ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

A Survey on Optimizing Cauchy Reed-Solomon Codes for Fault-Tolerant Storage Applications

Varsha A. Abhang, Anil D.Gujar

Department of Computer Engineering, TSSM'S BSCOER, Narhe, Pune, India

ABSTRACT: In this paper, we propose CaCo, an effective Cauchy coding approach for information in the cloud. To begin with, CaCo utilizes Cauchy framework heuristics to create a grid set. Second, for each grid in this set, CaCo utilizes XOR plan heuristics to create a progression of calendars. At last, CaCo chooses the most limited one from all the delivered plans. In such a way, CaCo can distinguish an ideal coding plan, inside the capacity of the current best in class, for a subjective given repetition setup. By influence of CaCo's inclination of straightforwardness to parallelize, we support fundamentally the execution of the determination procedure with bottomless computational assets in the cloud. We execute CaCo in the Hadoop conveyed record framework and assess its execution by contrasting and "Hadoop-EC" created by Microsoft research. Our trial results demonstrate that CaCo can acquire an ideal coding plan inside worthy time. Moreover, CaCo beats Hadoop-EC by 26.68-40.18 percent in the encoding time and by 38.4-52.83 percent in the translating time at the same time.

KEYWORDS: Cloud Storage, Fault Tolerance, Reed-Solomon Codes, Cauchy Matrix, XOR Scheduling.

I. INTRODUCTION

In this segment, we first give a general outline of Cauchy Reed-Solomon Coding. At that point, we give a portrayal in short of the exploration and related work on creating Cauchy networks and encoding with timetables. From these portrayals, we can make clearer the inspiration of our work. Distributed storage is developed of various reasonable and untrustworthy parts, which prompts a decline in the general interim between disappointments (MTBF). As capacity frameworks develop in scale and are sent over more extensive systems, part disappointments have been more basic, also, prerequisites for adaptation to internal failure have been further expanded. Along these lines, the disappointment assurance offered by the standard levels has been no more adequate by and large, also, capacity planners are thinking about how to endure bigger quantities of disappointments. For instance, Google's cloud capacity , Windows Azure Storage Ocean Store , Disk Reduce , HAIL, and others all endure at any rate three disappointments. Our study finds that with low likelihood, one coding plan picked by Rules of thumb, for a given redundancy configuration.

II. LITERATURE SURVEY

1) *Paper Name:*-An Efficient XOR-Scheduling Algorithm for Erasure Codes Encoding

Abstract:-

Enhancing the execution of encoding, the more regular operation. It does as such by planning the operations of XOR based codes to upgrade their utilization of store memory. The system XOR scheduling what's more, exhibit how it applies to a wide assortment of existing codes. We direct an execution assessment of booking these codes on an assortment of processors what's more, demonstrate that XOR-booking essentially makes strides upon the customary methodology.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

2) *Paper Name:-A* **Heuristics for optimizing matrix-based erasure codes for fault-tolerantstorage systems.**

Abstract:-
This paper develops an algorithm for finding the optimum solution and analyzes the performance of two known heuristics on a set of encoding matrices. In this Paper, we referred the solution, XOR operations required during the encoding process.We referred the solution, XOR operations required during the encoding process.

3) *Paper Name:-* **Redundant data composition of peers in p2p streaming systems using Cauchy Reed-Solomon codes.**

Abstract:-
This Paper applies Cauchy Reed Solomon codes to P2P streaming systems which is different from classical Reed Solomon Codes. Due to Churn in these systems, streaming data composition is vital for system management.. In this paper, we referred the solution, In Cauchy RS codes, they use Cauchy matrices instead of Vandomonde matrices and convert GF multiplications to XORs.

4) *Paper Name:-***T-Code: 3-Erasure Longest Lowest-Density MDS Codes.**

Abstract:-
CRS coding involves only XOR operations, the number of XORs required by a CRS code directly upon the performance of encoding or decoding. Cache behavior and device latency, reducing XORs is a reliable and effective way to improve the performance.

