Received: 1 April 2021 | Revised: 25 July 2021 | Accepted: 12 September 2021

DOI: 10.1049/cmu2.12289

ORIGINAL RESEARCH PAPER

Global repair bandwidth cost optimization of generalized regenerating codes in clustered distributed storage systems

Shushi Gu[1,2] | Fugang Wang[1] | Qinyu Zhang[1,2] | Tao Huang[3] | Wei Xiang[2,4]

[1] School of Electronic and Information Engineering, Harbin Institute of Technology (Shenzhen), Shenzhen, China

[2] Network and Communication Centre, Peng Cheng Laboratory, Shenzhen, China

[3] College of Science and Engineering, James Cook University, Cairns, QLD, Australia

[4] Cisco-La Trobe Centre for AI and IoT, La Trobe University, Melbourne, VIC, Australia

Correspondence
Shushi Gu, Harbin Institute of Technology (Shenzhen), Shenzhen 518055, China, Peng Cheng Laboratory, Shenzhen 518052, China; Peng Cheng Laboratory, Shenzhen 518052, China.
Email: gushushi@hit.edu.cn

This paper was presented in part at the 2019 11th International Conference on Wireless Communications and Signal Processing.

Funding information
National Natural Sciences Foundation of China, Grant/Award Numbers: 62027802, 61831008; Guangdong Science and Technology Planning Project, Grant/Award Numbers: 2018B030322004, 2021A1515011572; Shenzhen Natural Science Fund, Grant/Award Number: JCYJ20200109112822953; Shenzhen Natural Science Fund - Stable Support Plan Program, Grant/Award Number: GXWD20201230155427003-20200824081029001

Abstract

In clustered distributed storage systems (CDSSs), one of the main design goals is minimizing the transmission cost during the failed storage nodes repairing. Generalized regenerating codes (GRCs) are proposed to balance the intra-cluster repair bandwidth and the inter-cluster repair bandwidth for guaranteeing data availability. The trade-off performance of GRCs illustrates that, it can reduce storage overhead and inter-cluster repair bandwidths simultaneously. However, in practical big data storage scenarios, GRCs cannot give an effective solution to handle the heterogeneity of bandwidth costs among different clusters for node failures recovery. This paper proposes an asymmetric bandwidth allocation strategy (ABAS) of GRCs for the inter-cluster repair in heterogeneous CDSSs. Furthermore, an upper bound of the achievable capacity of ABAS is derived based on the information flow graph (IFG), and the constraints of storage capacity and intra-cluster repair bandwidth are also elaborated. Then, a metric termed global repair bandwidth cost (GRBC), which can be minimized regarding of the inter-cluster repair bandwidths by solving a linear programming problem, is defined. The numerical results demonstrate that, maintaining the same data availability and storage overhead, the proposed ABAS of GRCs can effectively reduce the GRBC compared to the traditional symmetric bandwidth allocation schemes.

1 | INTRODUCTION

With the demands of massive data storage, large-scale storage systems are often built on hundreds or even thousands of storage servers around the world and composed of multiple racks or clusters, which is also called Clustered Distributed Storage Systems (CDSSs), for example Amazon Dynamo [1] and Microsoft Azure [2]. One of main design goals of CDSSs is ensuring the file availability against nodes failures. To guarantee data reliability, pre-storing additional data redundancy is an advantageous approach in practical systems [3], [4], that is replication and erasure coding. Compared to replication, erasure coding has a better storage efficiency, while a data collector can reconstruct the lost data file by connecting arbitrary multiple storages nodes in file retrieval process. In order to ensure storage efficiency, traditional erasure coding expends large repair bandwidth. For example, Reed-Solomon (RS) codes [5], a typical maximum distance separable (MDS) code, has to download the chunks whose size is several times that of the lost data. To reduce the repair bandwidth, regenerating codes (RCs) are introduced by Dimakis et al. [6], which can effectively reduce the repair bandwidth with the same storage-efficiency of MDS codes.

As repairing the failure node in CDSSs, system will occupy two kinds of bandwidth resources. First, in the host cluster (in which node failed) and the remote help clusters, object nodes will download the repair data from alive nodes that occupies intra-cluster bandwidth. Second, the object nodes in the remote help clusters send repair data to the newcomer that occupies inter-cluster bandwidth. Generally speaking, the intra-cluster bandwidth is considered as fully available and cheap, while inter-cluster bandwidth is considered as scarce and expensive [7]. For the same requesting frequency, the available inter-cluster bandwidth is about 1/5 to 1/20 of the intra-cluster bandwidth. In some extreme cases, this proportion can even deteriorate to 1/240 [8]. In order to distinguish between intra-cluster repair bandwidth and inter-cluster repair bandwidth, and to reduce the inter-cluster repair bandwidth, generalized regenerating codes (GRCs), as an extension of regenerating codes (RCs), are proposed by Prakash et al. [9], which can reduce inter-cluster bandwidth by increasing intra-cluster bandwidth. Since inter-cluster bandwidth is much more expensive than intra-cluster bandwidth, the entire cost of node repair can be reduced effectively.

There are two problems needed to be solved for the current node repair in CDSSs. (i): All coding strategies (including GRCs) and system models only distinguish the bandwidth differences between intra-cluster and inter-cluster, and the inter-cluster repair adopts a symmetric repair model (newcomer download the same size of data from different remote helper clusters). For example, HDFS of Hadoop adopts the erasure coding like RSs and LRCs, and the both strategies are symmetric for inter-cluster repair. (ii): Most studies only focus on the optimal of inter-cluster bandwidth, simply treats bandwidth and cost equally. However, in the practical heterogeneous CDSSs, due to the different communication distances, stabilities of bandwidth and prices of transmission [10], transmitting the same size of data spends widely differently on different inter-cluster links.

It is noted the difference of exact repair and functional repair: Under exact repair, the content of the repaired node is identical to that of the failure node. While under functional repair, repaired node is not necessarily store the same content as the failure node, but has the same function as the failure node, which means that the repair content permits data collection and repair of additional failed nodes. This paper focuses on the functional repair of GRCs.

On the basis of above motivations, our main contributions are summarized as follows:

- We propose an asymmetric bandwidth allocation strategy (ABAS), and derive its achievable upper bound of file storage capacity by information flow graph (IFG). Moreover, for achieving the upper bound, we elaborate the constraints of intra-cluster bandwidths in the host cluster and remote help clusters, respectively, and obtain their lower bounds.
- We give the definition of global repair bandwidth cost (GRBC), which quantifies the performance of node repair. Then we derive GRBCs of ABAS and symmetric repair strategy for GRCs, and get the formula expressions between GRBC and other parameters. By solving a linear programming problem, we minimize GRBC and get the optimal parameter settings with intra-cluster bandwidth constraints.
- We provide the numerical results by simulation to compare GRBCs among ABAS of GRCs, symmetric repair of GRCs and RCs. Finally, we prove that ABAS of GRCs is an effective solution for heterogeneous storage system in reducing repair cost of node failures.

The rest of this paper is organized as follows. Section 2 introduces the related work and some basic knowledge. Section 3 presents the asymmetric bandwidth allocation strategy. Based on the information flow graph, the capacity upper bound of file storage size and the lower bound of intra-cluster bandwidth are derived under functional repair. Section 4 introduces the definition of GRBC, and optimizes GRBC by solving a linear programming problem under different parameters. The numerical comparison are illustrated in Section 5. Section 6 concludes this paper. The key notations used are summarized in Table 1.

2 | RELATED WORK AND PRELIMINARY

2.1 | Related work

For heterogeneous CDSSs, there exits many coding methods and system models. Locally Repairable Codes (LRCs) [11] are a type of non-MDS codes suitable for hierarchical storage system, which are currently used in large-scale distributed storage such as Microsoft Azure. In addition [12], considers the size of the local group, indicating that any node in a local group can be repaired by any other nodes in the same group, and each local group is protected by a MDS code. The capacity of LRCs is increased, but the size of the repair set is still limited to a small range [13]. considers a LRC construction where a node has multiple mutually exclusive repair sets of size, and improves the maximum bit rate to a certain extent [14]. proposed a generic transformation for any MDS codes to achieve optimal repair access for a single-node repair, which achieved optimality in both repair and update for normal MDS codes.

Theoretical studies on RC [15] in a hierarchical network as cluster topology also emerges endlessly [16–18]. distinguish the differences in bandwidth costs between internal and external racks in actual multi-rack storage systems [16]. and [17] allow nodes within a rack to participate in transmitting help data, thereby reducing the repair bandwidth overhead of the rack. In the Double Regenerating Codes (DRCs) proposed by [18], each rack stores multiple coded symbols (not one symbol) of a data block in different storage nodes. To repair a failed node, it must first regenerate newcomer in each rack. A node in each rack collects the encoded data from all nodes in the rack and re-encodes it. The regenerated data is transferred to the failed rack and the content of the failed node is recovered. Pernas et al. in [19] propose a variant system model similar to the two-rack model in [17], in which the storage costs of the nodes in the two racks are different, and the feasibility of minimum bandwidth regeneration code is verified under this model. Calis and

TABLE 1 Key notations

Notations	Descriptions	Notations	Descriptions
n	Number of clusters in a system	m	Number of storage nodes in each cluster
k	Number of clusters for file reconstruction	d	Number of remote helper clusters
ℓ	Number of local helper nodes in host cluster	α	Storage overhead on each node
β_{ij}	Inter-cluster repair bandwidth from j^{th} cluster to i^{th} cluster	C	Repair bandwidth cost
γ	Intra-cluster repair bandwidth in host cluster	ℓ'	Number of local helper nodes in each of remote help clusters
γ'	Intra-cluster repair bandwidth in remote helper cluster	ρ_{ij}	Cost ratio of transmission from j^{th} cluster to i^{th} cluster
B	Size of storage file	ε_i	Number of the out-nodes when FIG cut passed i^{th} cluster which belongs to sink set

Koyluoglu et al. [20] also consider a two-tier storage model, which consists of blocks (similar to clusters) of several storage nodes. Different aggregation methods are used for data collection and node repair. They assume that the entire block where the faulty node is located may be unavailable, and only use the nodes in other blocks for repair, without distinguishing between bandwidth costs within and outside the block [21, 22]. propose rack-aware regenerating code, which can divide the storage system into racks, based on [9], and provide the coding constructions. Shah et al. [23] propose the Flexible Regenerating Codes (FRCs) based on the work in [15]. This scheme allows the data collector to connect any number of nodes to recover the entire file, as long as it can meet the condition that the total amount of data obtained by the new node through other help nodes is greater than or equal to the stored file size. At the same time, the help data obtained by the new node from each help node is only required to be less than the link capacity, and the total help data amount is greater than or equal to the preset parameters. Cooperative Regenerating Codes (CRCs) is proposed by Hu et al. [24], which can effectively reduce the inter-cluster bandwidth by sharing data among newcomers.

