISA-L Of HDFS
[38]	Cost Storage Network Bandwidth Resources.	RS(3,3), (12,12), (24,24),	Fog	C++
[39]	Reconstructio n Overhead Network Transfer	RS(4,2),(12 ,2)	Cloud	Mapreduce Hadoop Facebook
[40]	Recovery Performance	MDS EC(8,2),(4, 2),(2,1)	Cloud	Servers Exstorusfs Glusterfs 3.7.20
[41]	Network Traffic Recovery Performance	RS(4, 2)	Cloud	Bigdatabench Hadoop Java Intel Isa-L Wikipedia Matrix Pagerank Google Web Graph Facebook Social Network Movielens 20m Pascal Large-Scale Challenge
[42]	Cost And Repair Time	Google Colossus(9, 6) HDFS-Xorbas(16, 10) Yahoo Objects Store(11, 8) Hadoop3.0 (14, 10) Rdpcode (10, 6) DFC(35, 20)	Cloud	Hrepair Ucloud Hadoop 3.0.3's
[43]	Average Recovery Bandwidth	RS(14, 10)	Cloud	Solomon Ds-Sim Simulator
[44]	Retrieval Time Decoding Time	Carousel Codes(12, 6, 10, 6) RS(12, 6)	Cloud	C++ With ISAL Apache Hadoop Amazon
[45]	Retrieval Time Encoding /Decoding Time	RS(10,4) XORBAS LRC(10 , 6) MXOR(10 , 7)	Cloud	National Research Cloud Nectar Hadoop

E. Storage Efficiency Solutions

One of the storage efficiency challenges affecting the lost data recovery process is determining the size and number of data blocks. In [46] authors proposed a framework for erasure coding videos in cloud systems based on Approximate Code to distinguish the importance of data, applying triple parities to the important data and applying single / double parities to unimportant data .compared to

erasure codes such as RS, LRC, STAR and TIP-Code, Their results showed the approximate storage cost was improved by 20.8%, Code reduces the number of parities by up to 55% and increased data recovery by 5%. In [20] authors proposed a hybrid fog and cloud storage system scheme using erasure code to encrypt the file's condensed data and distributed it across multiple cloud servers .

Table IV. Objectives and Properties of the Proposed Solutions to Improve Storage Efficiency

Papers	Objectives	Erasure Code	Platform	Tools/ Programming Language / Dataset
[46]	Storage Efficiency Increase Data Recovery	Approximate Codes Over APPR.RS/L RC /TIP/STAR (K, 1, 2, 4), APPR.RS/ LRC/TIP/ST AR(K,1,2, 6)	Cloud	Hadoop Systems Youtube
[47]	Storage Efficiency	RS(5,4), RS(5, 3)	Fog	Java Library Blackblaze_RS Amazon S3 Alibaba Cloud OSS BaiduCloud BOS Tencent_Cloud COS Qiniu_Cloud KODO

V. RESULTS

In this paper, the main impetus was to review the most recent research that have been published in prominent journals in the last five years from 2015 to 2020 and discussed solutions to the most important challenges related to the implementation of erasure codes in cloud and fog computing. Figure 4 illustrates the maximum amount of research was in solutions to the challenges and problems of implementing erasure code in cloud and fog environments related to network performance such as Latency, Network traffic, Recovery Bandwidth and, Encoding/Decoding time, whereas very little research focuses on solving challenges related to storage efficiency, updating data and ensuring consistency. Figure 5 shows the percentage of platforms used in the papers.

