

무선 분산 저장 시스템을 위한 효율적인 브로드캐스트 복구 기법

김정현*

Efficient Broadcast Repair for Wireless Distributed Storage Systems

Junghyun Kim*

요약

본 논문에서는 무선 분산 저장 시스템에서 패킷 소실에 대처할 수 있는 효율적인 브로드캐스트 기반 복구 방식을 제안한다.

Key Words : Cloud storage services, Distributed storage systems, Broadcast repair, Retransmission, Packet erasure channel

ABSTRACT

In this paper, we propose an efficient broadcast repair scheme to cope with packet loss in wireless distributed storage systems.

1. 서론

클라우드 서비스가 대중화됨에 따라 방대한 양의 데이터를 수많은 서버들에 분산 저장하고 관리하는 분산 저장 시스템이 주목받고 있다. 클라우드에서 요구하는 분산 저장 시스템은 단순히 데이터를 분산시켜 저장하는 기능을 제공하는 것뿐만 아니라 하드웨어의 장애에 유연하게 대처하여 서비스가 중단되지 않도록 하고 적절한 병렬 처리를 통해 서비스가 요구하는 성능도 만족시킬 수 있어야 한다¹⁾. 이러한 요구 사항을 만족하기 위하여 빈번하게 발생하는 장애를

빠르게 복구하는 다양한 장애 복구 기법들이 제안되었다²⁻⁴⁾.

특히 최근에는 무선 장비의 사용 증가와 센서 네트워크의 발전에 힘입어 무선 분산 저장 시스템이 새롭게 주목받고 있다. 무선 분산 저장 시스템에서는 무선 통신 채널에서의 브로드캐스트 특성을 활용하여 기존의 유선으로 구축된 분산 저장 시스템보다 더욱 효율적인 장애 복구가 가능하다⁵⁻⁷⁾. 그러나 기존의 브로드캐스트 기반 복구 방식은 도움 노드가 진입 노드에게 한 번의 전송만으로 복구에 필요한 패킷을 오류 없이 전송 가능하다는 가정 하에서 동작한다. 따라서 본 논문에서는 이러한 기존 제안 기법의 한계점을 극복하고자 재전송을 고려한 브로드캐스트 기반 효율적 복구 방식을 제안한다.

II. 브로드캐스트 복구 기법

그림 1에 표현된 바와 같이 4개의 저장 노드로 구성된 서버에 4개의 데이터 패킷 A_1, A_2, B_1, B_2 을 분산 저장하는 시스템을 가정하자. 분산 저장 시스템은 임의의 2개의 노드에서 동시에 문제가 발생하더라도 복구가 가능하도록 선형 연산을 적용한 복사본을 서버의 여러 노드에 분산하여 저장한다. 만일 해당 서버의 2번째와 4번째 저장 노드에 문제가 발생하여 복구가 필요한 경우 1번째와 3번째 저장 노드가 도움 노드로 동작한다. 도움 노드들은 보유하고 있는 패킷을 수신 노드를 특정하지 않고 브로드캐스트한다. 복구를 위해 투입된 2개의 진입 노드들은 브로드캐스트된 패킷을 수신한 후 수신한 패킷들에 선형 연산을 적용하여 문제가 발생한 노드들에 저장되어 있던 패킷들을 재생성하여 저장한다.

브로드캐스트 복구 기법을 사용할 수 없는 유선 분산 저장 시스템의 경우, 그림 1과 같은 상황에서 4개의 패킷을 두 진입 노드들에게 각각 전달해야 하므로 8개의 패킷 전송이 요구된다. 즉, 2배의 통신 자원이 필요하게 된다. 무선 분산 저장 시스템의 브로드캐스트 복구 기법은 통신 자원을 효율적으로 사용할 수 있다는 커다란 장점에도 불구하고 채널의 상태에 따라 수신 실패가 빈번히 발생할 수 있다는 한계가 존재한다. 따라서 이에 대한 대책이 추가적으로 필요하다.