Shen [25] proposes cluster-aware scattered repair based on RS codes, which researches the best help blocks that makes inter-cluster bandwidth smallest. [26] and [27] propose the block replication strategy, choosing different data blocks to minimize the inter-rack transmission bandwidth. For the cost issue, Yu et al. [28] and Ernvall et al. [29] consider the heterogeneous distributed storage system. The trade-off between system storage cost and the download cost of node repair data, the capacity and security of the system under this model are both studied [30]. analyzes the trade-off relationship between download cost and repair bandwidth. There are two node clusters with different download costs, the download cost required for the new node to rebuild data depends on its location. Each helper node should find the optimal path, possibly through other intermediate nodes, to reach the new-born node, in order to use all kinds of link capabilities to transmit all the helper node data needed for repair in the shortest possible time. Qu et al. [31] propose the Asymmetric Regenerating Codes (ARCs) to minimize the repair bandwidth of node repair. Its model is similar to flexible regenerative codes. All these repair models are based on the non-heterogeneous system.

At the same time, in order to minimize the download cost of help data during node repair in heterogeneous systems [32], studies the system with different node storage capacities, considers the different cost of downloading data from different nodes, and proves that the approximately uniform repair (quasi-uniform repair) method can achieve the optimal repair bandwidth cost [33–36]. study the problem of minimizing the cost of RCs in a tandem network topology. Among them [33], considers the situation that some node may not be able to transmit directly, and put forward the best repair plan under the multi-hop distributed storage system. Through different repair paths, the RC encoding principle based on precise repair and the optimal problem of node repair are given [34]. discusses the topology design of the distributed storage system, comprehensively considering the differences in storage capacity, storage cost, and data transmission cost between different nodes in the normal distributed storage system, and it proposes a new topology model. By introducing the data transmission cost matrix in the normal distributed system, the best data restoration cost method is obtained [35]. proposes a series network model in which there is only one link between adjacent nodes. When a single node is repaired, only adjacent nodes will transmit data. And by comparing with the general network model structure, it is concluded that no matter where the failed node is, compared with the general network repair model, the series structure can get the best repair cost [36]. compares the performance differences between centralized storage systems and distributed storage systems, and considers the transmission cost as a convex optimization problem. Using the original decomposition and dual decomposition methods, decoupling is a local solution to minimize the transmission cost.

From the above analysis, it can be known that most of the research on RC repair costs in heterogeneous conditions currently focuses on normal storage networks, although it takes into account the differences in link bandwidth costs and network topology.

2.2 | Preliminary

We will present the natural GRCs for CDSSs. The detailed encoding process of GRCs is introduced in [9], we mainly

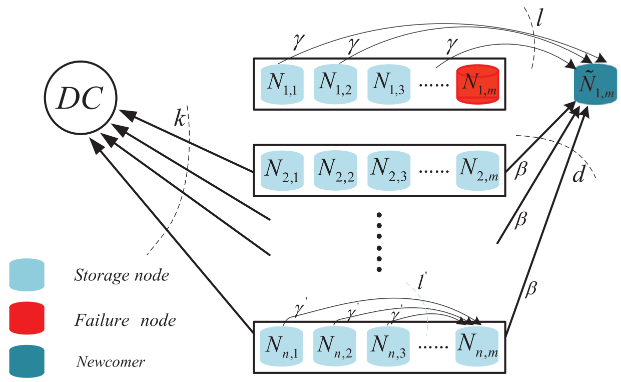


FIGURE 1 System model of node repair and data collection for GRCs in CDSSs. Data collection needs to download the contents of k clusters. Each of l nodes transmits γ symbols in the host cluster, each of d remote clusters transmits β symbols and each of l' nodes transmits γ' symbols in remote help clusters which are needed when repairing a single failure node

introduce and optimize the performance indicators for GRCs under functional repair, with the purpose of reducing the repair cost. The system consists of n clusters, with m nodes in each cluster. The system is fully connected meaning that any two nodes with in a cluster are connected via an intra-cluster link, and any two clusters are connected via an inter-cluster link by a pair of dedicated nodes. A node in a cluster can communicate with another node in another cluster via the corresponding inter-cluster link. A data file with the size of B symbols will be encoded into $nm\alpha$ symbols, and stored in these nm nodes with each node stores α symbols. The symbols are assumed to generated from a finite field $\mathbb{F}_q$ of q elements. The clustered storage system allows data failure by data repair and executes file reconstruction by data collection. For data collection, GRCs satisfy the MDS property in inter-clusters, which means the entire content of any k clusters is sufficient enough for reconstructing the original file data. For data repair, we only focus on the situation of one single failure node in the paper. We describe GRCs by parameters d, β and l, assume that the failure node and its replacement (which is called the newcomer) are located in the same cluster. And we call the cluster where the failure node located as the host cluster. The newcomer downloads β symbols each from any set of d other clusters, which is called remote help clusters and β is assumed to be a function of the $m\alpha$ symbols in the cluster. Every node in the cluster is required to compute these β symbols and then transmit these β symbols to the host cluster. Further, newcomer downloads contents from any set of l other nodes (which are called local helper nodes) in the host cluster. Therefore, the size of symbols $d\beta$ represents the entire inter-cluster repair bandwidth. The repair process is shown in Figure 1. We further parameterize GRCs by parameters $\{(n, k, d), (\alpha, \beta), (m, \ell)\}$.

For more details, intra-cluster bandwidth is needed to download repair data from l local help nodes in the host cluster, and to connect l' local help nodes with each to transmit γ' symbols for computing β in remote help clusters. In this model, the newcomer downloads γ $(\gamma \le \alpha)$ symbols from each of the l local help nodes from the host cluster and β symbols each from d remote help clusters. For a remote help cluster, we assume that the β symbols are just a function of l' and γ'. We make the assumption that any set of l' nodes can be used to compute the β symbols. Further, we limit the amount of data that each of these l' nodes can contribute to at most γ' $(\gamma' \le \alpha)$ symbols. Next, we introduce some performance indicators of GRCs.

2.2.1 | Upper bound of file size B

After giving the parameters of GRCs, the upper bound of file size B under functional repair can be expressed as 1:

$$B \le B^* \triangleq \ell k\alpha + (m-\ell)\sum_{i=0}^{k-1}\min\{\alpha, (d-i)^+\beta\}, \tag{1}$$

where the notation a^+ denotes $\max(a, 0)$, for any integer a.

2.2.2 | Lower bound of intra-cluster bandwidth

To achieve the upper bound of B, intra-cluster bandwidth γ and γ' need to be constrained. We present the lower bounds of intra-cluster bandwidth γ, l' and γ', under the assumption that 1 is achieved with equal sign. For these two parameters γ and γ', when one achieves the lower bound, the other is considered as α. In [9], for the optimal GRCs under function repair, its parameters satisfy as following:

$$\gamma \ge \gamma^* = \alpha - (d-k+1)^+\beta, \tag{2}$$

$$l' = m, \tag{3}$$

and

$$\gamma' \ge \frac{\beta}{m-l}. \tag{4}$$

3 | ASYMMETRIC BANDWIDTH ALLOCATION STRATEGY

The practical CDSSs are composed of multiple clusters distributed among large-scale data centres in different geographical locations. Data transmission across clusters is often affected by link bandwidth, network topology and pricing management. Availabilities and communication qualities of inter-cluster bandwidths are diverse, in other words, inter-cluster bandwidths of CDSSs are heterogeneous. The detailed encoding method of GRCs is shown in [9]. And in this section, we propose an asymmetric bandwidth allocation strategy (ABAS) of GRCs to optimize the repair process in order to reduce repair cost.

3.1 | Repair model of ABAS

As shown in Figure 2, we consider a (n, k, d, m, ℓ) asymmetric bandwidth allocation strategy (ABAS) of GRCs, which consists

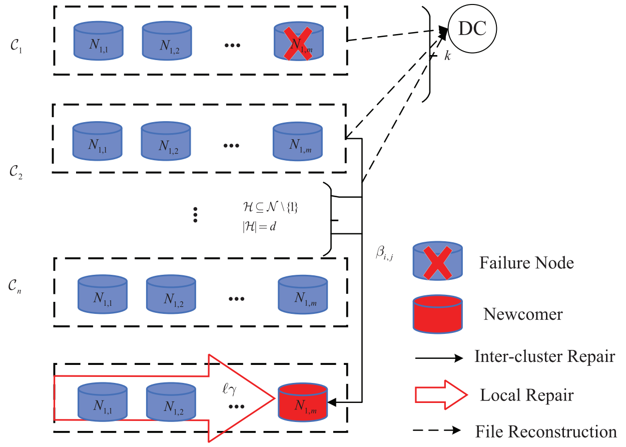


FIGURE 2 ABAS of GRCs. The contents of any k clusters is enough to reconstruct the original file. When repairing a single failure node, each of l nodes transmits γ symbols in the host cluster and each of d remote clusters transmits β_{ij} symbols. In remote help clusters, l' nodes with each transmit γ' symbols to calculate β_{ij}

of n clusters, with each cluster contains m storage nodes. We represent the 1st cluster to n^{th} cluster by $\mathcal{C}_1, \mathcal{C}_2, \dots, \mathcal{C}_n$, and denote j^{th} node in i^{th} cluster by $N_{i,j}$. When a certain node failure happens in $\mathcal{C}_i, i \in [n]$, a newcomer will replace the failed node by downloading help data sized γ from each of ℓ local helper nodes, and $\beta_{i,h}$ inter-cluster helper data from each of remote helper cluster $\mathcal{C}_h, h \in \mathcal{H}_i$, where $\mathcal{H}_i$ represents the set of indexes of remote helper clusters that the host cluster $\mathcal{C}_i$ request. Note that, $\mathcal{H}_i \subseteq [n] \setminus i$ and the cardinality of $\mathcal{H}_i$ is d. Hence, we can easily get that, when repairing a single node in $\mathcal{C}_i$, the inter-cluster repair bandwidth is a d-dimensional vector $\boldsymbol{\beta} = [\beta_{i,1}, \dots, \beta_{i,h}]^{\mathrm{T}}, h \in \mathcal{H}_i$. We permit $\beta_{i,h}$ from $\mathcal{C}_h$ are possibly a function of ℓ' and γ'. Moreover, newcomer acquires γ data from each of other ℓ nodes in the host cluster $\mathcal{C}_i$. For data collection, the whole contents from any k clusters is enough for reconstructing the original file.