Table II summarizes the objectives of the improvement and properties of the proposed research solutions to solve some of the challenges related to updating data and ensuring consistency. Table III summarizes the objectives of the improvement and properties of the proposed research solutions to solve some of the challenges related to network performance. Table IV summarizes the objectives of the improvement and properties of the proposed research solutions to solve some of the challenges related to storage efficiency. Table V shows comparison of scientific papers solutions according to the main research challenges and problems.
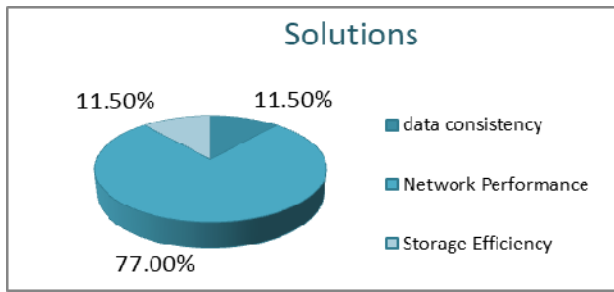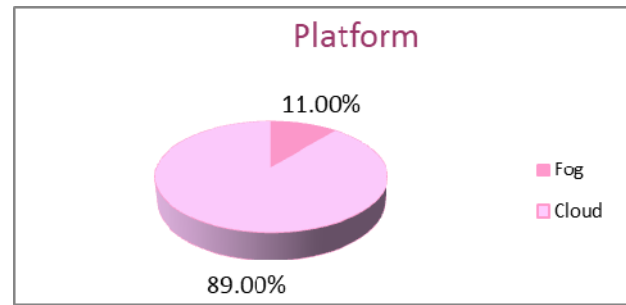
Figure 4.   Categories of Paper Solutions.

Figure 5.   A Platform for Solutions.

Table V.   Comparison of scientific papers' solutions according to the main research challenges and problems

Papers	Update Time	Data Consistency	Cost Retrieval	Recovery Time	Recovery Bandwidth	Network traffic	Encoding /Decoding Time	Throughput	latency	Cost Storage
[30]	√	√	–	–	–	–	–	–	–	–
[31]	√	–	–	–	–	–	–	–	–	–
[32]	–	–	–	√	–	–	–	–	√	–
[33]	–	–	–	–	–	–	–	–	√	–
[34]	–	–	–	–	–	–	–	–	√	–
[35]	–	–	–	–	–	–	–	√	√	–
[36]	–	–	–	–	–	–	–	–	√	–
[37]	–	–	–	√	√	–	–	–	–	–
[38]	–	–	–	–	√	–	–	–	–	√
[39]	–	–	–	–	–	√	–	–	–	–
[40]	–	–	–	√	–		–	–	–	–
[41]	–	–	–	√	–	√	–	–	–	–
[42]	–	–	√	√	–	–	–	–	–	–
[43]	–	–	–	–	√		–	–	–	–
[44]	–	–	–	√	–	–	√	–	–	–
[45]	–	–	–	√	–	–	√	–	–	–
[46]	–	–	√	–	–	–	–	–	–	√
[47]	–	–	–	–	–	–	–	–	–	√

VI.   CONCLUSION AND FUTURE DIRECTIONS

Recently, many researchers have focused on improving the performance and efficiency of the erasure method and solving several challenges that coincided with the emergence of the method in the cloud and fog environment, however, there is still a lack of comprehensive surveys. This paper sheds light on the papers published in the past five years (2015 - 2020) and classifies the proposed solutions according to the challenges that they discussed in three categories: (data consistency solutions, network performance solutions, storage efficiency solutions). In addition, this paper discusses the main idea of the solutions and compares the proposed solutions on the basis of various parameters such as (Objective - environments, tools, and programming languages used - types of coding).

According to this survey, the implementation of the erasure code approach in most of the presented papers is still limited to data archiving and cold data storage due to the delays associated with the encryption and decryption process. Very few studies have provided solutions to use the erase code to store hot data with the help of many modern technologies and models that improve data latency. Therefore, improving the performance and efficiency of erasing the code for storing hot data is one of the most important future trends in this field. Moreover, improving

data placement based on task scheduling and customer requirements can contribute to reducing delays. Data-block placement, especially in distributed storage systems, greatly affects the process of data access, which requires more research in optimizing data-block placement and resource allocation. There are many future approaches that some researchers have touched upon, including the use of accurate proactive methods of predicting failure that can improve data refresh or retrieval, by identifying specific data blocks to update or implement data migration to specific nodes to improve traffic. Encoding data at multiple levels can improve the recovery of lost data and the fog environment hierarchy can be exploited in this regard. Research has focused on implementing encryption erasure in computing environments to reduce storage cost, fault tolerance, and recovery process. Energy efficiency and security enhancement using erasure code one of the most open issues.

