SELECTIVE DELETION

IN A BLOCKCHAIN

REFERENCE

- Hillmann, Peter, et al.
  "Selective Deletion in a Blockchain."
  arXiv preprint arXiv:2101.05495 (2021).

ABSTRACT

ABSTRACT

- 블록체인 크기의 지속적이고 일정한 증가가 문제
  - 특히 원치 않았던 데이터가 문제인데
  - 자연스러운 방법으로는 제거할 수도 없기 때문

ABSTRACT

- 이 문제에 대항하기 위해
  - 블록체인의 한 엔트리에 대한 선택적 제거 방법을 제시

ABSTRACT

- 이를 위해
  - 정기적으로 요약(summary) 블록들을 생성함
  - 체인의 이전 데이터가 요약되고 새 블록에 저장됨
    - 원치 않은 정보는 제거해서

ABSTRACT

- 이러한 컨셉은
  - 블록 구조
  - 네트워크 구조
  - 합의 알고리즘과 무관함
- 현존하는 블록체인에 적용할 수도 있음

ABSTRACT

- 이로부터
  - 블록체인 기술이 더 많이/잘 응용될 수 있을 것

ABSTRACT

- Luke’s comment:
  - 원 논문이 가지고 있는
  - 용량 문제를 완화한다는 관점이나
  - 신뢰 노드가 필요한 점 등은
  - ‘삭제’ 콘셉트와는 별개로 다루는 것이 맞다고 판단해 배제함
  - 실제 블록체인이 아닌 클라이언트-서버에 기반한 실험 및 평가 역시 생략
  - 궁금하신 분들은 원 논문을 참고바람

INTRODUCTION

INTRODUCTION

- 분산 원장 기술이 블록체인 기술의 등장으로부터 잘 알려져왔음
  - 비잔틴 장군 문제도 잘 풀어내고
  - 이더리움 등으로부터 이제 프로그램 구동도 가능함
    - 스마트 컨트랙트로 잘 알려짐
- 이러한 보안과 추적가능성의 강점, 제3자를 요구하지 않는 점 등으로부터
  - 블록체인은 4차산업혁명이나 서플라이 체인(supply chain) 등
  - 널리 적용 되고/되려하고 있음

INTRODUCTION

- 블록체인에서 가장 중요한 속성은
  - 데이터의 불변성에 기반한
  - 신뢰
- 데이터를 블록의 형태로 수집해
  - 해시를 통해 임의 위변조를 막음
  - 블록과 블록 안의 콘텐츠의 순서가 고정됨

INTRODUCTION

- 블록체인 기술이 계속 발전하고 있지만 다음과 같은 문제가 있음:
  - 불변성으로 인해 불법이나 원치 않은 콘텐츠가 공공 블록체인에 계속 노출
  - 블록체인이 빠르게 성장해 용량 문제가 대두
  - 법적 문제 … 지워질 권리 등
- 실제로 성적인(pornographic) 콘텐츠 등을 비트코인이 포함하고 있음

INTRODUCTION

- 이러한 문제는
  - 비트코인 트랜잭션이 비트코인과 무관한 데이터를 저장할 수 있게 하고
  - 풀노드가 전체 블록체인 데이터를 무조건 다운로드해야하는 점으로부터 문제가 됨

INTRODUCTION

- 그렇기 때문에 지워질 권리가 고려됨
  - 개인 데이터 문제에서 특히
- 그러나 이러한 문제를 어떻게 다뤄야 할 지 명확하지 않음

INTRODUCTION

- 본 논문에서는 원치 않은 데이터를 동작하는 블록체인에서 제거하는 방법을 제시
  - 블록체인에서 이 데이터를 직접 지우거나
  - 잊게 함 (forget)
  - 이를 통해 체인의 신뢰를 유지
- 이 방법은
  - 네트워크 인프라와 구분되고
  - 합의 알고리즘과도 구분됨

REQUIREMENTS

REQUIREMENTS

- 신빙성(Authenticity): 데이터 엔트리의 진위를 확인하는 작업
- 데이터 여분/잉여(Redundancy of Data): fail-safe를 위한 데이터의 중복 저장
- 요청에 따른 삭제(Delete on request): 적은 노력(effort)으로 엔트리나 블록을 선택적 제거
- 확장성(Scalability): 시스템은 유저 수나 블록체인 크기와 무관하게 실행됨