3.2 | Information flow graph description

In this section, we describe the repair and reconstruction process in ABAS by a directed acyclic graph $G = (\mathcal{V}, \mathcal{E})$, that is information flow graph (IFG). $\mathcal{V}$ denotes the set of nodes in IFG and $\mathcal{E}$ denotes the set of edges in IFG. The storage capacity of one node is α. Time is divided into stages, and the stages is denoted by non-negatives integers. Upon the failures of a storage node, we will repair it and advance to the next stage. So, from stage $s-1$ to stage s, a newcomer will replace the failed node successfully. In $G = (\mathcal{V}, \mathcal{E})$, an actual node $N_{i,j}$, where $1 \le i \le n, 1 \le j \le m$ is represented by a pair of node $N_{i,j}^{\text{in}}$ and $N_{i,j}^{\text{out}}$, and there exists an edge of capacity α from $N_{i,j}^{\text{in}}$ to $N_{i,j}^{\text{out}}$, denoted as $e(N_{i,j}^{\text{in}} \to N_{i,j}^{\text{out}})$. Also, we assume every cluster has an additional control node responsible for communicating with other clusters, denoted by N_i^{ctrl}, which has an edge with capacity α from each $N_{i,j}^{\text{out}}$). In actual, the control node can be regarded

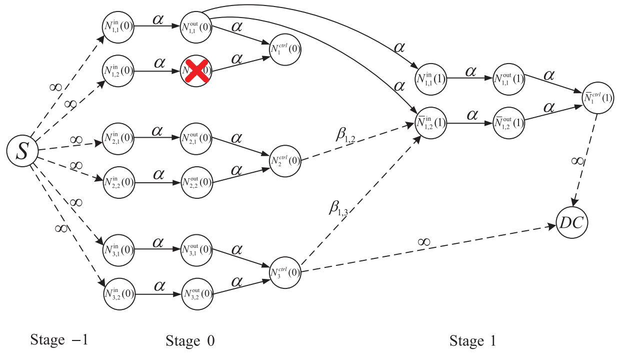


FIGURE 3 An IFG example of ABAS for a ($n = 3, k = 2, d = 2, m = 2, \ell = 1$) GRC

as any node in a cluster. The source node S represents the position of original file encoded into nm nodes, and it connects every initial $N_{i,j}^{\text{in}}$ in an edge with infinity capacity, represented by $e(\mathrm{S} \to N_{i,j}^{\text{in}})$. DC represents the data collector, executing the file reconstruction by connecting control nodes from any k clusters with an infinity capacity edge $e(N_i^{\text{ctrl}} \to \mathrm{DC})$, which means any k clusters' contents can reconstruct the original file.

In the repair process, it is noted that each cluster at any stage only has m available nodes. When an actual node fails (we denote it as $N_{i,j}$), it becomes unavailable. We call the cluster that contains failed node as the host cluster. We replace $N_{i,j}$ by a newcomer, denoted as $\overline{N_{i,j}}$, which connects the remote helper cluster $\mathcal{C}_{h_i}, h_i \in \mathcal{H}_i$ to acquire the remote help data and connect to other ℓ nodes in the cluster it lies to acquire local help data. In the IFG, the remote help data is denoted by β_{i,h_i} and the connection is represented by an edge $e(N_{h_i}^{\text{ctrl}} \to \overline{N_{i,j}})$. Use $\mathcal{L}, \mathcal{L} = \{N_{i,r}^{\text{out}}, 1 \le r \le m, r \ne j\}$ to denote the set of the local helper nodes, and the local helpers are denoted as the edges $e(N_{i,r}^{\text{ctrl}} \to \overline{N_{i,r}}^{\text{in}})$ with capacity α. Figure 3 gives an IFG example of ABAS for a ($n = 3, k = 2, d = 2, m = 2, \ell = 1$) GRC. In such IFG, with every failed node is replaced by a newcomer, and a new available cluster will replicate the other $m-1$ nodes in the cluster and simultaneously has a pair of node, an edge $e(N_{i,r}^{\text{in}} \to N_{i,r}^{\text{out}}), 1 \le r \le m, r \ne j$ and a control node $\overline{N_i}^{ctrl}$ responsible for the external communication. We use the notation $\mathcal{C}_i(s)$ to the available cluster of i^{th} cluster in stage s, $0 \le s \le n_i$. Moreover, we use $N_{i,j}^{\text{in}}(s)$, $N_{i,j}^{\text{out}}(s)$, $\overline{N_i}^{\text{ctrl}}(s)$, $1 \le j \le m$ to denote the nodes in $\mathcal{C}_i(s)$. The pair of node $N_{i,j}^{\text{in}}(s)$, $N_{i,j}^{\text{out}}(s)$ can be simplified into $N_{i,j}(s)$. Moreover, we use $\mathcal{N}_i$ to denote all the nodes of i^{th} cluster after s repairs, including the available and unavailable clusters.

3.3 | The upper bound of capacity

In this section, we derive the upper bound of capacity that ABAS can achieve, by analyzing their IFGs under functional repair setting.

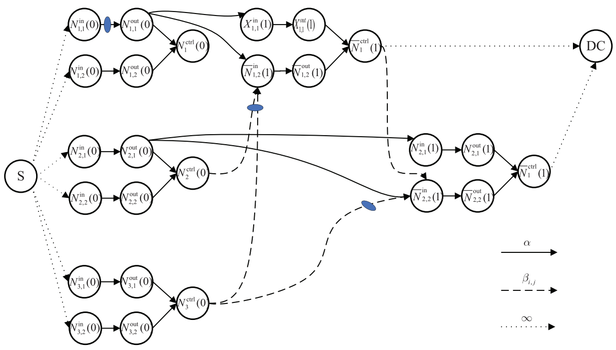


FIGURE 4 A min-cut example of ABAS for a $(n = 3, k = 2, d = 2, m = 2, \ell = 1)$ GRC's IFG

3.3.1 | Min-cut analysis

As Figure 4 shows, it is a min-cut example of ABAS for a $(n = 3, k = 2, d = 2, m = 2, \ell = 1)$ GRC's IFG. The node failure sequence is $(N_{1,2}, N_{2,2})$, and the DC connects the control node $\overline{N}_1^{\text{ctrl}}$ and $\overline{N}_2^{\text{ctrl}}$ to acquire the whole content of cluster $\mathcal{C}_1(1)$ and cluster $\mathcal{C}_2(1)$. Taking the whole single cluster as an entirety, we sort all the clusters topologically as $(\mathcal{C}_1(1), \mathcal{C}_2(1))$. At this time, for the 1^{th} repair, the min-cut experiences the edge $e(N_{1,1}^{\text{in}}(0) \rightarrow N_{1,1}^{\text{out}}(0))$ with cut-value α. Additionally, the min-cut also experiences these two edges $e(N_2^{\text{ctrl}}(0) \rightarrow \overline{N}_{1,2}^{\text{in}}(1))$ and $e(N_3^{\text{ctrl}}(0) \rightarrow \overline{N}_{1,2}^{\text{in}}(1))$ with cut-value $(\beta_{1,2} + \beta_{1,3})$ or the edge $e(N_{1,2}^{\text{in}}(1) \rightarrow N_{1,2}^{\text{out}}(1))$ with cut-value α. At this moment, we already attribute $\mathcal{C}_1$ to the sink. Thus the first node brings a contribution of $\alpha + \min\{\alpha, \beta_{1,2} + \beta_{1,3}\}$ to the cut. For the 2^{th} repair, the edges that the min-cut experiences have $e(N_{2,1}^{\text{in}}(0) \rightarrow N_{2,1}^{\text{out}}(0))$ with cut-value α, and one of the two edges $e(N_3^{\text{ctrl}}(0) \rightarrow \overline{N}_{2,2}^{\text{in}}(1))$ with cut-value $\beta_{2,3}$ or $e(N_{2,2}^{\text{in}}(1) \rightarrow N_{2,2}^{\text{out}}(1))$ with cut-value α. The min-cut contribution from the second node is $\alpha + \min\{\alpha, \beta_{2,3}\}$. Therefore, the min-cut value is $(2\alpha + \min\{\alpha, \beta_{1,2} + \beta_{1,3}\} + \min\{\alpha, \beta_{2,3}\})$. From Figure 4, we can observe that for the repair of $N_{1,2}$, the index set of the tail of edge from remote helper cluster to $\overline{N}_{1,2}^{\text{in}}(1)$ is 2, 3, and for the repair of $N_{2,2}$, it turns to 3.

