이더리움 기반 블록체인 확장성 연구

나지원*, 추민지*, 민연아*, 백영태O
가천대학교 소프트웨어학과*
김포대학교 멀티미디어학과O
e-mail: poquwk@gmail.com*, minchu9611@gmail.com*, yah0612@gachon.ac.kr*, hannae@kimpo.ac.krO

A Study on Ethereum based block-chain scalability

Ji-won Na*, Min-ji Choo*, Min Youn-A*, Baek Yeong-TaeO
Dept. of Software, Korea-Gachon University*
Dept. of Multimedia Kimpo UniversityO

● 요 약 ●

본 논문에서는 이더리움 기반 블록체인 확장성 문제해결을 위해 데이터 저장방법인 Merkle patricia tree를 응용하여 데이터에 따라 트리 이원화 사용을 제안한다. 이 연구는 시스템 자원인 CPU와 메모리를 효율적으로 사용하여 트랜잭션 처리량을 최대화하고, 작업시간을 최소화하기 위함이다. 본 논문에서는 기존 이더리움 블록체인의 트랜잭션 처리속도와 확장성 향상을 목표로 하며. 기존의 방식과 비교하여 제안을 분석한다.

키워드: 머클패트리샤트리(Merkle patricia tree), 확장성(scalability), 이더리움(ethereum), B+ tree

I. Introduction

4차 산업혁명 시대에 들어오면서 의료, 금융, 소셜 등 다양한 분야의 데이터를 디지털화 시키는 빅데이터가 주목받고 있다. 유출에 민감한 데이터들이 다수 포함되어 있어 데이터를 안전하게 지킬 수 있는 보안 기술을 찾고 있다. 그 중에서도 빅데이터와 블록체인을 결합하여 보안성을 높일 수 있도록 하는 연구가 활발하게 진행되고 있는 추세이다.

하지만 데이터의 크기가 점점 거대해지고 있기 때문에 모든 데이터를 블록체인에 결합시키기엔 블록의 용량이 너무 커지고 있다. 그에 따른 수수료 또한 높아지게 된다는 단점이 생긴다. 또한 블록체인은 데이터의 양이 많아질수록 속도가 느려진다는 문제점이 있다. 이런 확장성의 문제는 블록체인의 고질적인 단점으로, 이를 극복하기 위한 다양한 해결책이 나오고 있다.

따라서, 본 연구는 실제로 사용되고 있는 블록체인의 확장성을 높일 수 있는 방법들과 장단점에 대해 알아보고, 그 중에서도 이더리움 머클패트리샤트리 구조를 응용하여 블록체인의 확장성을 높일 수 있도록 하는 방안을 제시하고자 한다.

II. Preliminaries

1. Related works

1.1 Segregated Witness(SegWit)

SegWit은 비트코인 전용으로 확장성과 TPS(초당 거래량)을 높이기 위한 방법이다. 비트코인의 트랜잭션은 입력과 출력으로 구성되어 있는데, 입력에는 이전 거래의 세부 정보와 개인키(서명)가 포함되어 있다. 입력은 트랜잭션 전체의 약 60~70% 가량의 큰 용량을 차지하는데, 거래 정보는 트랜잭션에 필수 요소이지만 서명은 트랜잭션의 유효성 검사 시에만 필요한 요소이다. 트랜잭션에서 서명이 차지하는 부분을 없애고 거래 정보로 채우면 꽤 많은 거래 정보를 담을 수 있기때문에 거래 정보와 서명을 분리하게 되었는데 이를 SegWit이라고 한다. 즉 서명을 입력에서 제외하여 하나의 블록의 크기를 1MB고정시키고, 별도의 증인 필드를 만들어 서명을 포함시킨다. 증인 필드는 트랜젝션의 끝부분에 포함시킨다.