RELATED WORK

RELATED WORK

- 이미 여러 접근 방법이 제시된 바 있으며
  - 일부는 요구사항을 부분적으로 만족하곤 함

RELATED WORK

- 가장 유망한 기술은 오프체인 트랜잭션
  - 오프체인 트랜잭션이 메인체인에 기록되지는 않음
  - 메인체인과 상호작용에 비용, 시간이 많이 들어감

RELATED WORK

- 키(key)를 오프체인에서 관리하고
  - 비대칭 암호화된 데이터만 블록체인에 저장
  - 블록체인 길이를 줄이거나 데이터 크기를 축소시키지는 못함
  - 불법적 콘텐츠를 지우지 못함

RELATED WORK

- 로컬 저장분의 프루닝(pruning)
  - 글로벌 분산 블록체인의 문제를 해결하지 못함

RELATED WORK

- 새 블록체인으로의 하드포크
  - 원치 않은 데이터를 제거하고 하드포크
  - 시간이 매우 오래 소요됨

RELATED WORK

- 결국 현존하는 기술들은
  - 분산 블록체인의 삭제 불가능 문제를 해결하지는 못함
  - 요구사항을 전부 충족하지 못함

CONCEPT

CONCEPT

- 살아있는/동작하는(living) 블록체인에서 정보를 지울 수 있는 방법 소개
  - 이는 데이터베이스 관점에서
  - 블록체인을 고신뢰뿐만 아니라
  - 다른 속성들을 가질 수 있게 함

CONCEPT

- 요구사항들을 만족시키기 위해
  - 요약 블록(Summary)을 도입
  - 정기적인 인터벌로

BACKGROUND

- 블록의 개별 엔트리를 그냥 직접 삭제해버리는 것은
  - 블록체인의 해시 체인을 망가뜨림
  - 이는 모순, 붕괴, 신뢰의 손실을 발생시킴
- 블록체인에서 요청에 따른 개별 엔트리를 삭제하기 위해서는
  - 블록체인이 의도적으로 정보를 잊거나 잃어버릴 수 있어야 함
  - 그러면서 신뢰 수준을 유지해야 함

EXTENSION WITH SUMMARY BLOCKS

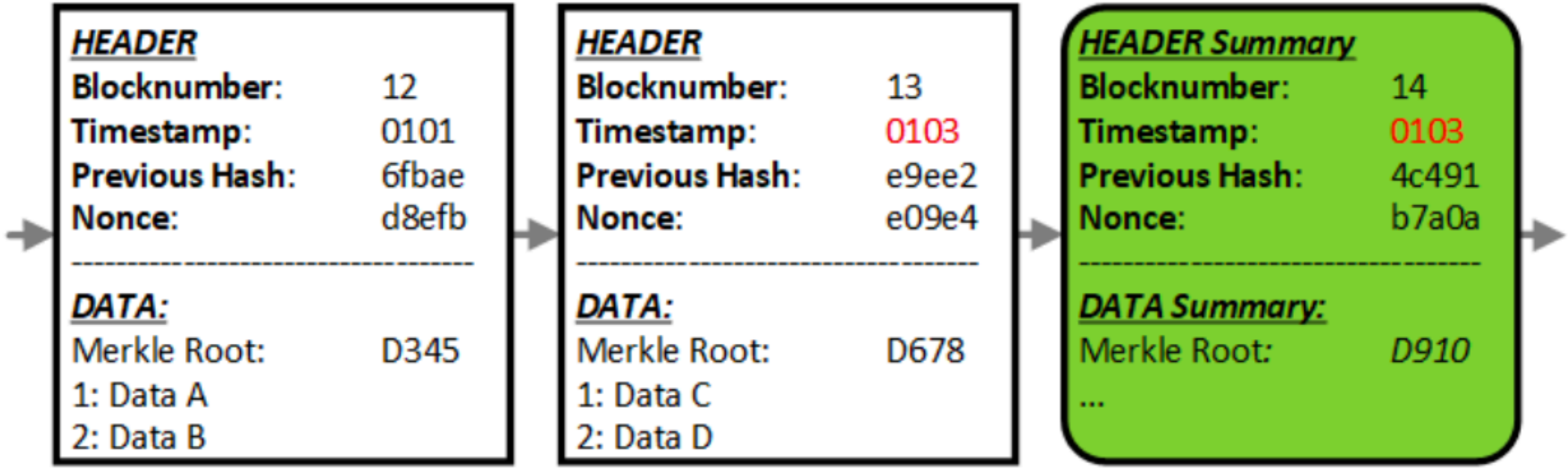
- 블록체인은 특별한 블록 타입인
  - 요약 블록 Σ으로부터 확장됨
  - 결정론적인 정보로 구성됨
  - 외부의 추가적 정보는 포함되지 않음