Now we begin to analyze the min-cut of ABAS for GRCs in general case, whose brief diagram is shown in Figure 5. Following Prakash' proof of Theorem 1, we also consider a failure sequence of $k(m-\ell)$ failures and repairs that $N_{i,\ell+1}, N_{i,\ell+1}, \dots, N_{i,m}$ fails successively from $i = 1$ to $i = k$. The corresponding nodes can be described as $N_{1,\ell+1}(0), N_{1,\ell+2}(1), .., N_{1,m}(m-\ell-1), N_{2,\ell+1}(0), \dots, N_{2,m}(m-\ell-1), \dots, N_{k,m}(m-\ell-1)$. The newcomer of each failed node $N_{i,\ell+t}(t-1), 1 \leq t \leq m-\ell$, needs to contact the local helper nodes $N_{i,1}(t-1), N_{i,2}(t-1), \dots, N_{i,\ell}(t-1)$ in the host cluster and the remote helper clusters $\mathcal{C}_1(m-\ell), \dots, \mathcal{C}_{i-1}(m-\ell)$ which have been repaired and other $d - \min\{i-1, d\} = (d-i+1)^+$ available clusters to acquire help data. Let DC connect to the last k available clusters $\mathcal{C}_1(m-\ell), \dots, \mathcal{C}_k(m-\ell)$. The edges that the min-cut experiences including that:

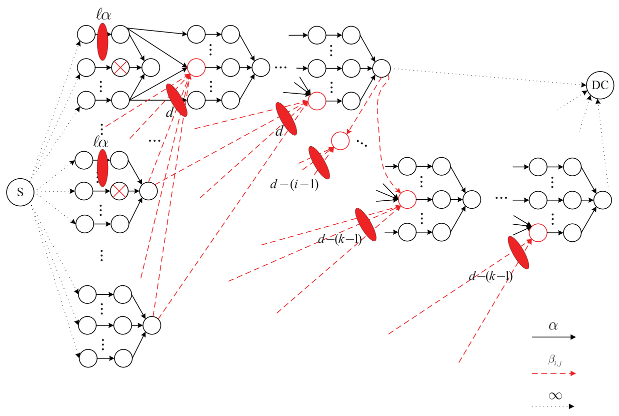


FIGURE 5 The brief diagram about the min-cut of asymmetric repair GRCs in general case

- $\{e(N_{i,j}^{\text{in}}(0) \rightarrow N_{i,j}^{\text{out}}(0)), i \in [k], j \in [\ell]\}$, whose total capacity is $\ell k\alpha$.
- Either the set of edges $\{e(N_{i,\ell+t}^{\text{in}}(t) \rightarrow e(N_{i,\ell+t}^{\text{out}}(t)), i \in [k], t \in [m-\ell]\}$, each of them has capacity α. Or the set of inter-cluster help edges $\{e(N_{h_i}^{\text{ctrl}}(0) \rightarrow N_{i,\ell+t}^{\text{in}}(t)), h_i \in \mathcal{H}_i\}$, where $\mathcal{H}_i \subseteq [n]\backslash[i], |\mathcal{H}_i| = d - \min\{i-1, d\}$. The total cut-value is given by (5):

$$\sum_{i=1}^{k}\left\{l\alpha + \sum_{j=\ell+1}^{m}\min\left\{\alpha, \min_{\mathcal{H}_i}\sum_{h_i\in\mathcal{H}_i}\beta_{i,h_i}\right\}\right\} = lk\alpha + (m-\ell)\sum_{i=1}^{k}\min\left\{\alpha, \min_{\mathcal{H}_i}\sum_{h_i\in\mathcal{H}_i}\beta_{i,h_i}\right\}. \tag{5}$$

Hence, we already have derived the upper bound of capacity B that ABAS for GRC can achieve under functional repair setting as Theorem 1 demonstrates.

Theorem 1. *The capacity B of ABAS for GRC having parameters (n, k, d, m, ℓ) under functional repair is upper bounded by 6:*

$$B \leq B^* \triangleq \ell k\alpha + (m-\ell)\sum_{i=1}^{k}\min\left\{\alpha, \min_{\mathcal{H}_i}\sum_{h_i\in\mathcal{H}_i}\beta_{i,h_i}\right\}, \tag{6}$$

where, $\mathcal{H}_i \subseteq [n]\backslash[i], |\mathcal{H}_i| = d - \min\{i-1, d\}$, which expresses the index set of the tail in the edge from remote helper clusters to the newcomer that the min-cut experiences, when repairing i^{th} cluster.

3.3.2 | The reachability of the upper bound

This part mainly involves the illustration about the reachability of the capacity upper bound B in 6 that for any valid IFG. We assume that any S-DC cut can divide the IFG G into two sets $(\mathcal{W}, \overline{\mathcal{W}})$, which represent the source and sink set, respectively. Since the capacity of edge DC is infinity, the source set $\mathcal{W}$ certainly contains the control nodes of last available clusters.

In a directed acyclic graph, if there exists a directed edge $e(A \to A')$, A must appears before A' in the topological sorted result. Therefore, following the topological sorted results, we can easily obtain our IFG's observation about the upper bound B for any S-DC cut $(\mathcal{W}, \overline{\mathcal{W}})$ as shown in (7):

$$\text{mincut(S-DC)} \geq \sum_{i=1}^{k} \left(\varepsilon_i \alpha + \sum_{j=\varepsilon_i+1}^{m} \min \left\{ \alpha, (\ell - j + 1)^{+} \alpha \right. \right.$$
$$\left. \left. + \min_{\mathcal{H}_i} \sum_{h_i \in \mathcal{H}_i} \beta_{i,h_i} \right\} \right), \tag{7}$$

where ε_i denotes the number of the out-nodes, which belong to $\overline{\mathcal{W}}$, for the i^{th} cluster that the cut passed, the right side of 7 can be simplified into

$$\max(\varepsilon_i, \ell) k\alpha + (m - \max(\varepsilon_i, \ell)) \sum_{i=1}^{k} \min \left\{ \alpha, \min_{\mathcal{H}_i} \sum_{h_i \in \mathcal{H}_i} \beta_{i,h_i} \right\}, \tag{8}$$

we also can get:

$$\text{mincut(S-DC)} \geq \max(\varepsilon_i, \ell) k\alpha + (m - \max(\varepsilon_i, \ell))$$
$$\sum_{i=1}^{k} \min \left\{ \alpha, \min_{\mathcal{H}_i} \sum_{h_i \in \mathcal{H}_i} \beta_{i,h_i} \right\}$$
$$\geq lk\alpha + (m - \ell) \sum_{i=1}^{k} \min \left\{ \alpha, \min_{\mathcal{H}_i} \sum_{h_i \in \mathcal{H}_i} \beta_{i,h_i} \right\}. \tag{9}$$

It can be known that, for the arbitrary cut of the information flow graph, regardless of the order of the repair sequence, the lower bound of the min-cut value is shown in this section, so the upper bound of capacity of ABAS for GRCs is always reachable. By analyzing the GRCs' IFG of ABAS, the upper bound of capacity always can be reached as long as the repair times has a bound. Hence, without loss of generality, we can assume the cost that transmit unit data between C_i and C_j gradually increases from $j = 1$ to $j = n$, to optimize the repair process, (6) can be simplified into

$$B \leq B^* \triangleq \ell k\alpha + (m - \ell) \sum_{i=1}^{k} \min \left\{ \alpha, \sum_{j=i+1}^{d+1} \beta_{i,j} \right\}. \tag{10}$$

3.4 | The constraint of local repair bandwidth

In the previous section, we have proved that the optimality of capacity of ABAS for GRCs. However, our assumption ignores the influences about intra-cluster repair bandwidth. In this section, we will separately consider the maximum capacity that can be achieved when the intra-bandwidth is constrained. It is known that, the intra-cluster repair bandwidth includes two parts, that is γ denotes intra-cluster repair bandwidth in host cluster and γ' denotes intra-cluster repair bandwidth in remote helper clusters. It is noted that, when we achieve one of them, the another is considered equal to α.

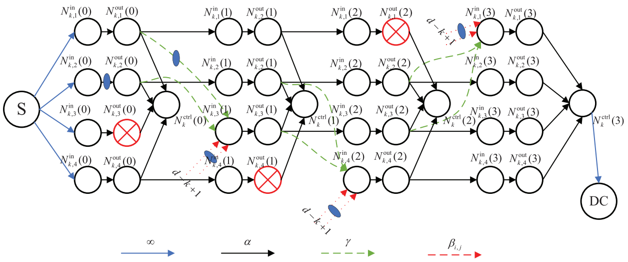


FIGURE 6 The failures and repairs in k clusters

3.4.1 | The lower bound of γ

Theorem 2 describes that for reaching the maximum capacity, the intra-cluster repair bandwidth in host cluster has a minimum constraint to satisfy.

Theorem 2. *For ABAS of GRCs under functional repair setting having parameters (n, k, d, m, ℓ), when $d \geq k, \gamma' = \alpha, \ell' = m$, to reach the maximum capacity, the intra-cluster repair bandwidth in host cluster γ is lower bounded by 11:*

$$\gamma \geq \gamma^* \triangleq \alpha - \min \left\{ \alpha, \sum_{j=k+1}^{d+1} \beta_{i,j} \right\}. \tag{11}$$

The Proof of Theorem 2 is similar to Theorem 2.

Proof. Consider a failure sequence of $k(m - \ell) + 1$ failures and repairs. The previous $k(m - \ell)$ repairs is the same as in Theorem 1, the main difference is that in the k^{th} cluster, and the later $m - \ell$ repair sequence $N_{k,\ell+1}, N_{k,\ell+1}, \ldots, N_{k,m}$. The node $N_{k,1}$ also needs to be repaired, thus the failure sequence in k^{th} cluster is denoted as $N_{k,\ell+1}(0), N_{k,\ell+2}(1), .., N_{k,m}(m - \ell - 1), N_{k,1}(m - \ell)$. For the repair of $N_{k,\ell+1}(0)$, the set of local helper node is $\{N_{k,1}(0), N_{k,1}(0), \ldots, N_{k,\ell}(0)\}$, while for the other node $N_{k,(\ell+t+1)}(t), 1 \leqslant t \leqslant m - \ell$, their set of local helper node is $\{N_{k,2}(t), N_{k,3}(t), \ldots, N_{k,\ell+1}(t)\}$. $C_1(m - \ell), C_2(m - \ell), \ldots, C_{\min(d,k-1)}(m - \ell)$ represents the remote helper clusters to repair the node failure in k^{th} cluster. Figure 6 gives an example of an IFG about k clusters with $(m = 4, \ell = 2)$. In Figure 6, 3rd 4th, 1st node fail successively, when 3rd or 4th node fails, 1st, 2nd nodes undertake the local helper nodes, while 1st node fails, 2nd, 3rd nodes undertake the local helper nodes. The DC contacts $\mathcal{N}_1(m - \ell), \mathcal{N}_2(m - \ell), \ldots, \mathcal{N}_k(m - \ell + 1)$ to reconstruct the original file, the min-cut can be analyzed by divided into two parts:

1) From 1st to $(k-1)^{\text{th}}$ clusters, it is the same as Theorem 1. The cut-value is $(k-1)\ell\alpha + (m - \ell)\sum_{i=1}^{k-1} \min\{\alpha, \sum_{j=i+1}^{i+(d+i+1)^+} \beta_{i,j}\}$.