5) *Paper Name:-* **Rethinking Erasure Codes for Cloud File Systems: Minimizing I/O for Recovery and Degraded Reads.**

Abstract:-
Cloud record frameworks are transitioning from replication to deletion codes. This procedure has uncovered new measurements on which to assess the execution of various coding plans: the measure of information utilized as a part of recuperation and when performing debased peruses.Show that the distinctions enhance I/O execution practically speaking for the huge square sizes utilized as a part of cloud document frameworks. A few cloud frameworks have embraced Reed Solomon (RS) codes, due to their all inclusive statement also, their capacity to endure bigger quantities of disappointments.

III. SYSTEM ARCHITECTURE AND OVERVIEW

A.EXISTING SYSTEM APPROACH

The Design of CaCo not proves efficacious in the improvement of performance, but also has good scalability. First, CaCo incorporates all existing matrix and schedule heuristics so that it is able to discovery a comparable Solution. Compared to the enumeration algorithm, CaCo has much lower complexity. Second , if there are new heuristics for generating Cauchy matrices or Scheduling to be derived in the future, we can easily add them to CaCo. If some other matrices are later found better for scheduling

Disadvantages :-

- Very General And Fail to Consider individual circumstances
- Potentially misleading results. The ROT may be used for important decisions that should be submitted for more complete analysis.

B. PROPOSED SYSTEM APPROACH

In Proposed System CaCo, a new approach that incorporates all existing matrix and schedule heuristics, and thus is able to identify an optimal coding scheme within the capability of the current state of the art for a given

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

redundancy configuration. The selection process of CaCo has an acceptable complexity and can be accelerated by parallel computing. It should also be noticed that the selection process is once for all. The experimental results demonstrate that CaCo outperforms the "Hadoop-EC" approach by 26.68-40.18% in encoding time and by 38.4-52.83% in decoding time simultaneously.

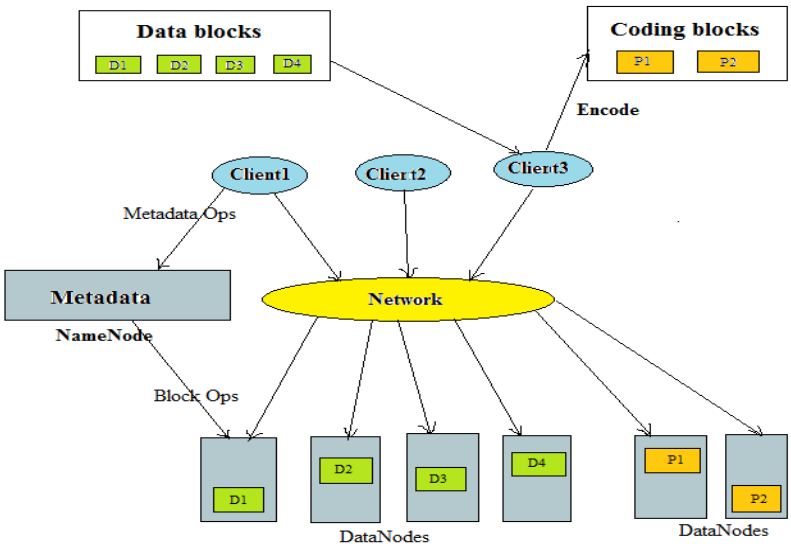

Fig : System Architecture

Advantages:-

- Cloud storage systems always use different redundancy configurations
- There is a large gap in the number of XOR operations, and no single combination performs best for all redundancy configuration

IV. CONCLUSION

Distributed storage frameworks dependably utilize distinctive repetition setups (i.e., k; m; w), contingent upon the coveted equalization amongst execution and adaptation to internal failure. For various mixes of grids and calendars, there is an expansive crevice in the quantity of XOR operations, and no single mix performs best for all excess arrangements. In this paper, we propose CaCo, another methodology that fuses all current framework and timetable heuristics also, consequently can recognize an ideal coding plan inside the ability of the present best in class for a given excess arrangement.