EXTENSION WITH SUMMARY BLOCKS

- 요약 블록은 주기적으로 형성됨
  - 가령 10 블록마다

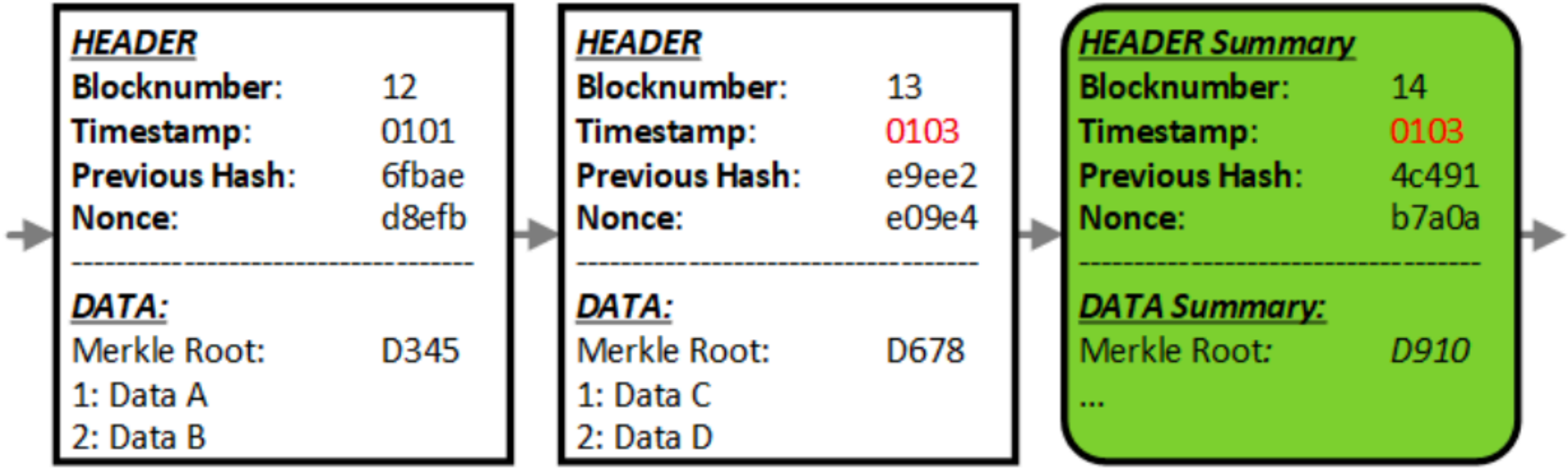
EXTENSION WITH SUMMARY BLOCKS

- 요약 블록은 블록체인상에 일반적인 블록처럼 통합됨
  - 블록 번호도 1 증가해서 등록됨



EXTENSION WITH SUMMARY BLOCKS

- 요약 블록은 블록체인상에 일반적인 블록처럼 통합됨
  - 타임스탬프는 이전 블록을 따름



EXTENSION WITH SUMMARY BLOCKS

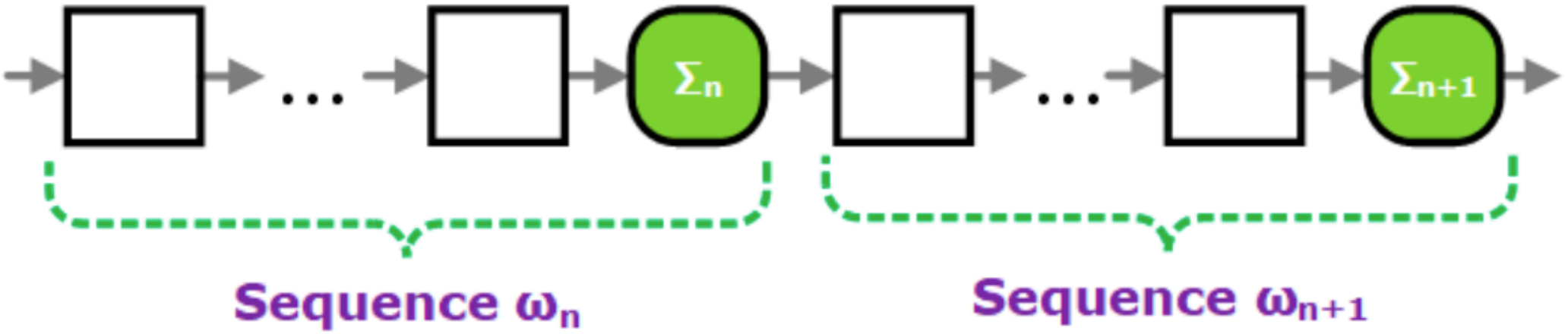
- 합의를 통하기 때문에
  - 요약 블록은 각 노드에게 동일하게 나타날 것
  - 따라서 요약 블록은 그 해시를 비교함으로써
    - 동기화의 확인을 위해 사용될 수 있음