1.2 블록 크기의 확장

말 그대로 블록의 크기를 확장하여 하나의 블록에 들어가는 트랜잭션의 개수를 늘리는 방법이다. 가장 단순하고 구현이 간단한 방법이다. 하지만 이는 단위 시간 당 더 많은 트랜잭션을 처리해야 한다는 것을 의미하기 때문에 비용적인 부담으로 인해, 중앙화가 될 수 있다는 단점이 있다.

1.3 Sharding

Sharding은 네트워크를 샤드(Shard) 단위로 나누어서 블록체인의 데이터를 분산 저장하여, 하나의 노드가 검증해야 하는 데이터를 분할하는 On-chain방식이다. 기존의 이더리움 검증 방식은 모든 노드가 전체 블록을 저장하고 트랜잭션을 처리해야 하기때문에 속도가 느려진다.

하지만 Sharding 방식을 사용하면 샤드 단위로 블록체인 데이터를 나누어 저장하고 트랜잭션을 처리하기 때문에 하나의 노드가 검증해야 하는 데이터의 양이 줄어들어 속도가 전체적으로 향상된다.

III. The Proposed Scheme

이더리움의 데이터는 Merkle patricia tree 형태로 저장되는데, 이는 Merkle tree와 Patricia tree의 결합형태로 Key-value 맵핑과 각 노드가 hash 값을 보유하는 특징을 가지고 있다. Merkle patricia tree는 extension node, branch node, leaf node 3가지 종류의 노드로 구성되어 있는데, 본 논문에서는 B+ tree와 접목시켜 저장되는 데이터의 종류나 상황에 따라 사용되는 트리의 이원화를 제안하여 branch node의 공간 효율성과 데이터 검색 효율성을 증대시키고자 한다.

branch node는 이더리움의 데이터가 16진수로 이루어져있기 때문에 한번 생성이 되면 value값을 저장할 공간을 포함해 17개 값을 저장할 수 있는 배열이 생성된다. 그러나 이는 공통된 키 값을 가진 데이터가 있으나 데이터의 개수 자체가 적을 때 공간을 낭비하는 상황으로 이어진다. 또한 leaf node 간의 연관성이 없기 때문에 기본적인 탐색에는 관계없으나 순차 탐색에는 효율적이지 않은 형태를 가지고 있다. 하지만 이러한 단점은 B+tree와 접목시키면 상쇄가 가능하다.

기존 B+tree에서 접목시킬 부분은 크게 두가지이다. leaf node들끼리 순차성을 유지하며 연결리스트 형태로 이어져 있다는 점과 leaf node가 아닌 내부 노드 키 값의 개수 범위가 있다는 점이다. Merkle patricia tree에서 branch node는 17개의 배열로 구성되어 있었으나 들어오는 데이터 양이 적으면 branch node가 가질 수 있는 키 값의 개수의 범위를 지정하고, leaf node들끼리 리스트로 연결시킨 B+tree와 접목시킨 트리를 사용한다. 이렇게 되면 branch node에 있던 빈공간을 줄이고 기본 탐색부터 순차탐색까지 효율적으로 실행 가능하다. 데이터의 양이 많을 때에는 앞서 말한 트리를 사용하면 트리의 깊이가 깊어질 수 있기 때문에 기존 Merkle patricia tree를 사용한다.

IV. Conclusions

본 논문에서는 데이터 양에 따라 트리를 이원화하여 사용하는 것을 제안하고, 이더리움 블록체인의 확장성 향상에 의의를 둔다. 연구를 순차적으로 진행하며 시뮬레이션을 통해 기존 Merkle patricia tree만 사용했을 때와 성능을 비교하고, 분석을 진행할 예정이다.

REFERENCES

[1] On the scalability of Blockchain, 2018.8, FathElrahman AwadElkarim, Hanyang University Graduate school.

[2] Colin Percival. Stronger key derivation via sequential memory-hard functions. pages 1– 16, 2009.

[3] Ethereum, “https://www.ethereum.org/learn/