* 이 성과는 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 연구임(No. 2019R1G1A110000212)

* First and Corresponding Author : (ORCID:0000-0003-0265-5169) Soonchunhyang University, Department of Big Data Engineering, kimjh@sch.ac.kr, 조교수, 종신회원

논문번호 : 202008-208-A-LU, Received August 24, 2020; Revised September 1, 2020; Accepted September 2, 2020

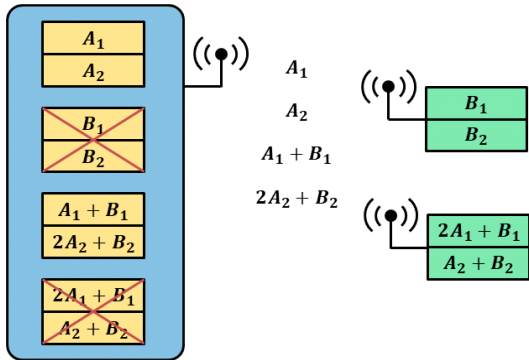


그림 1. 브로드캐스트 복구 기법.
Fig. 1. An example of broadcast repair.

III. 제안하는 브로드캐스트 복구 기법

본 논문에서는 앞서 소개한 브로드캐스트 복구 상황에서 수신 실패가 발생하는 경우를 고려한 효율적인 브로드캐스트 복구 기법을 제안한다. 재전송 여부는 일정 시간 동안 패킷 전송 후 수신에 실패한 진입 노드들로부터 피드백을 수신하여 재전송 패킷을 다시 생성한 후 재전송을 수행한다. 한 번의 재전송 주기 동안 최적의 재전송 패킷 조합을 생성하기 위하여 다음과 같이 재전송 행렬을 정의한다.

$$R = \begin{bmatrix} r_{1,1} & \cdots & r_{1,|P|} \\ \vdots & \ddots & \vdots \\ r_{|N|,1} & \cdots & r_{|N|,|P|} \end{bmatrix} \quad (1)$$

여기서 N 은 진입 노드의 인덱스 집합, P 는 복구를 위해 이전에 전송된 패킷의 인덱스 집합이다. 그림 1의 예에서 모든 진입 노드가 수신에 실패한 경우 $N=2$ 이고 $P=4$ 이다. 재전송 행렬 R 의 각 원소는 재전송이 필요한 경우 1로, 필요하지 않은 경우 0으로 설정한다. 예를 들어 2번째 노드가 3번째 패킷에 대하여 수신 실패하여 재전송이 필요한 경우 $r_{2,3} = 1$ 이다.

무선 분산 저장 서버는 재전송 행렬을 기반으로 재전송 패킷 조합을 생성하고 순서대로 전송한다. 이때 진입 노드들이 수신한 재전송 패킷으로부터 원하는 패킷을 최대한 많이 생성해 낼 수 있도록 기존 전송 패킷들의 이진 연산을 통해 최적의 재전송 패킷을 생성한다. 최적의 재전송 패킷을 생성하는 과정은 알고리즘 1에 표현되어 있다.

진입 노드들로부터 수신한 피드백을 통해 생성된 재전송 행렬 R 을 입력으로 알고리즘이 시작된다. 재전송 패킷 조합에서 제외할 패킷 인덱스 집합 \bar{P} 를 공

알고리즘 1. 효율적인 재전송 패킷 생성 기법

```

1: Input:  $R \in \mathbb{F}_2^{N \times P}$ 
2:  $\bar{P} \leftarrow \emptyset$ 
3:  $k \leftarrow 0$ 
4: for  $p \in P - \bar{P}$  do
5:    $k \leftarrow k + 1$ 
6:    $c_k \leftarrow b_p$ 
7:    $\hat{P} \leftarrow \{p\}$ 
8:   for  $p' \in P - \hat{P}$  do
9:      $s \leftarrow 0$ 
10:    for  $p'' \in \hat{P}$  do
11:      for  $n \in N$  do
12:         $s \leftarrow s + (r_{n,p'} \times r_{n,p''})$ 
13:      if  $s = 0$  then
14:         $c_k \leftarrow c_k \oplus b_{p'}$ 
15:         $\hat{P} \leftarrow \hat{P} \cup \{p'\}$ 
16:       $\bar{P} \leftarrow \bar{P} \cup \hat{P}$ 
17: return  $\mathbf{c} = \{c_1, \dots, c_k\}$ 
    
```