2) In k^{th} cluster,
   a) For $[e(N_{k,1}^{\text{out}}(0) \rightarrow N_{k,\ell+1}^{\text{in}}(1))]$, the cut-value is γ.
   b) For $[e(N_{k,j}^{\text{in}}(0) \rightarrow N_{k,j}^{\text{out}}(0)), \forall j \in [2, \ell]]$, the cut-value is $(\ell - 1)\alpha$.
   c) Either the edge set $[\{e(X_{k,\ell+1}^{\text{in}}(t+1) \rightarrow X_{k,\ell+t+1}^{\text{out}}(t+1)), 0 \leqslant t \leqslant (m-\ell)\}]$ or $\{e(N_{i'}^{\text{ctrl}}(0) \rightarrow N_{k,\ell+t+1}^{\text{in}}(t+1))\}$, where i' belongs to the subtraction between the index set of remote helper clusters that $X_{k,\ell+t+1}^{\text{in}}(t+1), 0 \leqslant t \leqslant (m-\ell)$ connects. Let $d \geq k$, the cut-value of this part is $(m-\ell+1)\min\{\alpha, \sum_{j=k+1}^{k+(d-k+1)^+} \beta_{i,j}\}$.

The total min-cut value can be described as 12:

$$\begin{aligned}\text{mincut(S-DC)} = (k-1)\ell\alpha + (m-\ell)\sum_{i=1}^{k}\Bigg\{\min(\alpha,\\ \sum_{j=i+1}^{d+1}\beta_{i,j})\Bigg\} - \alpha + \min\Bigg\{\alpha, \sum_{j=k+1}^{i+(d-i+1)^+}\beta_{i,j}\Bigg\}\\ + \gamma + (\ell-1)\alpha + (m-\ell+1)\min\Bigg\{\alpha, \sum_{j=k+1}^{k+(d-k+1)^+}\beta_{i,j}\Bigg\},\end{aligned} \tag{12}$$

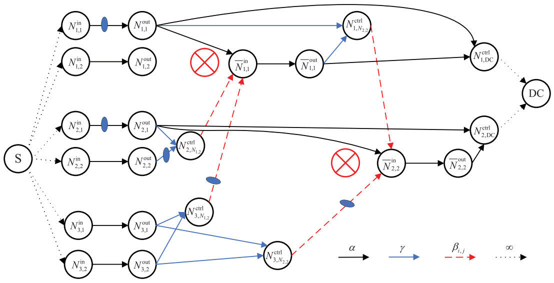
after being simplified, it turns to 13:

$$\begin{aligned}\text{mincut(S-DC)} &= k\ell\alpha + (m-\ell)\sum_{i=1}^{k}\min\left(\alpha, \sum_{j=i+1}^{d+1}\beta_{i,j}\right)\\ &\quad - \alpha + \min\left\{\alpha, \sum_{j=k+1}^{d+1}\beta_{i,j}\right\}\\ &= B^* - \alpha + \min\left\{\alpha, \sum_{j=k+1}^{d+1}\beta_{i,j}\right\} + \gamma.\end{aligned} \tag{13}$$

According to Theorem 1, for a system with bounded failures, no matter what the failure sequence is, the upper bound of the reachable capacity of ABAS for GRCs is shown in 10, so the left side of 13 should be larger than or equal to 10. Therefore, we can get 11.

Next, we further prove that when $\gamma \geq \gamma^*$, the upper bound of any valid IFG under any repair sequence is still B^*. Substitute the intra-cluster repair bandwidth in 7 with γ, we can get the 14:

$$\begin{aligned}\text{mincut(S-DC)} \geq \sum_{i=1}^{k}\left(\varepsilon_i\alpha + \sum_{j=\varepsilon_i+1}^{m}\min\left\{\alpha, (\ell-j+1)^+\gamma\right.\right.\\ \left.\left.+ \min_{\mathcal{H}_i}\sum_{h_i\in\mathcal{H}_i}\beta_{i,h_i}\right\}\right).\end{aligned} \tag{14}$$



S
$N_{1,1}^{\text{in}}$
$N_{1,1}^{\text{out}}$
$N_{1,2}^{\text{in}}$
$N_{1,2}^{\text{out}}$
$N_{1,N_{2,2}}^{\text{ctrl}}$
$\overline{N}_{1,1}^{\text{in}}$
$\overline{N}_{1,1}^{\text{out}}$
$N_{1,\text{DC}}^{\text{ctrl}}$
DC
$N_{2,1}^{\text{in}}$
$N_{2,1}^{\text{out}}$
$N_{2,2}^{\text{in}}$
$N_{2,2}^{\text{out}}$
$N_{2,N_{1,2}}^{\text{ctrl}}$
$N_{2,\text{DC}}^{\text{ctrl}}$
$\overline{N}_{2,2}^{\text{in}}$
$\overline{N}_{2,2}^{\text{out}}$
$N_{3,N_{1,2}}^{\text{ctrl}}$
$N_{3,1}^{\text{in}}$
$N_{3,1}^{\text{out}}$
$N_{3,2}^{\text{in}}$
$N_{3,2}^{\text{out}}$
$N_{3,N_{2,2}}^{\text{ctrl}}$
α
γ
$\beta_{i,j}$
∞

FIGURE 7 An example of IFG about ABAS for GRCs having parameters $(n = 3, k = 2, d = 2, m = 2, \ell = 1)$

Then, we simplify 14 to 15:

$$\begin{aligned}\text{mincut(S-DC)} \geq \sum_{i=1}^{k}\left(\varepsilon_i\alpha + \sum_{j=\varepsilon_i+1}^{m}\min\left\{\alpha,\right.\right.\\ \left.\left.(\ell-j+1)^+\gamma + \sum_{s=i+1}^{i+(d-i+1)^+}\beta_{i,s}\right\}\right).\end{aligned} \tag{15}$$

At this moment, $\gamma \geq \gamma^*$, we get 16:

$$(\ell-j+1)^+\gamma + \sum_{s=i+1}^{i+(d-i+1)^+}\beta_{i,s} \geqslant \alpha. \tag{16}$$

No matter $j \leqslant \ell, i \leqslant k$, 11 can be obtained. □

3.4.2 | The lower bound of γ'

Theorem 3 illustrates that for reaching the maximum capacity, the intra-cluster repair bandwidth in remote helper cluster has a minimum constraint to satisfy.

Theorem 3. *For ABAS of GRCs under functional repair setting having parameters (n, k, d, m, ℓ), when $d \geq k, \gamma = \alpha, \ell' = m$, to reach the maximum capacity, the intra-cluster repair bandwidth in remote helper cluster γ' is lower bounded by (17):*

$$\gamma' \geqslant \frac{\min\left\{\sum_{j=i+1}^{k}\beta_{i,j}, \alpha - \sum_{s=k+1}^{d+1}\beta_{i,s}\right\}}{(k-i)(m-\ell)}, \forall i \in [1, k-1]. \tag{17}$$

Proof. Similarly, consider same failure and repair sequences as Theorem 1 proof does, Figure 7 gives an example of IFG about ABAS for GRCs having parameters $(n = 3, k = 2, d = 2, m = 2, \ell = 1)$. The min-cut of such IFG consists of the following parts:

1) The edge set $\{e(N_{i,j}^{\text{in}} \rightarrow N_{i,j}^{\text{out}}), i \in [k], j \in [\ell]\}$, the cut-value is $\ell k\alpha$;
2) For each $i, i \in [k]$, the min-cut experiences either the edge set $\{e(\overline{N}_{i,j}^{\text{in}} \rightarrow \overline{N}_{i,j}^{\text{out}}), j \in [\ell+1, m]$ with cut-value

$(m-\ell)\alpha$, or $\{\{e(X_{i_1}^{\mathrm{ctrl}} \to \overline{X}_{i,j}^{\mathrm{in}})\} \cup \{e(X_{i_2,j}^{\mathrm{out}} \to X_{i_2}^{\mathrm{ctrl}})\}, i_1 \in [k+1, d+1], i_2 \in [i+1, k], j \in [\ell+1, m]\}$ with cut-value $(m-\ell)[\sum_{s=k+1}^{d+1} \beta_{i,s} + (\min\{k, d+1\} - i) + (m-\ell)\gamma']$.

So we obtain the total min-cut value as (18):

$$\mathrm{mincut(S\text{-}DC)} = \ell k\alpha + (m-\ell)\sum_{i=1}^{k} \min\left\{\alpha, \sum_{s=k+1}^{d+1} \beta_{i,s} + (\min\{k, d+1\} - i)^{+}(m-\ell)\gamma'\right\}. \tag{18}$$

According to Theorem 1, we have (19):

$$\mathrm{mincut(S\text{-}DC)} \geqslant B^{*} = k\ell\alpha + (m-\ell)\sum_{i=1}^{k} \min\left\{\alpha, \sum_{j=i+1}^{d+1} \beta_{i,j}\right\}. \tag{19}$$

Substitute (6) into 19, we can obtain (20):

$$\min\{\alpha, \sum_{s=k+1}^{d+1} \beta_{i,s} + (\min\{k, d+1\} - i)^{+}(m-\ell)\gamma'\} \\ \min\left\{\alpha, \sum_{j=i+1}^{d+1} \beta_{i,j}\right\}, \forall i \in [k]. \tag{20}$$

Then we have (21) and (22), for $\forall i \in [k]$,

$$\begin{cases} \sum\limits_{s=k+1}^{d+1} \beta_{i,s} + (\min\{k, d+1\} - i)^{+}(m-\ell)\gamma' \geqslant \sum\limits_{j=i+1}^{d+1} \beta_{i,j} \\ \sum\limits_{s=k+1}^{d+1} \beta_{i,s} + (\min\{k, d+1\} - i)^{+}(m-\ell)\gamma' \geqslant \alpha \end{cases}, \tag{21}$$

$$\begin{cases} \gamma' \geqslant \frac{\sum\limits_{j=i+1}^{k} \beta_{i,j}}{(k-i)(m-\ell)} \\ \gamma' \geqslant \frac{\alpha - \sum\limits_{s=k+1}^{d+1} \beta_{i,s}}{(k-i)(m-\ell)} \end{cases}, \forall i \in [1, k-1], \tag{22}$$

Finally, we get (17) proved. □

4 | GRBC OF GRC REPAIR FAILURES

4.1 | Global repair bandwidth cost

In CDSSs, data needs to be sent to the newcomer from the help nodes in different clusters when repairing the failure node. In the above section, we proposed ABAS of GRCs to optimize the data repair bandwidth during node repair, and derived the constraints of two local repair bandwidths γ and γ', and get their lower bounds for achieving the maximum capacity. For GRCs, when repairing the failure node, another two kinds of bandwidths will be considered, which are intra-cluster bandwidth and inter-cluster bandwidth. Both intra-cluster bandwidth and inter-cluster bandwidth directly determine the transmission cost when repairing the failed node. In actual channel transmission, the costs of transmitting data within the cluster and without the cluster are different. Similarly, different inter-cluster bandwidth costs are also different. Therefore, we propose a definition of global repair bandwidth cost (GRBC) to represent the total data transmission overhead of GRCs during node repair.

