저장 효율적인 하이퍼레저 패브릭 블록체인을 위한 소거 코드 기반 분산 저장 시스템

박소현 [1,2], 최병준 [1], 김창수 [1], 이명철 [1], 이일구 [2]
한국전자통신연구원 [1], 성신여자대학교 [2]

{sohyun.park, bjchoi92, cskim7, mclee}@etri.re.kr, iglee@sungshin.ac.kr

Erasure Code-based Distributed Storage System for Storage-Efficient Hyperledger Fabric Blockchain

Sohyun Park[1,2], Beongjun Choi[1], Changsoo Kim[1], Myungcheol Lee[1], Il-Gu Lee[2]
Electronics and Telecommunications Research Institute[1], Sungshin Women's Univ [2]

요 약

하이퍼레저 패브릭은 데이터 경량 저장을 지원하는 비트코인이나 이더리움과 다르게 모든 노드가 전체 데이터를 저장하는 풀 노드이기 때문에 트랜잭션 데이터의 분산 저장을 통한 저장 공간의 경량화가 요구된다. 본 연구에서는 소거 코드(Erasure Codes)의 일종인 Reed-Solomon(RS) 코드를 기반으로 하이퍼레저 패브릭 원장을 분산 저장하여 메모리 사용 효율을 개선하고 데이터 복구 장애 허용 가능한 방법을 제안한다. MDS(Maximum Distance Separable) 및 시스터메틱 부호화 특성을 갖는 RS 코드를 이용하여 트랜잭션 데이터를 인코딩 및 분산 저장하면, 부호화된 데이터를 디코딩하지 않고 접근할 수 있으며, (10,6) RS 코드의 경우 저장 공간 부하를 기존 중복 저장 방식 대비 75% 줄일 수 있다. 또한, RS 코드는 패리티 개수만큼의 손실이 발생해도 블록을 복구할 수 있어 BFT(Byzantine Fault Tolerance)를 보장하도록 분산 저장 시스템을 설계할 수 있다.

Ⅰ. 서 론

블록체인은 네트워크에 참여하는 모든 노드가 동일한 데이터를 중복 저장하여 데이터 무결성과 투명성을 보장하는 분산 원장 기술이며, 블록체인에 기록되는 데이터는 임의로 삭제되거나 변경될 수 없다. 블록체인의 이러한 비가역적이고 중복하여 분산 저장하는 특징으로 인하여 블록체인 네트워크에 참여하는 노드의 저장 공간 문제가 대두되고 있다 [1].

블록체인 저장 공간 문제를 해결하기 위하여 라이트 노드, 샤딩(Sharding), 프루닝(Pruning) 등의 방법을 사용할 수 있지만, 최근에는 소거 코드(Erasure Codes, EC) 기반의 데이터 분산 저장 연구가 활발히 진행되고 있다 [2]. 소거 코드는 데이터에 패리티를 추가하여 데이터의 일부가 손실되는 상황에서 원본 데이터를 복원할 수 있도록 하는 기술이다. 블록체인 트랜잭션 데이터에 소거 코드를 적용하여 부호화하고 참여 노드들은 부호화된 데이터의 일부만 저장함으로써 노드가 저장해야 하는 데이터의 용량을 크게 줄이면서도 비잔틴 장애 내성(Byzantine Fault Tolerance, BFT)을 갖도록 설계할 수 있다. 이러한 이유로 소거 코드 부호화 기반의 분산 저장 방식은 기존 방식에 비해 블록체인의 분산성, 무결성, 보안성을 동시에 만족할 수 있다는 장점이 있다.

하이퍼레저 패브릭(Hyperledger Fabric)은 리눅스 파운데이션(Linux Foundation)에서 개발한 허가형 블록체인을 위한 오픈소스 분산 운영 시스템이다 [3]. 의료, 금융 등 데이터 보안이 중요한 산업 분야에서는 허가형 블록체인을 통해 안전하게 데이터를 공유하고 저장할 수 있다. 그러나, 하이퍼레저 패브릭은 모든 참여 노드가 풀 노드로서 동일한 크기의 원장을 중복저장해야 하기 때문에, 각 노드가 용량이 매우 큰 전체 트랜잭션 데이터를 모두 저장해야 한다는 문제점이 있다. 본 연구에서는 블록체인 네트워크 참여 노드의 트랜잭션 데이터 저장 공간 절약을 위하여, 소거 코드 기반 하이퍼레저 패브릭 원장 분산 저장 시스템을 설계하였다.