집합으로 설정하고 재전송 패킷 인덱스 변수 k 를 0으로 초기화한다. 집합 $P - \bar{P}$ 이 공집합이 될 때까지 다음을 반복한다. 집합 $P - \bar{P}$ 에서 원소 p 를 추출하고 k 를 1만큼 증가시킨다. 재전송 패킷 c_k 를 기존 전송 패킷 b_p 로 설정한다. 재전송 패킷 c_k 에 포함된 패킷 인덱스 집합 \hat{P} 을 $\{p\}$ 로 설정한다. 집합 $P - \hat{P}$ 에서 원소 p' 을 추출하고 판별값 s 를 0으로 설정한다. 선택된 p' 번째 패킷이 현재까지 재전송 패킷에 포함된 모든 패킷에 대해 대응되는 재전송 요청 노드가 겹치지 않으면 $s=0$ 이 된다. 이러한 경우 p' 번째 패킷을 재전송 패킷 c_k 에 이진 연산으로 더하여 업데이트한다. 또한 패킷 인덱스 집합 \hat{P} 에 p' 을 추가한다. 집합 $P - \hat{P}$ 가 공집합이 되면 \bar{P} 에 \hat{P} 을 추가한다. 집합 $P - \bar{P}$ 에서 다음 패킷을 선택하고 위 과정을 반복한다. 집합 $P - \bar{P}$ 가 공집합이 되어 더 이상 선택할 패킷이 없으면 생성된 재전송 패킷 집합 \mathbf{c} 를 반환하고 알고리즘이 종료된다.

그림 2는 알고리즘 1에 의해 재전송 패킷을 생성한 예이다. 이진 연산을 통해 전송 횟수를 최소화하면서 노드별 소실 패킷에 대한 재전송 패킷 수신 횟수를 최대화함을 확인할 수 있다. 예를 들어 3번째 노드는 c_5 를 3번, 4번째 노드는 c_6 를 4번 수신할 수 있는 기회가 주어진다. 따라서 패킷 소실 채널 환경에서 최소의 통신 자원을 사용하여 패킷 복구 성공률을 향상시킨다.

기존의 브로드캐스트 복구 기법에서는 수신 실패 상황을 고려하지 않아 직접 비교가 불가능하기 때문에 무선 통신 시스템에서 널리 사용되는 재전송

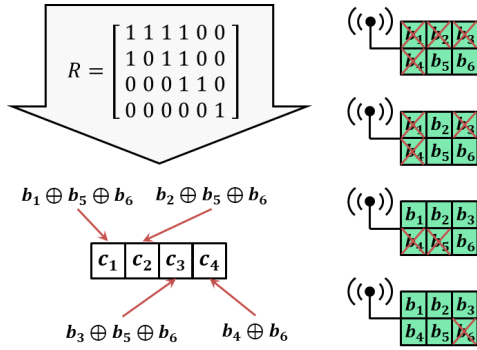


그림 2. 제안된 브로드캐스트 복구를 위한 재전송 패킷 생성.
Fig. 2. An example of retransmission packet generation for the proposed broadcast repair.

(automatic repeat request; ARQ) 기법⁸⁾과 선형 연산을 통해 재전송 횟수를 최소화하는 부호 기반 재전송 기법⁹⁾을 브로드캐스트 복구 기법에 적용하여 비교하였다.

그림 3은 재전송 기법을 브로드캐스트 복구에 적용한 예이다. 비교를 위해 그림 2와 동일한 수신 실패 상황을 가정하였다. 이 경우 진입 노드들이 수신에 실패한 패킷을 다시 전송받기 위하여 최소 6개의 패킷이 재전송되어야 한다. 즉, 제안된 브로드캐스트 복구 기법보다 더 많은 통신 자원을 사용하게 된다.