We first propose the definition of transmission cost factor. We define unit cost of intra-cluster bandwidth, which means the cost of occupying the unit intra-cluster bandwidth is 1. In most cases, the intra-cluster data transmission is wired, so we can ignore the difference in costs. Therefore, ρ represents the cost ratio of inter-cluster to intra-cluster bandwidths. Besides, when different clusters perform cross-cluster data transmission, the cost factor is considered to be different either. For simplifying the analysis, the transmission cost model when repairing a failure node can be described as a $n \times n$ symmetric matrix $\mathbf{\Psi}$ as 23, where $\rho_{i,j}$ represents the ratio of the cost of inter-cluster bandwidth from cluster j to cluster i and the cost of the intra-cluster bandwidth. When repairing a failure node, GRCs permit any other d clusters transmitting data to the newcomer, of which the total bandwidth cost has $\binom{n}{d}$ possibilities, represented by $\psi_{\mathcal{H}_i}$, where $|\mathcal{H}_i| = d$ represents the index set of remote helper cluster. Without loss of generality, we simplify $\psi_{\mathcal{H}_i}$ as ψ.

$$\mathbf{\Psi} = \begin{bmatrix} 1 & \rho_{12} & \cdots & \rho_{1i} & \cdots & \rho_{1n} \\ \rho_{21} & 1 & \cdots & \rho_{2i} & \cdots & \rho_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{i1} & \rho_{i2} & \cdots & 1 & \cdots & \rho_{1n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \rho_{n1} & \rho_{n2} & \cdots & \rho_{ni} & \cdots & 1 \end{bmatrix}. \tag{23}$$

For symmetric repair strategy of GRCs, due to ignoring the difference of inter-cluster bandwidth, matrix $\mathbf{\Psi}$ can be considered as a matrix consists of two elements. $\rho_{i,j}$ can be expressed as:

$$\rho_{i,j} = \begin{cases} 1, & i = j \\ \rho, & i \neq j \end{cases}. \tag{24}$$

Upon the above concepts, we define a more comprehensive target to describe the total bandwidth cost — *Global Repair Bandwidth Cost* (GRBC). In the above sections, we have introduced how the GRC repairs the failed node. When repairing a single node, in the intra-cluster repair, we need l nodes with each node transmit γ symbols to the newcomer. And in the inter-cluster repair, we need d remote clusters with each cluster transmit β symbols to the newcomer. In each remote cluster, there are l'

nodes, and each node transmits γ' symbols to calculate β. So, the function expression of the GRBC is as follows:

$$C_{\text{global}} = \ell\gamma + d\ell'\gamma' + \psi\beta. \tag{25}$$

For symmetric repair model, GRBC of GRCs satisfies the formula (25).

Under ABAS, given a CDSS with $\boldsymbol{\Psi}$ fixed, when repairing a node in i^{th} cluster, the newcomer contacts any d remote helper clusters from other $n-1$ clusters. Assume these d clusters' indexes forms $\mathcal{H}_i$, and $\beta_{i,h_i}, h_i \in \mathcal{H}_i$ denotes the inter-cluster repair bandwidth from remote helper cluster C_{h_i} to C_i. Therefore, we can get the GRBC of ABAS for GRCs when repairing node in i^{th} cluster as (26).

$$C_{\text{global}}(i) = \ell\gamma + d\ell'\gamma' + \sum_{h_i \in \mathcal{H}_i} \rho_{i,h_i}\beta_{i,h_i}. \tag{26}$$

4.2 | Optimization of global repair bandwidth cost

Considering the intra-cluster bandwidth, in order to reach the optimal capacity, it must satisfy the minimization constraint either in host cluster as shown in (11) or in remote helper cluster as shown in (17), and we change them into (27) and (28) respectively.

$$\gamma \geq \gamma(\boldsymbol{\beta}) = \alpha - \min\{\alpha, \boldsymbol{e}_{[k+1,d+1]}\boldsymbol{\beta}_{[k+1,d+1]}\}, \tag{27}$$

$$\begin{aligned} \gamma' \geqslant \gamma'(\boldsymbol{\beta}) = \min\{&\boldsymbol{e}_{[i+1,k]}\boldsymbol{\beta}_{[i+1,k]}, \\ &\alpha - \boldsymbol{e}_{[k+1,d+1]}\boldsymbol{\beta}_{[k+1,d+1]}\}/\xi_i, \forall i \in [k-1]. \end{aligned} \tag{28}$$

In (27) and (28), $\boldsymbol{\beta}$ is a n-dimensional inter-cluster repair bandwidth column-vector, $\boldsymbol{\beta} = [\beta_{i,1}, \beta_{i,2}, \ldots, \beta_{i,n}]^{\text{T}}, \beta_{i,i} = 0$, and $\boldsymbol{\beta}_{[k+1,d+1]}$ represents a sub-vector from the $(k+1)^{\text{th}}$ to the $(d+1)^{\text{th}}$ in $\boldsymbol{\beta}$, $\boldsymbol{\beta}_{[k+1,d+1]} = [\beta_{i,k+1}, \ldots, \beta_{i,d+1}]^{\text{T}}$. Meanwhile, $\boldsymbol{e}_{[k+1,d+1]}$ denotes a row-vector whose elements are all 1, and dimension is the same as $\boldsymbol{\beta}_{[k+1,d+1]}$. ξ_i and $\xi_i = (k-i)(m-l)$. Based on the following descriptions, we formulate the GRBC of ABAS for GRCs as linear programming problems $\mathcal{P}_1$ and $\mathcal{P}_2$, which correspond to the condition γ and γ' are limited by the lower bound, respectively.

$$\mathcal{P}_1 : \min_{\boldsymbol{\beta}} \quad c_i(\boldsymbol{\beta}) \triangleq \ell\gamma + md\alpha + \boldsymbol{\rho}_i\boldsymbol{\beta}_i, \tag{29}$$

$$\text{s.t.} \quad \boldsymbol{\beta} \geq 0, \tag{30}$$

$$\gamma(\boldsymbol{\beta}) \leq \gamma, \tag{31}$$

$$k\ell\alpha + (m-\ell)\sum_{i=1}^{k} \min\{\alpha, \min_{\mathcal{H}_i} \boldsymbol{e}_i\boldsymbol{\beta}_i\} \geq B. \tag{32}$$

$$\mathcal{P}_2 : \min_{\boldsymbol{\beta}} \quad c_i(\boldsymbol{\beta}) \triangleq \ell\alpha + md\gamma' + \boldsymbol{\rho}_i\boldsymbol{\beta}_i, \tag{33}$$

$$\text{s.t.} \quad \boldsymbol{\beta} \geq 0, \tag{34}$$

$$\gamma'(\boldsymbol{\beta}) \leq \gamma', \tag{35}$$

$$k\ell\alpha + (m-\ell)\sum_{i=1}^{k} \min\{\alpha, \min_{\mathcal{H}_i} \boldsymbol{e}_i\boldsymbol{\beta}_i\} \geq B. \tag{36}$$

In $\mathcal{P}_1$ and $\mathcal{P}_2$, $c_i(\boldsymbol{\beta})$ represents the GRBC for repairing arbitrary single node in i^{th} cluster, $\boldsymbol{\rho}_i$ represents a row-vector composed of $\rho_{i,j}$ from C_j to C_i, $\boldsymbol{\rho}_i = \{\rho_{i,j}, j \in \mathcal{H}_i\}$ and $\boldsymbol{\beta}_i$ represents the inter-cluster repair bandwidth column-vector for repairing node in i^{th} cluster, $\boldsymbol{\beta}_i = \{\beta_{i,j}, j \in \mathcal{H}_i\}$.

Furthermore, in order to simplify the constraints in $\mathcal{P}_1$ and $\mathcal{P}_2$, we consider in the system the cost ratio $\rho_{i,j}, i \neq j$ increases with j changes from 1 to n, that is $\rho_{i,1} \leq \rho_{i,2} \leq \ldots \leq \rho_{i,i-1} \leq \rho_{i,i+1} \ldots \leq \rho_{i,n}$. It is obvious that to minimizing the GRBC of repairing a single node, it is profitable for the host cluster to choose the cheaper help connection, which indicates the host cluster will contact the remote helper clusters between which $\rho_{i,j}$ are less. Therefore, we confirm the index set of the remote helper clusters that the host cluster C_i contacts should be $\mathcal{H}_i = [d+1]\backslash i$, the optimization problems $\mathcal{P}_1$ and $\mathcal{P}_2$ transform into $\mathcal{P}_3$ and $\mathcal{P}_4$.

$$\mathcal{P}_3 : \min_{\boldsymbol{\beta}} \quad c_i(\boldsymbol{\beta}) \triangleq \ell\gamma + md\alpha + \boldsymbol{\rho}_i\boldsymbol{\beta}_i, \tag{37}$$

$$\text{s.t.} \quad \gamma(\boldsymbol{\beta}) \leq \gamma, \tag{38}$$

$$\beta_{i,1} \geq \ldots \geq \beta_{i,i-1} \geq \beta_{i,i+1} \geq \ldots \geq \beta_{i,n} \geq 0, \tag{39}$$

$$k\ell\alpha + (m-\ell)\sum_{i=1}^{k} \min\{\alpha, \min_{\mathcal{H}_i} \boldsymbol{e}_i\boldsymbol{\beta}_i\} \geq B. \tag{40}$$

$$\mathcal{P}_4 : \min_{\boldsymbol{\beta}} \quad c_i(\boldsymbol{\beta}) \triangleq \ell\alpha + md\gamma' + \boldsymbol{\rho}_i\boldsymbol{\beta}_i, \tag{41}$$

$$\text{s.t.} \quad \gamma'(\boldsymbol{\beta}) \leq \gamma', \tag{42}$$

$$\beta_{i,1} \geq \ldots \geq \beta_{i,i-1} \geq \beta_{i,i+1} \geq \ldots \geq \beta_{i,n} \geq 0, \tag{43}$$

$$k\ell\alpha + (m-\ell)\sum_{i=1}^{k} \min\{\alpha, \min_{\mathcal{H}_i} \boldsymbol{e}_i\boldsymbol{\beta}_i\} \geq B. \tag{44}$$

It is obvious to see from $\mathcal{P}_3$ and $\mathcal{P}_4$ that, the constraint on the inter-cluster repair bandwidth vector has changed, and the reason is proved in the following description.