Ⅱ. 본론

하이퍼레저 패브릭의 원장은 그림 1 과 같이, 원장의 최신 상태에 빠르게 접근하기 위한 world state 데이터베이스와 실제 트랜잭션 데이터가 저장되는 블록체인으로 구성된다. World state 가 저장되는 상태 데이터베이스는 가장 최신의 데이터만을 기록하기 때문에 과거의 블록체인 트랜잭션 데이터부터 거슬러 올라가 계산하지 않고도 바로 최신 상태를 얻을 수 있다. 블록체인은 최초부터 현재까지의 트랜잭션 로그를 저장하고 있으며, 트랜잭션 로그는 변경 및 삭제가 불가능하기 때문에 트랜잭션이 발생함에 따라 저장해야 하는 트랜잭션 데이터의 양도 증가한다.

하이퍼레저 패브릭의 모든 노드는 블록을 자신의 파일 시스템에 블록파일(blockfile)의 형태로 저장한다. 하나의

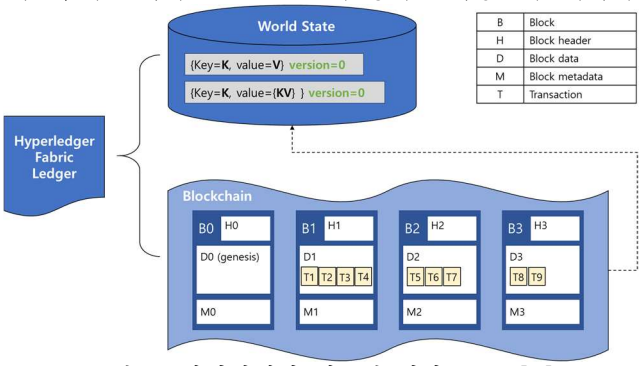

그림 1. 하이퍼레저 패브릭 원장 구조 [4]

블록파일 내에 여러 개의 블록이 저장되고, 설정한 블록파일의 최대 크기가 넘어가게 되면 현재의 블록파일을 닫고 다음 블록파일에 블록을 기록하게 된다. 모든 노드에 중복 저장되는 트랜잭션 데이터의 용량을 줄이기 위하여, 블록파일 내 블록 단위로 소거 코드를 적용하여 블록을 분산 저장할 수 있다.

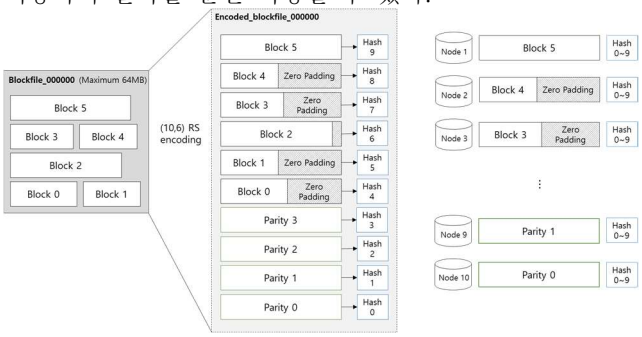

그림 2. 소거 코드 기반 하이퍼레저 패브릭 원장 분산 저장 방법

그림 2 는 소거 코드 기반 하이퍼레저 패브릭 원장 분산 저장 방법을 나타낸다. 대표적인 소거 코드 중 하나인 Reed-Solomon (RS) 코드를 사용하여 트랜잭션 데이터 인코딩을 구현할 수 있다. (n,k) RS 코드는 k개의 원본 데이터에 $n-k$개의 패리티를 추가하여 n개의 인코딩 청크를 생성한다. n 개의 노드는 각자 인코딩을 수행하여 n개의 인코딩 청크를 생성한 후 자신이 저장할 청크만을 남기고 삭제하여 저장 공간을 절약할 수 있다. 그림 2 의 예시에서는 블록파일(blockfile_000000)의 최대 크기가 64MB 이고 총 6 개의 블록이 저장되어 있다. 블록파일 내 블록의 크기는 각기 다르기 때문에, RS 인코딩을 위해 각 블록에 제로 패딩을 추가하여 모든 블록의 크기를 동일하게 만드는 전처리 과정을 수행하였다. 그림 2 에서는 가장 크기가 큰 Block 5 를 기준으로 제로 패딩을 추가했다. 이후, (10,6) RS 코드를 적용하여 인코딩을 수행하면 총 10 개의 인코딩 청크가 생성된다. 10 개의 인코딩 청크는 10 개의 노드에 하나씩 분산 저장된다.

그림 2 에서 보는 바와 같이, 인코딩된 데이터에 원본 데이터가 포함되도록 부호화하는 것을 시스터메틱 부호화(Systematic Coding)라고 한다. 시스터메틱 코드를 사용하면 디코딩을 하지 않아도 원본 데이터에 빠르게 접근할 수 있다. 자신이 가지고 있지 않은 특정 블록이 필요한 경우, 해당 블록을 가지고 있는 노드에 블록을 요청하여 원하는 블록에 빠르게 접근할 수 있다. 만약 블록이 손실되어 접근이 어려운 경우에는 전체 노드에 인코딩 청크를 요청하고, 그 중 k 개의 청크로 원본 데이터를 복구하여 필요한 블록에 접근할 수 있다. 모든 노드는 전체 블록 및 인코딩 청크의 해시 값을 계산하여 저장하기 때문에, 전송받은 블록의 무결성을 빠르게 검증할 수 있다.