VII.   REFERENCES

[1]   M. H., Padgavankar and S. R. Gupta, "Big data storage and challenges." International Journal of Computer Science and Information Technologies, vol 5, No. 2 ,pp. 2218-2223, 2014.

[2]   R. Nachiappan,J. Bahman, R. N. Calheiros, and K. M. Matawie.,"Cloud storage reliability for big data applications:

A state of the art survey." Journal of Network and Computer Applications vol. 97 pp. 35-47, 2017.

[3] C Ji, Y Li,W Qiu, U Awada and K Li ,"Big data processing in cloud computing environments." In 2012 12th international symposium on pervasive systems, algorithms and networks, pp. 17-23. IEEE, 2012.

[4] Y. Ashkan,C. Fung, T. Nguyen, K. Kadiyala, F. Jalali, A. Niakanlahiji, J. Kong, and J. P. Jue, "All one needs to know about fog computing and related edge computing paradigms: A complete survey.", Journal of Systems Architecture ,vol. 98, pp. 289-330, 2019.

[5] N. J. Navimipour, B. A. Milani, "Replica selection in the cloud environments using an ant colony algorithm.",Third International Conference on Digital Information Processing, Data Mining, and Wireless Communications (DIPDMWC), pp. 105-110. IEEE, 2016.

[6] C. Aatish, and A. Mungur. "Data Management in Erasure-Coded Distributed Storage Systems." , International Symposium on Cluster, Cloud and Internet Computing (CCGRID), pp. 902-907. IEEE, 2020.

[7] M. Gandelman, Y Cassuto – arXiv. preprint. "Treeplication: An Erasure Code for Distributed Full Recovery under the Random Multiset Channel.", arXiv preprint arXiv:2002.09895 ,2020.

[8] MP. Ramkumar, N. Balaji, GSRE Selvan and R. J. Rohini. "RAID-6 Code Variants for Recovery of a Failed Disk.", In Soft Computing in Data Analytics, pp. 237-245. Springer, Singapore, 2019.

[9] L Xu, J Bruck. "X-code: MDS array codes with optimal encoding." IEEE Transactions on Information Theory, vol..45, No.1, pp. 272-276 , 1999 .

[10] T. Miyamae, T. Nakao, K. Shiozawa, "Erasure code with shingled local parity groups for efficient recovery from multiple disk failures.", Workshop on Hot Topics in System Dependability ,2014.

[11] L. Pamies-Juarez, A. Datta and F. Oggier, "RapidRAID: Pipelined erasure codes for fast data archival in distributed storage systems." In 2013 Proceedings IEEE INFOCOM, pp. 1294-1302. IEEE, 2013.

[12] X. Li, Z. Yang, J. Li, R. Li, PPC. Lee, Q. Huang and Y. Hu. "Repair Pipelining for Erasure-Coded Storage: Algorithms and Evaluation." pp.1908.01527 , 2019.

[13] MU Bokhari, Q Makki, YK Tamandani, "A survey on cloud computing." , Big Data Analytics Springer ,vol. 654,pp. 149-164., 2018.

[14] TL Mokgetse, R Sridaran and R. Sridaran. "A Comparative Study of underlying Threats of Cloud Deployment Service Models." In 2019 6th International Conference on Computing for Sustainable Global Development (INDIACom), pp. 1218-1222. IEEE, 2019.