EXTENSION WITH SUMMARY BLOCKS

- 요약 블록은 전파될 필요가 없음
  - 결정론적으로 생성되므로
- 노드가 스스로 만들어야 하며 같은 헤더를 가져야만 함
  - 만일 해시가 다르다면 포크 상황에 해당하게 됨
  - 네트워크가 분리되게 됨

LIMITING THE SEQUENCES IN A BLOCKCHAIN

- 요약 블록이 도입되며
  - 블록체인이 시퀀스(sequence) 단위로 분리되는 것으로 보임



LIMITING THE SEQUENCES IN A BLOCKCHAIN

- 블록체인의 길이를 줄이는 방법
  - 프로토콜이나 쿼럼(Quorum)에서
  - 블록체인의 길이를 l_{max}로 정의하거나
  - 시퀀스 ω의 최대 길이를 설정하는 식으로

LIMITING THE SEQUENCES IN A BLOCKCHAIN

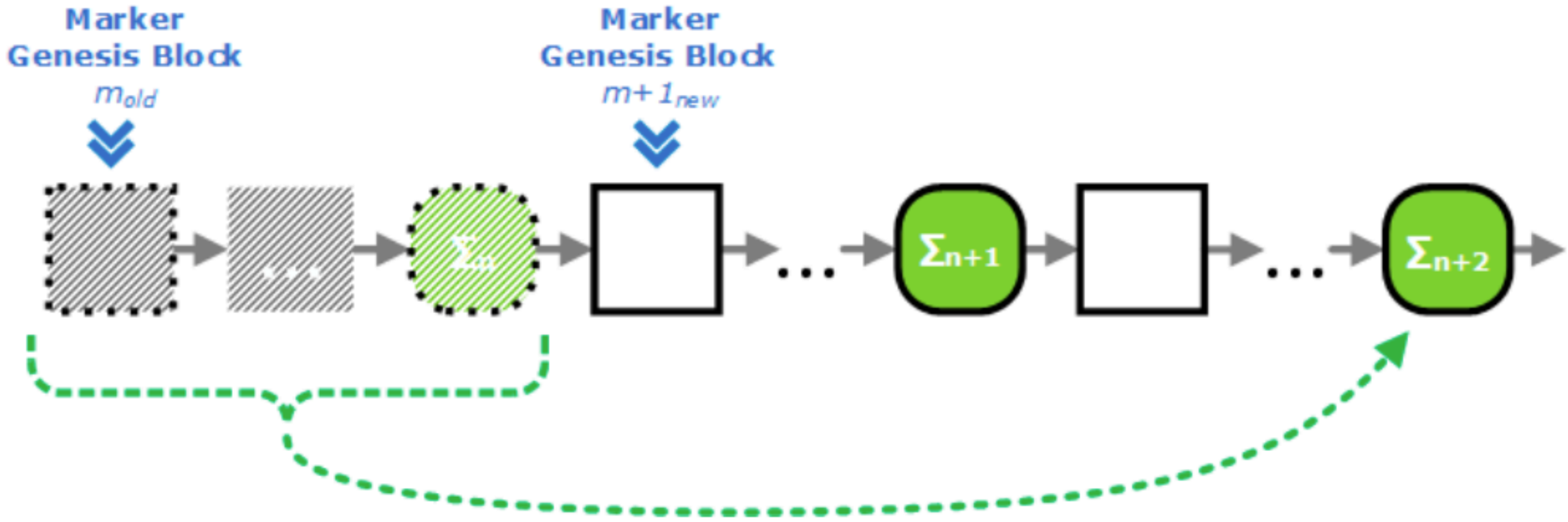
- 만일 블록체인이 l_{max}보다 길어지면
  - 가장 오래된 시퀀스가 다음 요약 블록에 통합됨
  - 이렇게 함으로써 첫 시퀀스의 블록 각각의 정보가
  - 새 요약 블록에 포함되도록 함
- 오래된 요약 블록의 정보가 지워짐
- 복사된 정보는 블록의 데이터 영역에 해당되게 되며
  - 논스와 이전블록해시 필드 등은 필요 없어짐