[표 1] (10,6) RS 코드 적용 시, 저장 방식에 따른 저장 공간 효율성 비교

	중복 저장 방식	분산 저장 방식
총 저장 공간 사용량	64 MB × 10 = 640 MB	16 MB × 10 = 160 MB
저장 공간 부하	640 MB/64 MB = 10	160 MB/64 MB = 2.5

표 1 은 기존의 트랜잭션 중복 저장 방식과 그림 2 의 (10,6) RS 코드 기반 분산 저장 방식의 저장 공간 효율성을 비교한다. 그림 2 의 가장 큰 블록인 block5 가 16 MB 라고 가정하고 해시 값은 고려하지 않았을 때, 전체 노드가 부담해야 하는 저장 공간 사용량이 줄어들어, 동일한 64 MB 의 데이터를 저장할 때 저장공간 부하를 기존 저장 방식 대비 75% 줄일 수 있다.

RS 코드는 MDS (Maximum Distance Separable) 코드의 한 종류이기 때문에, n 개의 인코딩 청크 중 패리티 개수만큼의 손실에 대해서도 오류 정정을 통한 원본 데이터 복구가 가능하다. 따라서, 패리티 개수만큼의 비잔틴 노드에 대한 복구 능력을 갖기 때문에 BFT 를 만족할 수 있어 블록체인 네트워크에서도 사용이 가능하도록 설계할 수 있다. 하이퍼레저 패브릭의 경우에는 합의 알고리즘으로 PBFT (Practical Byzantine Fault Tolerance)를 사용하는데, 비잔틴 노드가 f개일 때 전체 노드가 $3f+1$ 개 이상인 경우 BFT 를 만족한다. 그림 2 의 (10,6) RS 코드의 경우 10 개의 인코딩 청크 중 최대 4 개의 손실이 발생하더라도, 나머지 인코딩 청크 6 개의 조합으로 원본 블록을 생성할 수 있기 때문에 BFT 를 보장한다. 패리티가 증가하면 많은 손실에도 오류 정정이 가능하지만 더 많은 데이터를 저장해야 하기 때문에 저장 공간 효율성이 떨어지므로, 적용 환경의 가용성 특성에 맞게 적절한 부호율을 설계하는 것이 중요하다.

Ⅲ. 결론

본 연구에서는 하이퍼레저 패브릭 참여 노드의 데이터 저장 공간을 절약하고 데이터 복구 장애 허용 가능한 분산 저장 시스템을 설계하기 위해 소거 코드를 블록체인에 적용했다. (10,6) RS 코드를 적용하여 분산 저장했을 때, 기존의 중복 저장 방식 대비 저장 공간 부하를 75% 줄이면서 BFT 를 보장할 수 있다. RS 코드를 이용하여 트랜잭션 데이터를 인코딩 및 분산 저장하면 일부 인코딩 청크만으로 원본 블록을 다시 복구할 수 있다. 하지만, 데이터 복구를 위한 디코딩을 하기 위해서 많은 노드로부터 인코딩 청크를 받아와야 하기 때문에 복구 비용이 증가한다는 단점이 있다. 향후 연구 과제로는 복구 비용을 줄이고 고속 복구가 가능한 트랜잭션 데이터 분산 저장 기술을 개발하고자 한다.

ACKNOWLEDGMENT

이 논문은 2021 년도 정부(과학기술정보통신부)의 재원으로 정보통신기획평가원의 지원을 받아 수행된 연구임 (No. 2021-0-00136, 다양한 산업 분야 활용성 증대를 위한 대규모/대용량 블록체인 데이터 고확장성 분산 저장 기술 개발).

참 고 문 헌

[1] Xie, J. et al. "A survey on the scalability of blockchain systems," IEEE Network, pp. 166-173, 2019.

[2] Qi, X et al. "BFT-Store: Storage partition for permissioned blockchain via erasure coding," 2020 IEEE 36th International Conference on Data Engineering (ICDE), pp. 1926-1929, April. 2020.

[3] Androulaki, E. et al. "Hyperledger fabric: a distributed operating system for permissioned blockchains," Proceedings of the thirteenth EuroSys conference, pp. 1-15, April. 2018.

[4] Hyperledger Fabric, “Ledger,” 2020, (https://hyperledger-fabric.readthedocs.io/en/latest/ledger/ledger.html).