TELKOMNIKA Indonesian Journal of Electrical Engineering
Vol.12, No.5, May 2014, pp. 4030 ~ 4037
DOI: http://dx.doi.org/10.11591/telkomnika.v12i5.5100

Research of Probability Symmetric Allocation Storage in Distributed Storage System

Shengrong Lu*[1], Hanming Li[2]
School of Mathematics and Computer Science, Longyan University,
Dongxiao Rd. N, Xinluo DistrictLongyan, Fujian, China, 364012, +86-0597-2793778
*Corresponding author, e-mail: strun@live.cn[1], lhm399@qq.com[2]

Abstract

The goal of optimal allocation is to increase the stored data availability subject to minimize the storage budget. The symmetric allocation based on the network coding is proved to be optimal without considering the nodes availability in distributed storage system. Because of network conditions and node inherent property, each node has different availability. This paper focuses on the optimization distributed data storage problem with nodes availability. Based on probability model of storage system, we re-define the symmetric allocation as the probability symmetric allocation, and proposed probability symmetric allocation model and strategy which are proved to be optimal in the general condition based on SVM. Comparing to the symmetric allocation proposed by Leong D. et al., The proposed probability symmetric allocation scheme improves the data availability, and is more practical method for distributed storage system.

Keywords: *distributed storage allocation, network coding, probability symmetric*

1. Introduction

The environment in the network, sensor network, data center network based on the data center, data distributed storage security and efficiency is the focus of current research. Data backup method because the traditional cost more; data synchronization difficulties are not suitable for distributed network storage system with high reliability. Network coding [1] data storage to provide distributed network storage of high reliability, low cost based, widespread concern and research. In the distributed data storage in [2] based on network coding, original data is extended code into a plurality of encoded data blocks, and the coded data block allocated to different data storage nodes. When users access data, use the data receive node from the data storage node obtaining encoding data blocks corresponding to the [3], restore the original data by calculating. The use of coded data block allocation strategy is the key problem in distributed data storage based on network coding.

In order to solve the allocation, coded data block of Leong D., A. G. Dimakis proposed the uniform code block allocation model of [4], the encoding of the data block amount assigned to each storage node, and the model is proved to be optimal, which ensures the availability of data, data allocation overhead is minimal. In further study, Leong D., A. G. Dimakis et al. set the same failure rate for data storage node, namely the data receive node with the same probability of access data storage node to retrieve a data block. At this time, uniform allocation strategy is still an optimal allocation strategy. Because the storage node itself has a certain failure rate, storage nodes of different types, different time and the failure rate of different use of the environment. In considering the node availability, Leong D., uniform code block allocation model proposed by Dimakis A. G. for lack of considering the availability of storage nodes, or if the same node failure rate, rather than the optimal mode.

In reality, inherent properties of the storage node network environment [11] and the node itself determines the data receiver access storage nodes have a certain probability of failure. For example, network communication link failure, the node itself online storage duration, hacker attack causes the data receiver cannot access data storage node. At this time, data storage nodes with a certain probability are the data receiver access node availability, i.e...

This paper focused on the consideration of distributed data storage optimization storage node availability in case of. Through the method of probability theory, establish the storage node

Received November 2, 2013; Revised December 27, 2013; Accepted January 16, 2014

probability distribution based on data model. Based on the redefined uniform distribution model, the probability of network coding based on uniform distribution model, presents data storage probability distribution strategies and methods, and demonstrates the proposed method is optimal. Compared with Leong D., A. G. Dimakis proposed a uniform distribution model, model and the method proposed in this paper considers the node failure rate, improve the availability of data storage, distributed storage system accords with the actual.

2. Research Method

2.1. Related Work

In distributed storage systems, network coding is a common coding mode, in order to improve the data availability and reliability. The current academic research and discuss how to utilize the data redundancy, to improve the reliability of data. Build redundant data including simple backup, erasure coding and network coding method. Compared to the simple backup mode, in the same storage overhead, erasure codes can provide higher data reliability [5]. At the same time, network coding is applied to distributed storage, can balance the storage space and bandwidth. Distributed data storage based on network coding, is a potential way.

With different mechanism, for a simple backup, data source only needs to be fully replicated original data stored in the data storage node. For the distributed storage based on network coding, data storage nodes requires a data block received coded data, storage space, data bandwidth usage are better than the simple backup [3]. Based on the MDS (n, K) distributed data storage method coding can effectively reduce redundant data, balance the storage space and network bandwidth, can effectively reduce the data distribution, storage and access the use of storage space and network bandwidth, improve the reliability of [6] data.