그림 4는 부호 기반 재전송 기법을 브로드캐스트 복구에 적용한 예이다. 비교를 위해 그림 2와 동일한 수신 실패 상황을 가정하였다. 이 경우 진입 노드들이 수신에 실패한 패킷을 다시 전송받기 위해 필요한 최소 재전송 패킷 수는 4개로 제안된 브로드캐스트 복구 기법과 동일하다. 그러나 노드별 소실 패킷에 대한 재전송 패킷 수신 횟수를 살펴보면 모든 노드가 각 패킷에 대하여 단 한 번의 수신 기회만 갖는다. 따라서

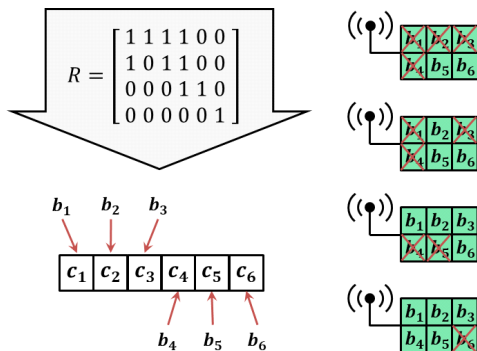


그림 3. 재전송 기반 브로드캐스트 복구를 위한 재전송 패킷 생성.
Fig. 3. An example of retransmission packet generation for the ARQ-based broadcast repair.

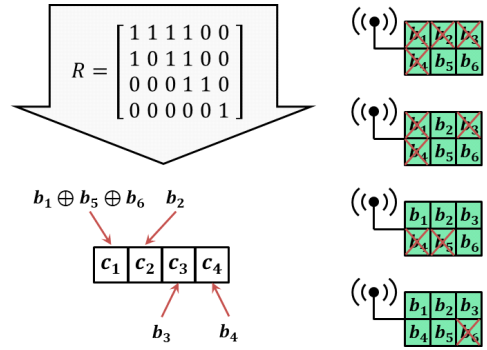


그림 4. 부호 기반 브로드캐스트 복구를 위한 재전송 패킷 생성.
Fig. 4. An example of retransmission packet generation for the code-based broadcast repair.

부호 기반 브로드캐스트 복구 기법은 제안된 브로드캐스트 복구 기법과 동일하게 최소의 통신 자원을 사용하지만 패킷 복구 성공률은 상대적으로 저하됨을 알 수 있다.

IV. 성능검증

모의실험 환경은 무선 분산 저장 서버에서 일정한 개수의 노드에 문제가 발생하여 복구가 필요한 상황을 가정하였다. 또한 복구를 위하여 필요한 패킷을 브로드캐스트 방식으로 전송한 뒤 일정 확률로 수신 실패가 발생한다. 재전송을 포함한 총 전송 횟수를 최대 전송 횟수 2회로 제한하였다.

성능 비교 대상으로 앞서 소개한 재전송 기반 브로드캐스트 복구 기법과 부호 기반 브로드캐스트 복구 기법을 고려하였다. 이하 논문에서 두 비교 기법과 제안하는 기법을 각각 A-BR, C-BR, P-BR로 표시한다. 성능지표는 최대 전송 횟수만큼 패킷 전송 후 모든 진입 노드가 패킷 복구에 성공한 비율을 사용하였다.

그림 5는 복구에 필요한 노드를 6개, 패킷을 12개로 고정하고 무선 채널에서 패킷 손실 확률을 변화시키면서 복구 성공 비율을 비교한 결과이다. 전 범위에서 제안한 P-BR 기법이 가장 좋은 성능을 보였으며 패킷 손실 확률이 높아질수록 비교 기법은 급격한 성능 저하를 보였다.

그림 6은 동일한 환경에서 패킷 손실 확률을 0.1로 고정하고 노드의 수를 변화시키면서 복구 성공 비율을 비교한 결과이다. 전 범위에서 제안한 P-BR 기법이 가장 좋은 성능을 보였으며 노드 수가 증가할수록 성능 격차가 커짐을 확인할 수 있다.