[15] H. Hassan, A. I. El-Desouky, A. Ibrahim, E. M. El-Kenawy and R. Arnous, "Enhanced QoS-Based Model for Trust Assessment in Cloud Computing Environment," ,IEEE Access, vol. 8, pp. 43752-43763, 2020.

[16] Y. Hu, B. Deng, F. Peng, B Hong, Y. Zhang, and D. Wang, "A survey on evaluating elasticity of cloud computing platform." , World Automation Congress, pp. 1-4. IEEE, 2016.

[17] A. Pradhan, SK. Bisoy, PK. Mallick, "Load Balancing in Cloud Computing: Survey.",Innovation in Electrical Power Engineering, Communication, and Computing Technology, pp. 99-111, 2020.

[18] S. Mustafa, B. Nazir, A. Hayat, SA Madani ,"Resource management in cloud computing: Taxonomy, prospects, and challenges.",Computers & Electrical Engineering,vol. 47, pp. 186-203,2015.

[19] S Basu, A Bardhan, K Gupta, P Saha, M. Pal, M. Bose, K. Basu, S. Chaudhury, and P. Sarkar. "Cloud computing

[20] S, Khanagha, S. Ansari, S. Paroutis, and L. Oviedo. "Mutualism and the dynamics of new platform creation: A study of cisco and fog computing." Strategic Management Journal , 2020.

[21] P. Hu, S. Dhelim, H. Ning, T. Qiu ,"Survey on fog computing: architecture, key technologies, applications and open issues." ,Journal of network and computer applications,vol. 98 ,pp. 27-42, 2017.

[22] T. Huang, W. Lin, Y. Li, LG. He, SL Peng, "A latency-aware multiple data replicas placement strategy for fog computing." Journal of Signal Processing Systems vol. 91,No.10,pp. 1191-1204, 2019.

[23] MI. Bala, MA. Chishti, "Survey of applications, challenges and opportunities in fog computing." International Journal of Pervasive Computing and Communications ,vol. 15 No. 2, pp. 80-96. 2019.

[24] JM. Pérez, F. García-Carballeira, J. Carretero, A. Calderón, and J. Fernández. "Branch replication scheme: A new model for data replication in large scale data grids." Future Generation Computer Systems,vol. 26,No. 1 ,pp. 12-20, 2010.

[25] L. Liang, H. He, J. Zhao, C. Liu, Q. Luo and X. Chu, "An Erasure-coded Storage System for Edge Computing." IEEE Access ,vol.8, 2020.

[26] J. Yue, M. Xiao,"Coding for Distributed Fog Computing in Internet of Mobile Things." IEEE Transactions on Mobile Computing ,2020.

[27] H. Bao, Y. Wang, F. Xu,"An Adaptive Erasure Code for JointCloud Storage of Internet of Things Big Data." IEEE Internet of Things Journal ,vol. 7,No. 3 pp. 1613-1624, 2019.

[28] OT. Lee, SDM. Kumar, and P. Chandran ,"Erasure coded storage systems for cloud storage—challenges and opportunities." In 2016 International Conference on Data Science and Engineering, pp. 1-7. IEEE, 2016

[29] J. Li, B. Li, "Erasure coding for cloud storage systems: A survey." Tsinghua Science and Technology vol.18, No.3,pp. 259-272, 2013.

[30] GJ. Akash, OT. Lee, SDM. Kumar, P. Chandran, and A. Cuzzocrea, "Rapid: A fast data update protocol in erasure coded storage systems for big data." In 2017 17th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing, pp. 890-897. IEEE, 2017.

[31] X. Pei, Y. Wang, X. Ma, F. Xu, "T-update: A tree-structured update scheme with top-down transmission in erasure-coded systems." In IEEE Infocom 2016-the 35th Annual IEEE International Conference on Computer Communications, pp. 1-9. IEEE, 2016.

[32] P. Li, J. Li, RJ. Stones, G. Wang, Z. Li, and X. Liu, "Procode: A proactive erasure coding scheme for cloud storage systems.", Symposium on Reliable Distributed Systems (SRDS), pp. 219-228. IEEE, 2016.