Distributed storage method based on network coding has received widespread attention in academic circles, the [1-3] distributed data encoding and storage method of massive. But the distribution of code blocks of data research, still less. The traditional method is to each data block code stored in each storage node, in fact is the assignment of [4, 7] data storage method of equivalent.

In fact, in distributed data storage, based on availability, network storage node topology environment can improve the efficiency and usability of the data storage block allocation code data. The research work in this area is still relatively small. The main work concentrates on two aspects: 1) The data distribution method based on efficiency; 2) Data allocation method based on usability.

In the data distribution method based on efficiency, MSR (minimum-storage regenerating, minimum memory regeneration code), is to reduce the data storage space, improve the repairing efficiency data coding modes of distribution as the target. MBR (minimum-bandwidth regenerating, minimum bandwidth regenerating code) is considered from the angle of channel bandwidth, data coding to reduce the link bandwidth allocation strategy for target, see [8, 9]. Tree type data allocation method based on the [10] Jun proposed by Li et al., this paper put forward a data allocation strategies from the perspective of bandwidth, in order to improve the repair efficiency of data.

In the data distribution method based on availability, Leong D. and Dimakis A.G. gives the definition of uniform distribution, and proposes the [4, 7] storage strategy in distributed storage system with uniform distribution. Uniform distribution strategy based on the ideal model, without considering the storage space, storage node failure probability, the number of each storage node storage equal data. In the ideal model, the author proves that the uniform distribution is optimal. Data receiving node access data storage nodes for the data block to restore the original data is completely successful. In considering the working node to its efficiency, Leong D. proposed the data distribution method of the probability of [7, 12]. This method is that the probability of failure for all storage nodes is the same; the essence of this method with the method of literature [4] is in fact consistent. The author proves that the method is only the premise of the same node failure probability is the optimal.

The advantages of the two methods is the storage allocation method is simple; the main disadvantage is the lack of data distribution node to consider its own characteristics, network environment, does not meet the needs of practical applications, in these methods, assuming ideal conditions.

In reality, the individual storage node determines the data visitors cannot access data storage nodes, all with the same probability, failure probability and the present study did not pay attention to the data storage node. Data storage node failure probability of different directly affects the reliability of recovering the original data and the probability of system, uniform distribution strategy is no longer optimal allocation strategy. System reliability is an important goal of distributed storage system, is an important problem of distributed storage system. Method of allocating focusing coding data from general block, considering the failure probability of storage nodes in the data distribution, according to the storage node failure probability of different storage node data distribution method.

2.2. Probability Distribution Model Based on Network Coding

Distributed storage systems, storage nodes are different in different failure probability. Uniform data distribution method of the traditional is no longer optimal. Data distribution method is more feasible to assign a different number of data as the data storage node according to the node availability of different. Based on considering the storage node availability, the probability of network coding based on uniform distribution model.

2.2.1. Establish the Hypothesis and Model

Distributed network storage system is studied in this paper by the data source node, M data storage nodes and data receiving nodes.

The use of MDS (n, k) [2, 3] encoding, encoding original data to be coded data block. May be the original data unit size, which is divided into k blocks, $F_1,\dots,F_k$ use become n block coding matrix code, $B_1,\dots,B_n$ These blocks, $F_1,\dots,F_k$ the linear combination of The coefficient vector Bi is $(a_{i1},\dots,a_{ik})^T$, so we can get:

$$\begin{pmatrix} a_{11} & L & a_{1k} \\ M & O & M \\ a_{n1} & L & a_{nk} \end{pmatrix} \times \begin{pmatrix} F_1 \\ M \\ F_k \end{pmatrix} = \begin{pmatrix} B_1 \\ M \\ B_n \end{pmatrix}$$

Data allocation assigns the *B*i to data storage node is different, the code allocation scheme as shown in Figure 1.

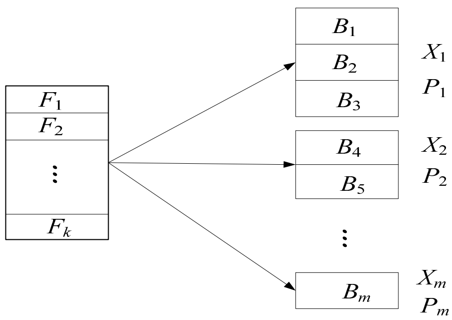

Figure 1. Data Allocation Code Based on the MDS (n, k)