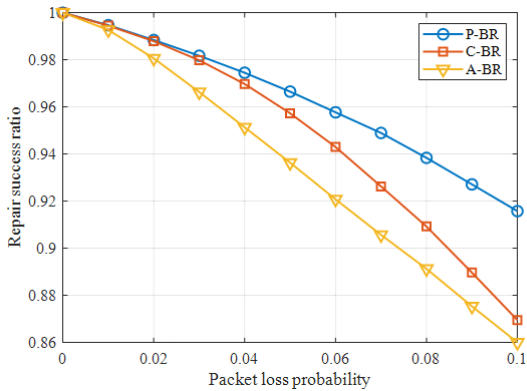


그림 5. 패킷 손실 확률에 따른 복구 성공 비율.
Fig. 5. Repair success ratio versus packet loss probability.

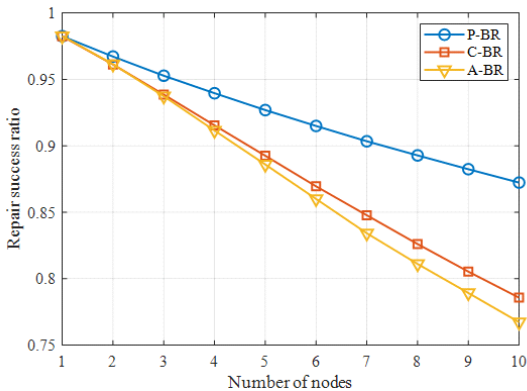


그림 6. 노드 수에 따른 복구 성공 비율.
Fig. 6. Repair success ratio versus number of nodes.

V. 결 론

본 논문에서는 무선 분산 저장 시스템에서 패킷 손실에 대처하기 위하여 재전송 기반의 효율적인 브로드캐스트 복구 방식을 제안하였다. 모의실험을 통해 제안한 방식이 기존의 재전송 기법들을 브로드캐스트 복구 방식에 적용한 것보다 탁월한 성능을 가짐을 보였다. 따라서 제안한 브로드캐스트 복구 방식은 향후 다양한 무선 분산 저장 시스템에 적용될 것으로 기대된다.

References

- [1] A. G. Dimakis, P. B. Godfrey, Y. Wu, M. J. Wainwright, and K. Ramchandran, "Network coding for distributed storage systems," *IEEE Trans. Inf. Theory*, vol. 56, no. 9, pp. 4539-4551, Sep. 2010.
- [2] M.-Y. Nam and H.-Y. Song, "Binary locally repairable codes with minimum distance at least 6 based on partial t-Spreads," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1683-1686, Aug. 2017.
- [3] J.-H. Kim and H.-Y. Song, "Alphabet-dependent bounds for locally repairable codes with joint information availability," *IEEE Commun. Lett.*, vol. 21, no. 8, pp. 1687-1690, Aug. 2017.
- [4] J.-H. Kim and H.-Y. Song, "Hypergraph-based binary locally repairable codes with availability," *IEEE Commun. Lett.*, vol. 21, no. 11, pp. 2332-2335, Nov. 2017.
- [5] A. S. Rawat, O. O. Koyluoglu, and S. Vishwanath, "Centralized repair of multiple node failures with applications to communication efficient secret sharing," *IEEE Trans. Inf. Theory*, vol. 64, no. 12, pp. 7529-7550, Dec. 2018.
- [6] P. Hu, C. W. Sung, and T. H. Chan, "Capacity of wireless distributed storage systems with broadcast repair," *IEEE Trans. Commun.*, vol. 67, no. 2, pp. 1004-1017, Feb. 2019.
- [7] M. Zorgui and Z. Wang, "Centralized multi-node repair regenerating codes," *IEEE Trans. Inf. Theory*, vol. 65, no. 7, pp. 4180-4206, Jul. 2019.
- [8] A. Goldsmith, *Wireless Communications*, Cambridge University Press, 2005.
- [9] C. W. Sung, K. W. Shum, L. Huang, and H. Y. Kwan, "Linear network coding for erasure broadcast channel with feedback: Complexity and algorithms," *IEEE Trans. Inf. Theory*, vol. 62, no. 5, pp. 2493-2503, May 2016.