Proof. Assume the optimal solution of $\mathcal{P}_1$ or $\mathcal{P}_2$ is $\boldsymbol{\beta}_1^* = [\beta_{i,1}^*, \ldots, \beta_{i,i-1}^*, 0, \beta_{i,i+1}^*, \ldots, \beta_{i,n}^*]$, which satisfies the constraint

(32) or (36). If there exists $\beta^*_{i,i_1} \geq \beta^*_{i,i_2}$, for some $i_1 \leq i_2$, exchange the values of β^*_{i,i_1} and β^*_{i,i_2}, the feasible solution set for the optimization problems $\mathcal{P}_1$ or $\mathcal{P}_2$ remains unchanged, since the constraints in $\mathcal{P}_1$ or $\mathcal{P}_2$ unchanged. However, the optimization object function will decrease, because $\rho_{i,j}$ multiplied by a larger β^*_{i,i_1} reduces, which violates the previous hypothesis that $\boldsymbol{\beta}^*_1 = [\beta^*_{i,1}, \dots, \beta^*_{i,i-1}, 0, \beta^*_{i,i+1}, \dots, \beta^*_{i,n}]$ is the optimal solution. Thus, for an optimal solution $\boldsymbol{\beta}^*_1$, it should meets that $\beta_{i1} \geqslant \dots \geqslant \beta_{i,i-1} \geqslant \beta_{i,i+1} \dots \geqslant \beta_{in} \geqslant 0$. In other words, we multiply the larger $\rho_{i,j}$ with a smaller $\beta_{i,j}$, then GRBC can be optimized. □

It is noted that, when repairing different nodes in any cluster i, if it chooses different remote helper clusters every time, the $c_i(\boldsymbol{\beta})$ will differ, hence we further formulate the average GRBC for repairing the failure nodes in r clusters as shown in (45).

$$c_{\text{avg}}(\boldsymbol{\beta}) = \frac{1}{r} \sum_{i=1}^{r} c_i(\boldsymbol{\beta}). \tag{45}$$

Therefore, we formulate two more LP problems $\mathcal{P}_5$ and $\mathcal{P}_6$. It is noted that

$$\mathcal{P}_5 : \min_{\boldsymbol{\beta}} \quad c_{\text{avg}}(\boldsymbol{\beta}) \triangleq \ell\gamma + md\alpha + \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{\rho}_i \boldsymbol{\beta}_i, \tag{46}$$

$$\text{s.t.} \quad \gamma(\boldsymbol{\beta}) \leq \gamma, \tag{47}$$

$$k\ell\alpha + (m-\ell) \sum_{i=1}^{k} \min\{\alpha, \boldsymbol{e}_{[i+1,d+1]} \boldsymbol{\beta}_{[i+1,d+1]}\} \geqslant B. \tag{48}$$

$$\beta_{i,1} \geqslant \beta_{i,2} \geqslant \dots \geqslant \beta_{i,j} \geqslant \dots \geqslant \beta_{i,n} \quad \geqslant 0 \forall i \neq j, i \in [n], j \in [n] \tag{49}$$

$$\mathcal{P}_6 : \min_{\boldsymbol{\beta}} \quad c_{\text{avg}}(\boldsymbol{\beta}) \triangleq \ell\alpha + md\gamma' + \frac{1}{k-1} \sum_{i=1}^{k-1} \boldsymbol{\rho}_i \boldsymbol{\beta}_i, \tag{50}$$

$$\text{s.t.} \quad \gamma'(\boldsymbol{\beta}) \leq \gamma', \tag{51}$$

$$k\ell\alpha + (m-\ell) \sum_{i=1}^{k} \min\{\alpha, \boldsymbol{e}_{[i+1,d+1]} \boldsymbol{\beta}_{[i+1,d+1]}\} \geqslant B, \tag{52}$$

$$\beta_{i,1} \geqslant \beta_{i,2} \geqslant \dots \geqslant \beta_{i,j} \geqslant \dots \geqslant \beta_{i,n} \geqslant 0 \quad \forall i \neq j, i \in [n], j \in [n]. \tag{53}$$

It is easily to point out that the average GRBC expressed in $\mathcal{P}_5$ and $\mathcal{P}_6$ is just like the GRBC of GRCs under symmetric repair.

5 | NUMERICAL RESULTS

In this section, we will provide the numerical results by simulation diagrams. These numerical results are based on the constraints of intra-cluster bandwidths in host cluster and remote help cluster (γ and γ'), respectively. First, we compare the GRBC between ABAS of GRCs, symmetric repair of GRCs and RCs for different coding parameters and different cost factors $\boldsymbol{\rho}_i$. And then, we optimize ABAS of GRCs by changing parameters to further reduce the GRBCs.

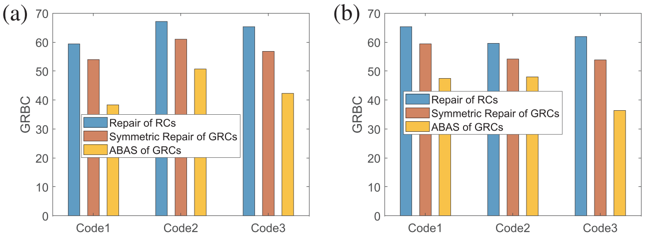


FIGURE 8 GRBCs comparison among different repair coding strategies under different construction parameter settings

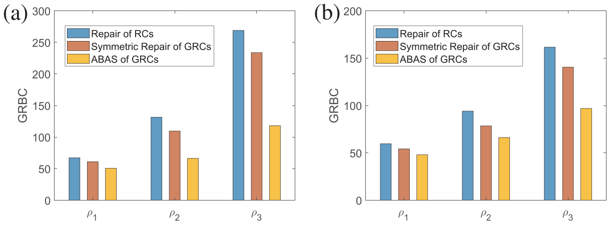


FIGURE 9 GRBCs comparison among different repair coding strategies under different cost factors of heterogeneous systems

5.1 | Comparison of GRBCs between different repair coding strategies

Numerical results are provided to show the comparison among three repair processes for different GRCs and RCs parameter settings. The results are both based on the Intra-cluster repair bandwidth constraints of host and helper clusters as γ and γ', respectively.

First, we will verify the advantages of ABAS of GRCs in reducing GRBC compared to other coding constructions under different parameters. In GRCs, the three construction parameter settings are Code1 ($n = 5, k = 3, d = 4, m = 4$), Code2 ($n = 7, k = 4, d = 5, m = 6$) and Code3 ($n = 7, k = 4, d = 6, m = 6$). We set $B = 36, l = 3, l' = m$ and $\boldsymbol{\rho} = [1 \ \ 2 \ \ 5 \ \ 10 \ \ 20 \ \ 50 \ \ 100]$. Notice that parameters n, k and d of RCs are the same as those of GRCs.

As Figure 8 shows, when $\boldsymbol{\rho} = [1 \ \ 2 \ \ 5 \ \ 10 \ \ 20 \ \ 50 \ \ 100]$, for the above three parameter settings, ABAS of GRCs reduces GRBC effectively compared with other repair strategies no matter γ or γ' is constrained. Moreover, we can get that the reduction of ABAS of GRCs is suitable for all other valid parameters.

Then, we will show the influence of cost ratio $\boldsymbol{\rho}$. Similarly, we choose the GRC with parameters of ($n = 7, k = 4, d = 5, m = 6, l = 3, l' = m$). We set three kinds of $\boldsymbol{\rho}$, which are $\rho_1 = [1 \ \ 2 \ \ 5 \ \ 10 \ \ 20 \ \ 50 \ \ 100]$, $\rho_2 = [1 \ \ 3 \ \ 15 \ \ 20 \ \ 30 \ \ 100 \ \ 200]$ and $\rho_3 = [1 \ \ 5 \ \ 20 \ \ 50 \ \ 100 \ \ 200 \ \ 500]$.

In Figure 9, these three cost ratio increases gradually, which quantified the heterogeneity of inter-cluster

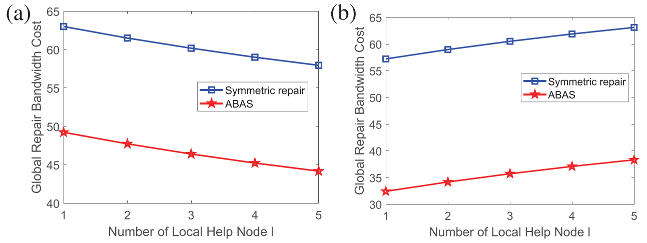


FIGURE 10 Relationship between optimized GRBCs and l with two intra-cluster repair bandwidth constraints

transmissions. When $\boldsymbol{\rho}_1 = [1 \quad 2 \quad 5 \quad 10 \quad 20 \quad 50 \quad 100]$, the reduction of cost is about 20.8%, and when $\boldsymbol{\rho}_3 = [1 \;\; 5 \;\; 20 \;\; 50 \;\; 100 \;\; 200 \;\; 500]$, the reduction achieves 40.28%. We can get a conclusion, if the heterogeneity of inter-cluster transmissions is larger, ABAS performs more effectively in reducing GRBC.

5.2 | GRBC optimization for ABAS

In this part, numerical results will be given to show the function between GRBC and parameters of GRCs. In GRCs, the parameter settings are $n = 7$, $k = 4$, $d = 5$, $m = 6$, $l' = m$ with l values among 1, 2, 3, 4 and 5, and the cost ratio $\boldsymbol{\rho}_i$ is set as $\boldsymbol{\rho}_i = [1 \;\; 2 \;\; 5 \;\; 10 \;\; 20 \;\; 50 \;\; 100]$. Results will be obtained with intra-bandwidth γ and γ' constraints, respectively. Compared (a) and (b), we can get the converse conclusion. In Figure 10a, GRBC of symmetric repair and ABAS decreased with the increase of l when γ is constraint. However, Figure 10b shows that when γ' is constraint, by increasing the number of l, GRBC increased either. Therefore, for reducing GRBC, if γ is constraint, we need to reduce l, and conversely, if γ' is constraint, l need to be increased.