The source data object D for stored data. Through MDS (n, k) data of D coding method, coding and data block allocated to each storage node. Assigned to each storage node data is represented as x_i, $1 \le i \le m$ Availability of storage node is the probability of nodes are available, and the node by manufacturers, technology, operation time, operation environment and so on, can be obtained by a node sampling. The availability of each storage node is different, but in a relatively short period of time is fixed. The probability of failure data storage node x_i is F_{xi}, the availability probability is p_i=1-F_{xi}

2.2.2. Uniform Distribution of Probability Model

This paper proposes to use for data distribution node availability strategy based on the proposed model, probability distribution. Probability distribution model is a generalization of the traditional uniform distribution, the distributed storage system considering node failure rate. The allocation model as follows:

$$\begin{aligned} X &= \{x_i\}_{i=1}^{m} \\ &= \{x_1, \ldots, x_m\} \\ &= \{f(p_1), \ldots, f(p_m)\} \end{aligned} \quad (1)$$

$$\sum_{i=1}^{m} x_i = \sum_{i=1}^{m} f(p_i) = T \quad (2)$$

$$f(p_i) > f(p_j), \forall p_i, p_j \in [0,1], p_i > p_j \quad (3)$$

Use $X = \{x_i\}_{i=1}^{m}$ to represent the data distribution, distribution function f (pi) is assigned to the i encoded data node, $f(p_i)$ is defined as the interval [0, 1] one-way increasing function, change the node p_i availability. *T is* storage space for data storage node.

The use of data distribution model is proposed in this paper, a data block for high-availability node allocation more, make the data receiver can block access to higher reliability, with a higher probability of recovering the original data object.

2.2.3. Data Recovery Method

Data receiving node can successfully recover the original data, the access to the data storage node acquires data is not less than the original data. The original data unit size, namely the data receive node access data quantity of not less than 1.

Assuming that the data receive node successfully restored the original data need to access r data storage node, *r* subset, use $|r|$to represent the R subset of the number of elements, namely the storage node number. S was defined as the successful recovery of data events.

$$P(S) = \sum_{|r|=1}^{m} P(|r|) \cdot I[\sum_{i \in r} x_i \geq 1]$$

Among them,

$$P(|r|) = \prod_{i \in r} p_i \prod_{j \in M \backslash r} (1 - p_j)$$

Is true I [G] =*1*, G is false I [G] =*0*.

Distributed storage system of the *m* data storage node, composed of a plurality of *r*subset, $M \backslash r$ represents a r subset of a M set. The set was defined r subset of all the successful repair of the set S_r, said to be able to set a subset of successful recovery, $|S_r|$ is the number of elements in the collection of successful repair that is composed of *m* data nodes *r* the number of subsets.

3. Results and Analysis

This section based on the storage node availability, analysis using probability distribution model, probabilistic data receive node successful recovery, data availability. At the same time, this article in the presence of node failures, data availability, uniform distribution and probability distribution evaluation with probability form.

3.1. Data Availability

For distributed memory systems in the presence of node availability, data storage node exists failure rate, so the data receive node successfully recover the probability of original data objects is not the only value. For the distributed data storage, the reliability of the system is a distributed system is the most important goal.

For the distributed storage system, the use of probability distribution model, the amount of data the *m* data storage nodes to store is different, according to the node availability changes, therefore, a number of each subset of elements of success is uncertain, $|\boldsymbol{r}| \in [1, m]$, as long as the subset of the nodes on the amount of data of $\sum_{i \in r} x_i \geq 1$ can meet. In distributed storage systems, data availability in each subset of successful recovery, data availability of each *r* subset of *r* concentration probability of all nodes is available.

Theorem 1: The existence of storage node failure rate of distributed storage system, data availability for (can recover data probability):

$$\sum_{|r|=1}^{m} \min\left(\frac{|r|T}{m}, 1\right) P(|r|)$$

Proof: The data receive node can be successfully restored the original data, the data size must meet the following conditions:

$$\sum_{i \in r} x_i = \sum_{i \in r} f(p_i) \geq 1$$

S_r said to set successful subset, which consists of a subset of the set of all ***r***. $|S_r|$ is the number of the elements of success in the collection. *S* was defined as the successful recovery of data events, A_i is exactly *i* nodes are access to things.

$$\begin{aligned} P(S) &= \sum_{|r|=1}^{m} P(S \mid A_i) \cdot P(A_i) \\ &= \sum_{|r|=1}^{m} \frac{|S_r|}{\binom{m}{|r|}} \cdot \prod_{i \in r} p_i \prod_{j \in M \setminus r} (1 - p_j) \\ &= \sum_{|r|=1}^{m} \frac{|S_r|}{\binom{m}{|r|}} \cdot P(|r|) \end{aligned}$$