[33] M. Abebe, K. Daudjee, B. Glasbergen, and Y. Tian, "Ec-store: Bridging the gap between storage and latency in distributed erasure coded systems.",International Conference on Distributed Computing Systems, pp. 255-266. IEEE, 2018.

[34] Al-Abbasi, Abubakr O., and Vaneet Aggarwal. "Mean latency optimization in erasure-coded distributed storage systems." In IEEE INFOCOM 2018-IEEE Conference on Computer Communications Workshops, pp. 432-437. IEEE, 2018.

[35] AO. Al-Abbasi, V. Aggarwal,"ProactiveCache: on reducing degraded read latency of erasure coded cloud storage." ,International Conference on Cloud Computing Technology and Science (CloudCom 2019) , pp. 223-230. 2019.

[36] AO. Al-Abbasi and V. Aggarwal, "TTLCache: Taming Latency in Erasure-coded Storage through TTL Caching."

IEEE Transactions on Network and Service Management ,vol. 17, No. 3, pp. 1582-1596, 2020.

[37] X. Xie, C. Wu, J. Gu, H Qiu, J. Li, M. Guo, , X. He, Y. Dong, and Y. Zhao. "AZ-Code: an efficient availability zone level erasure code to provide high fault tolerance in cloud storage systems.", pp. 230-243. IEEE, 2019.

[38] Y. Shu, M. Dong, K. Ota, J. Wu, and S. Liao, "Binary reed-solomon coding based distributed storage scheme in information-centric fog networks." In 2018 IEEE 23rd International Workshop on Computer Aided Modeling and Design of Communication Links and Networks, pp. 1-5. IEEE, 2018.

[39] J. Li and B. Li, "Zebra: Demand-aware erasure coding for distributed storage systems." In 2016 IEEE/ACM 24th International Symposium on Quality of Service (IWQoS), pp. 1-10. IEEE, 2016.

[40] DO. Kim, HY. Kim, YK. Kim and JJ. Kim, "Efficient techniques of parallel recovery for erasure-coding-based distributed file systems." Computing vol.101,No.12, pp. 1861-1884, 2019.

[41] X. Yao, CL. Wang, M. Zhang, "EC-Shuffle: Dynamic Erasure Coding Optimization for Efficient and Reliable Shuffle in Spark." ,CCGRID, pp. 41-51. 2019.

[42] H Bao, Y Wang and F Xu, "Reducing network cost of data repair in erasure-coded cross-datacenter storage.",Future Generation Computer Systems vol. 102 ,pp. 494-506, 2020.

[43] R. Nachiappan, B. Javadi, RN. Calheiros and K. M. Matawie. "Adaptive Bandwidth-Efficient Recovery Techniques in Erasure-Coded Cloud Storage.",European Conference on Parallel Processing, pp. 325-338. Springer, Cham, 2018.

[44] J. Li and B. Li ,"On data parallelism of erasure coding in distributed storage systems." , IEEE 37th International Conference on Distributed Computing Systems (ICDCS), pp. 45-56. IEEE, 2017.

[45] LJ. Mohan, PIS. Caneleo, U. Parampalli, and A. Harwood, "Geo-aware erasure coding for high-performance erasure-coded storage clusters." Annals of Telecommunications vol. 73, No.1-2 pp. 139-152, 2018.

[46] H. Jin, C. Wu, X. Xie, J. Li, M. Guo, H. Lin, and J. Zhang, "Approximate code: a cost-effective erasure coding framework for tiered video storage in cloud systems." , Proceedings of the 48th International Conference on Parallel Processing, pp. 1-10. 2019.

[47] Y. Fu, X. Qiu, J. Wang ,"F2MC: Enhancing Data Storage Services with Fog-toMultiCloud Hybrid Computing.", Performance Computing and Communications Conference (IPCCC), pp. 1-6. IEEE, 2019.