The determination procedure of CaCo has a satisfactory multifaceted nature and can be quickened by parallel processing. It ought to likewise be seen that the choice procedure is once for all. The trial results exhibit that CaCo outflanks the "Hadoop-EC" approach by 26.68-40.18 percent in encoding time and by 38.4-52.83 percent in unraveling time at the same time. At long last, we promptly recognize that diminishing XORs is by all account not the only approach to enhance the execution of a deletion code. Other code properties, similar to the measure of information required for recuperation and corrupted peruses may constrain execution more than the CPU overhead.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 4, Issue 12, December 2016

ACKNOWLEDGEMENTS

I am profoundly grateful to Prof.Gujar A.D.for his expert guidance and continuous encouragement throughout this work. I would like to express deepest appreciation towards Dr.Bilgi D.S., Principal BSCOER, Pune and Prof. PokaleN.B., HOD Computer Department, whose invaluable guidance supported me in completing this work.At last I must express my sincere heartfelt gratitude to Prof. Kodmelwar M.K.

REFERENCES

[1] Kevin M. Greenan ,Xiaozhou Li , "Flat XOR-based erasure codes in storage systems:Constructions, efficient recovery, and tradeoffs", ParaScale ,HP Labs IEEE 2010.
[2] Sheng Lin, Gang Wang, *Member, IEEE*, Douglas S. Stones, Xiaoguang Liu and Jing Liu, "T-Code: 3-Erasure Longest Lowest-Density MDS Codes", VOL. 28, NO. 2, FEBRUARY 2010.
[3] JianqiangLuo, LihaoXu, James S. Plank,"An Efficient XOR-Scheduling Algorithm for Erasure Codes Encoding", Dependable Syst. Netw., 2009, pp. 504–513.
[4] James Lee Hafner, VeeraDeenadhayalan, and KK Rao," Matrix Methods for Lost Data Reconstruction in Erasure Codes", IBM Almaden Research Center.
[5] Osama Khan, Randal Burns," Rethinking Erasure Codes for Cloud File Systems: Minimizing I/O for Recovery and Degraded Reads", Cheng Huang Microsoft Research.
[6] B. Calder, J. Wang, A. Ogus, N. Nilakantan, A. Skjolsvold, S.McKelvie, Y. Xu, S. Srivastav, J. Wu, H. Simitci, J. Haridas, C.Uddaraju, H. Khatri, A. Edwards, V. Bedekar, S. Mainali, R.Abbasi, A. Agarwal, M. F. u. Haq, M. I. u. Haq, D. Bhardwaj, S.Dayanand, A. Adusumilli, M. McNett, S. Sankaran, K.Manivannan, and L. Rigas, "Windows Azure storage: A highly available cloud storage service with strong consistency", inProc. 23rd ACMSymp. Oper. Syst. Principles, New York, NY,USA, 2011, pp. 143–157.
[7] J. Edmonds, "Paths, trees, and flowers," Can. J. Math., vol. 17, no.3, pp. 449–467, 1965.
[8] G. Xu, F. Wang, H. Zhang, and J. Li, "Redundant data composition of peers in p2p streaming systems using Cauchy Reed-Solomon codes," in Proc. 6th Int. Conf. Fuzzy Syst. Knowl. Discovery-Volume 2, Piscataway, NJ, USA, 2009, pp. 499–503.
[9] J. S. Plank, C. D. Schuman, and B. D. Robison, "Heuristics for optimizing matrix-based erasure codes for fault-tolerant storage systems," in Proc. 42Nd Annu. IEEE/IFIP Int. Conf. Dependable Syst. Netw, Washington, DC, USA, 2012, pp. 1–12.
[10] Guoxin Liu and HaiyingShen," Minimum-cost Cloud Storage Service Across Multiple Cloud Providers", Clemson, SC 29631, USA.