For each subset of ***r***, can be represented by inequalities:

$$\sum_{i \in r} x_i = \underbrace{x_{t1} + x_{t2} + \ldots + x_{tr}}_{r\ nodes}$$

All the elements in the S_r are a successful recovery of a subset of S_r, can be expressed by the following inequalities:

$$\begin{cases} x_{t1} + x_{t2} + \ldots + x_{tr} \geq 1 \\ \vdots \\ x_{S_r 1} + x_{S_r 2} + \ldots + x_{S_r r} \geq 1 \end{cases}$$

With inequalities, $a_1 x_1 + \ldots + a_m x_m \geq |S_r|$, because each storage node belongs to a subset of R, different $0 \leq a_i \leq \binom{m-1}{|r|-1}$, and on the type:

$$\begin{aligned}|S_r| &\le a_1x_1+...+a_mx_m\\ &\le \binom{m-1}{|r|-1}\sum_{i=1}^{m}x_i\\ &\le \binom{m-1}{|r|-1}T\end{aligned}$$

Because $|S_r|\le\binom{m}{|r|}$, is $|S_r|\le \min\left(\binom{m-1}{|r|-1}T,\binom{m}{|r|}\right)$

Therefore, the presence of reliability on distributed storage system node availability for:

$$P(S)=\sum_{|r|=1}^{m}\min\left(\frac{|r|T}{m},1\right)P(|r|)$$

System reliability in distributed storage systems, namely, data receiving node to the success probability of recovery is determined by a random variable p_i m.

3.2. Probability Distribution Model and Distribution Model Comparison

Distributed storage system, data distribution strategy is the main factor of affecting the distributed storage system. Uniform allocation strategy for Dimakis A.G. is proposed, the optimal allocation strategy in the ideal model. For the distributed storage system storage node availability, the probability distribution model is proposed in this paper, based on the storage node failure rate, the use of distribution function for the allocation of storage node does not equal amounts of data. In the storage node is available, uniform probability model is equivalent to the uniform distribution model in literature [4], in all the storage node availability equal, uniform distribution model probability model is equivalent to the document [7]. The uniform probability model is more suitable to the universal, uniform distribution case model is uniform distribution model probability.

Theorem 2: the presence of distributed storage system node availability, uniform distribution of uniform distribution model definition model is superior to [4] and [7] in the probability of literature.

The proof of theorem 2 before, first need a lemma.

Lemma 1: The minimum number of nodes with the least number of node probability distribution models of successful recovery using less than uniform distribution defined [4] and [7] use a successful recovery.

Proof: We use reduction ad absurdum proof.

The minimum number of nodes to evenly distribute the successful recovery of the assumed probability using r_1 use the appropriate distribution function, makes the stored data storage node is $f(p_1) > f(p_2) > \ldots > f(p_{r1})$.

Minimum number of nodes to evenly distribute the successful recovery of the use of a r_2, memory data storage node weight is $\frac{T}{m}$, $r_1>r_2$。The number of successful recovery nodes with the least at the receiving node, data access to the amount of 1:

$$f(p_1)+f(p_2)+...+f(p_{r_1})=r_2\cdot\frac{T}{m}=1$$

$$\Rightarrow f(p_1)-\frac{T}{m}+f(p_2)-\frac{T}{m}+...+\sum_{i=r_1-r_2+1}^{r_1}f(p_i)=0$$

Because of the $f(p_i)>\frac{T}{m}$, and $\sum_{i=r_1-r_2+1}^{r_1}f(p_i)>0$, so cannot be equal to zero, and the assumption of contradiction. There are $r_1<r_2$, minimum number of nodes with the least number

of nodes is uniform distribution model successfully restored using probability less than successful recovery using a uniform distribution.

Now the proof of theorem 2

Proof: Uniform distribution model, all data storage nodes to store the same amount of data, without considering the empty nodes, the amount of data of all storage nodes to store is $\frac{T}{m}$, an ideal situation successfully restored the number of nodes requires at least $\left\lceil 1/\frac{T}{m} \right\rceil = \left\lceil \frac{m}{T} \right\rceil$, at this time, $\frac{|r|T}{m} > 1$. Existing distributed storage system node availability, uniform distribution model of data availability for $\sum_{|r|=\left\lceil \frac{m}{T} \right\rceil}^{m} P(|r|)$.