6 | CONCLUSION

Based on the characteristics of heterogeneous network bandwidths between clusters in a clustered distributed storage system, this paper proposes an asymmetric bandwidth allocation strategy (ABAS) for generalized regenerating codes (GRCs), derives the upper bound of the capacity of ABAS for GRCs based on the information flow graph (IFG), and proves its capacity for any valid IFGs. In addition, based on the reachability of the upper bound, the lower bound of the intra-cluster repair bandwidth of the host cluster and the remote helper clusters are analyzed to obtain the minimum constraints.

In order to minimize the GRBC of ABAS for GRCs, the capacity upper bound constraint and the intra-cluster bandwidth minimum constraint are used as the subject constraints, and the vector composed of inter-cluster repair bandwidth between different clusters is used as an optimization variable. Linear programming problems of node repair and GRBC for the entire system are formulated and optimized. After numerical simulation to find the optimal solution, it was confirmed that even though the local bandwidth under ABAS performs no better than that under symmetric repair strategy, ABAS for GRCs can effectively reduce the GRBC under the condition of heterogeneous bandwidth costs between clusters compared to the symmetric repair strategy. The larger gap of cost coefficients, the better ABAS effect is, the maximum reduction in GRBC achieves 49.48%.

ACKNOWLEDGEMENTS

This work was supported by the National Natural Sciences Foundation of China under Grant 61831008 and Grant 62027802, the Guangdong Science and Technology Planning Project under Grant 2018B030322004 and Grant 2021A1515011572, the Shenzhen Natural Science Fund under Grant JCYJ20200109112822953, and Shenzhen Natural Science Fund (Stable Support Plan Program GXWD20201230155427003-20200824081029001).

CONFLICT OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

ORCID

Shushi Gu https://orcid.org/0000-0002-3897-5407
Tao Huang https://orcid.org/0000-0002-8098-8906
Wei Xiang https://orcid.org/0000-0002-0608-065X

REFERENCES

1. Beach, B.: AWS architecture overview. In: Pro Powershell for Amazon Web Services, pp. 1–6. Springer, Berkeley (2014)
2. Greenberg, A.: SDN for the cloud. In: Keynote of the 2015 ACM Conference on Special Interest Group on Data Communication, pp. 1–47. London (2015)
3. Facebook HDFS.: https://github.com/facebookarchive/hadoop-20 (2014). Accessed 11 Oct 2014
4. Jin, H., et al.: Approximate code: A cost-effective erasure coding framework for tiered video storage in cloud systems. In: Proceedings of ACM ICPP, pp. 1–10, New York (2019)
5. Reed, I., Solomon, G.: Polynomial codes over certain finite fields. J. Society Ind. Appl. Math. 8(2), 300–304 (1960)
6. Dimakis, A.G., Godfrey, P.B., Wu, Y., Wainwright, M.J., Ramchandran, K.: Network coding for distributed storage systems. IEEE Trans. Inf. Theory 56(9), 4539C4551 (2010)
7. Hu, Y., et al.: Optimal repair layering for erasure-coded data centers: From theory to practice. ACM Trans. Storage 13(4), 33 (2017)
8. Cai, C.X., Saeed, S., Gupta, I., Campbell, R.H., Le, F.: Phurti: Application and network-aware flow scheduling for multi-tenant mapReduce clusters. In: 2016 IEEE International Conference on Cloud Engineering (IC2E), pp. 161–170, Berlin (2016)
9. Prakash, Abdrashitov, V., Médard, M.: The storage versus repair-bandwidth trade-off for clustered storage systems. IEEE Trans. Inf. Theory 64(8), 5783–5805 (2018)
10. Norton, W.B.: Internet transit prices—Historical and projected. Report, DRPeering International (2010)
11. Gopalan, P., Huang, C., Simitci, H., Yekhanin, S.: On the locality of codeword symbols. IEEE Trans. Inf. Theory 58(11), 6925–6934 (2012)

12. Kamath, G.M., Prakash, N., Lalitha, V., Kumar, P.V.: Codes with local regeneration and erasure correction. IEEE Trans. Inf. Theory 60(8), 4637–4660 Aug. (2014)
13. Tamo, I., Barg, A., Frolov, A.: Bounds on the parameters of locally recoverable codes. IEEE Trans. Inf. Theory 62(6), 3070–3083 June (2016)
14. Hou, H., Lee, P.P.C., Han, Y.S.: Toward optimality in both repair and update via generic MDS code transformation. In: 2020 IEEE International Symposium on Information Theory (ISIT), pp. 560–565, Los Angeles (2020)
15. Dimakis, A.G., Godfrey, P.B., Wu, Y., Wainwright, M.J., Ramchandran, K.: Network coding for distributed storage systems. IEEE Trans. Inf. Theory 56(9), 4539–4551 (2010)
16. Sohn, J., Choi, B., Yoon, S.W., Moon, J.: Capacity of clustered distributed storage. IEEE Trans. Inf. Theory 65(1), 81–107 (2019)
17. Gastón, B., Pujol, J., Villanueva, M.: A realistic distributed storage system: The rack model. arXiv preprint, arXiv:1302.5657 (2013)
18. Hu, Y., Lee, P.P., Zhang, X.: Double regenerating codes for hierarchical data centers. In: 2016 IEEE International Symposium on Information Theory (ISIT), pp. 245–249, Barcelona (2016)
19. Pernas, J., Yuen, C., Gastón, B., Pujol, J.: Non-homogeneous two-rack model for distributed storage systems. In: 2013 IEEE International Symposium on Information Theory (ISIT), pp. 1237–1241. Istanbul, Turkey (2013)
20. Calis, G., Koyluoglu, O.O.: Architecture-aware coding for distributed storage: Repairable block failure resilient codes. arXiv preprint, arXiv:1605.04989, (2016)
21. Hou, H., Lee, P.P.C., Shum, K.W., Hu, Y.: Rack-aware regenerating codes for data centers. IEEE Trans. Inf. Theory 65(8), 4730–4745 (2019)
22. Zhang, Z., Zhou, L.: Rack-aware regenerating codes with fewer helper racks. http://arxiv.org/abs/2101.08738 (2021). Accessed 21 Jan 2021
23. Shah, N.B., Rashmi, K., Kumar, P.V.: A flexible class of regenerating codes for distributed storage. In: 2010 IEEE International Symposium on Information Theory (ISIT), pp. 1943–1947. Austin, TX, USA (2010)
24. Shum, K.W., Hu, Y.: Cooperative regenerating codes. IEEE Trans. Inf. Theory 59(11), 7229–7258 (2013)
25. Shen, Z., Shu, J., Huang, Z., Fu, Y.: Cluster-aware scattered repair in erasure-coded storage: Design and analysis. IEEE Trans. Comput. early access, (2020). https://doi.org/10.1109/TC.2020.3028353
26. Venkataramanachary, V., Reveron, E., Shi, W.: Storage and rack sensitive replica placement algorithm for distributed platform with data as files. In: 2020 International Conference on COMmunication Systems NETworkS (COMSNETS), pp. 535–538, Bengaluru (2020)
27. Zhang, Q., Zhang, S.Q., Leon-Garcia, A., Boutaba, R.: Aurora: Adaptive block replication in distributed file systems. In: 2015 IEEE 35th International Conference on Distributed Computing Systems, pp. 442–451, Columbus (2015)
28. Liu, P., Zheng, L., Yu, Q., Ye, H.: Tradeoff between storage cost and repair cost for cloud storage. In: 2018 IEEE 3rd International Conference on Cloud Computing and Big Data Analysis (ICCCBDA), pp. 169–173, Chengdu (2018)
29. Ernvall, T., El Rouayheb, S., Hollanti, C., Poor, H.V.: Capacity and security of heterogeneous distributed storage systems. IEEE J. Selected Areas Commun. 31(12), 2701–2709 (2013)
30. Akhlaghi, S., Kiani, A., Ghanavati, M.R.: Cost-bandwidth tradeoff in distributed storage systems. Comput. Commun. 33(17), 2105–2115 (2010)
31. Qu, S., Zhang, J., Wang, X.: Asymmetric regenerating codes for heterogeneous distributed storage systems. In: 2018 16th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), Shanghai (2018)
32. Armstrong, C., Vardy, A.: Distributed storage with communication costs. In: 2011 49th Annual Allerton Conference on Communication, Control, and Computing (Allerton), pp. 1358–1365, Monticello (2011)
33. Gerami, M., Xiao, M.: Exact optimized-cost repair in multi-hop distributed storage networks. In: 2014 IEEE International Conference on Communications (ICC), pp. 4120–4124, Sydney (2014)
34. Yu, Q., Sung, C.W., Chan, T.H.: Repair topology design for distributed storage systems. In: 2012 IEEE International Conference on Communications (ICC), pp. 7009–7013, Ottawa (2012)
35. Qin, S., Li, Z.: Network topology impacts on repair cost in distributed storage system with network coding. In: 2018 IEEE International Conference on Electronics and Communication Engineering (ICECE), Xi'an (2018)
36. Gerami, M., Xiao, M., Fischione, C., Skoglund, M.: Decentralized minimum-cost repair for distributed storage systems. In: 2013 IEEE International Conference on Communications (ICC), pp. 1910–1914, Budapest (2013)
37. Li, K., Gu, S., Wang, Y., Zhang, Q., Xiang, W.: Repair bandwidth cost of generalized regenerating codes for clustered distributed storage. In: 2019 11th International Conference on Wireless Communications and Signal Processing (WCSP), pp. 1–6 Xi'an (2019)
38. Chen, Y., Jain, S., Adhikari, V.K., Zhang, Z., Xu, K.: A first look at inter-data center traffic characteristics via Yahoo! datasets. In: 2011 Proceedings IEEE INFOCOM, pp. 1620–1628, Shanghai (2011)
39. Nygren, E., Sitaraman, R.K., Sun, J.: The Akamai network: A platform for high-performance internet applications. ACM SIGOPS Operating Syst. Rev. 44(3), 2–19 (2010)

How to cite this article: Gu, S., Wang, F., Zhang, Q., Huang, T., Xiang, W.: Global repair bandwidth cost optimization of generalized regenerating codes in clustered distributed storagesystems. IET Commun. 15:2469–2481 (2021). https://doi.org/10.1049/cmu2.12289