The minimum number of nodes with probability distribution model successfully restored to use for *t*, $t < \left\lceil \frac{m}{T} \right\rceil$ the availability of data, and a uniform distribution of model data availability is the difference:

$$\sum_{|r|=t}^{m} \min\left(\frac{|r|T}{m}, 1\right) P(|r|) - \sum_{|r|=\left\lceil \frac{m}{T} \right\rceil}^{m} P(|r|)$$

$$= \begin{cases} \sum_{|r|=t}^{\left\lceil \frac{m}{T} \right\rceil - 1} P(|r|) & \frac{|r|T}{m} > 1 \\ \sum_{|r|=t}^{m} \frac{|r|T}{m} P(|r|) - \sum_{|r|=\left\lceil \frac{m}{T} \right\rceil}^{m} P(|r|) & \frac{|r|T}{m} < 1 \end{cases}$$

When $\frac{|r|T}{m} > 1$, $\sum_{|r|=t}^{\left\lceil \frac{m}{T} \right\rceil - 1} P(|r|) > 0$,

Probability distribution, data availability is greater than the uniform distribution of data availability.

When $\frac{|r|T}{m} < 1$, a uniform distribution of the model failed to restore the original data.

To sum up, the probability distribution model of data availability is greater than the uniform distribution of data availability, there is a distributed storage system node availability, probability distribution strategy is better than the uniform allocation strategy.

4. Conclusion

Optimization of distributed data storage is the goal of ensuring safe storage data recovery, to improve the reliability of data storage system. Without considering the storage node availability, network coding data based on uniform distribution is shown to be optimal. But due to node failures and other factors, the storage node has different failure rate, based on the storage node probability model, defines the uniform distribution probability model, made the storage data probability distribution strategies and methods, and demonstrate the proposed method is optimal. Compared with Leong D., A. G. Dimakis proposed a uniform distribution model, model and the method proposed in this paper considers the availability of nodes, improves the effectiveness of data storage system, has more realistic significance.

Acknowledgements

The work is supported by the B class Natural Science project of Fujian province Education Department of China (JB12208).

References

[1] Dimakis Ag, Prabhakaran V, Ramchandran K. Decentralized Erasure Codes for Distributed Networked Storage. *IEEE Transactions on Information Theory*. 2006; 52(6): 2809-2816.

[2] Dimakis Ag, Ramchandran K, Wu Y. *A Survey on Network Codes for Distributed Storage*. Proceedings of the IEEE. 2011; 99(3): 476-489.
[3] Dimakis Ag, Godfrey Pb, Wu Y. Network Coding for Distributed Storage Systems. *IEEE Transactions on Information Theory*. 2010; 56(9): 4539-4551.
[4] Leong D, Dimakis Ag, Ho T. *Distributed storage allocation problems.* Network Coding, Theory, and Applications. Workshop on NetCod '09. Lausanne. 2009.
[5] Acedanski S, Deb S, Medard M. *How good is random linear coding based distributed networked storage*. Proceedings of the 1st. Workshop on Network Coding. Riva del Garda. 2005.
[6] Cadambe Vr, Jafar Sa, Maleki H. *Minimum Repair Bandwidth for Exact Regeneration in Distributed Storage*. Wireless Network Coding Conference (WiNC), IEEE. Bostn, MA, USA. 2010.
[7] Leong D Dimakis Ag, Ho T. *Distributed Storage Allocation for High Reliability*. Communications (ICC). IEEE International Conference on. Cape Town, South Africa. 2010.
[8] Rashmi KV, Shah NB, Kumar PV. *Optimal Exact-Regenerating Codes for Distributed Storage at the MSR and MBR Points via a Product-Matrix Construction*. Information Theory. IEEE Transactions on, 2011; 8(57): 5227-5239.
[9] Rashmi KV, Shah NB, Kumar PV. *Explicit and Optimal Exact-Regenerating Codes for the Minimum-Bandwidth Point in Distributed Storage*. Information Theory Proceedings (ISIT). IEEE International Symposium on. Austin, TX. 2010.
[10] Eong D, Dimakis Ag, Ho T. *Symmetric Allocations for Distributed Storage*. Global Telecommunications Conference (GLOBECOM 2010). Mia GLOBECOM 2010, Miami, FL, USA 2010 Mohan N, Undeland TM, Robbins WP. Power Electronics. New York: John Wiley & Sons. 2005.
[11] Xiaoying Wang. Meta-service Design and Implementation for ContentDelivery Network Cloud Environments. *TELKOMNIKA Indonesia Journal of Electrical Engineering*. 2012; 10(7): 1833-1842.
[12] Ren Shuxia, Zhao Zheng, Zou Xiaojian. An Indeterminacy Temporal Data Model based on Probability. *TELKOMNIKA Indonesia Journal of Electrical Engineering.* 2013; 11(11): 6686-6692.