LIMITING THE SEQUENCES IN A BLOCKCHAIN

- 제네시스 블록도 재설정해야 함
  - 현재 체인의 블록에 기반해서
  - 삭제될 시퀀스의 요약 블록 이후 블록이 됨

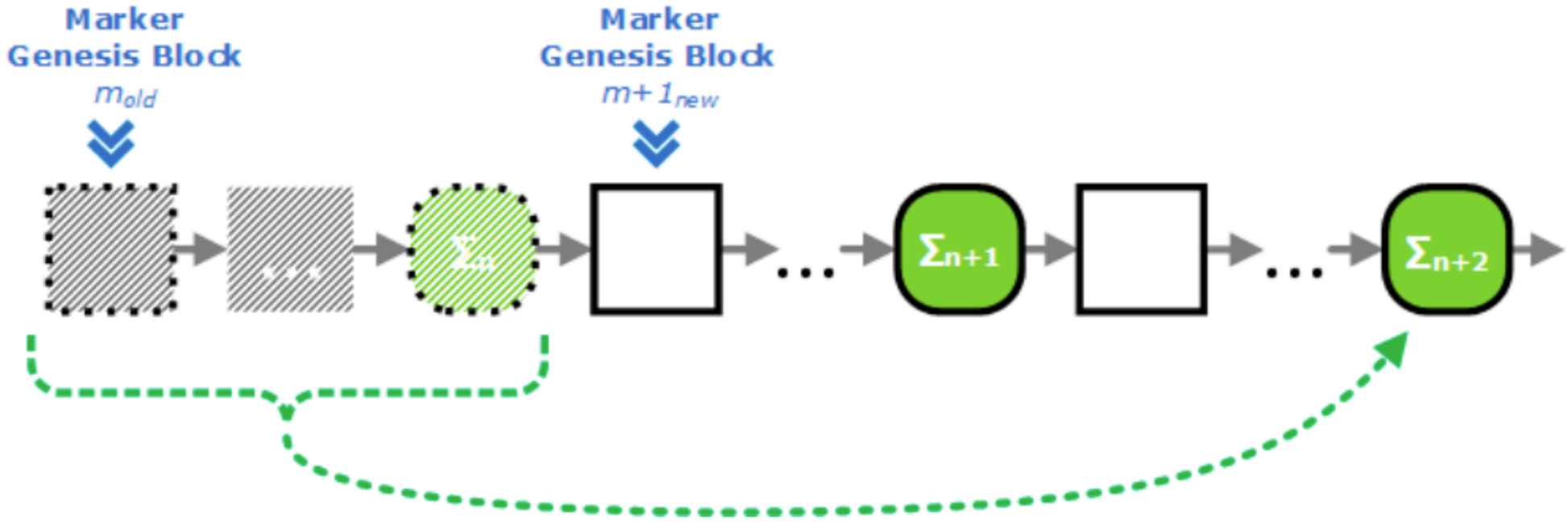
LIMITING THE SEQUENCES IN A BLOCKCHAIN

- 가장 오래된 시퀀스가 새 요약 블록에 요약되는 절차



LIMITING THE SEQUENCES IN A BLOCKCHAIN

- 정보는 블록체인에 계속 포함되어 있으므로
  - 체인은 여전히 유효하고 완전함



LIMITING THE SEQUENCES IN A BLOCKCHAIN

- 요약 블록의 데이터 영역은 다음과 같이 생김:
  - 콘텐츠가 원래 블록에서부터 복제됨
  - 블록 번호, 타임스탬프, 엔트리 번호가 유지됨

HEADER
Blocknumber: 42
Timestamp: 4711
Previous Hash: ….
Nonce: ….

DATA:
Merkle Root: ….
- Entry 1
  - Block number X
  - Timestamp of Block X
    - Entry number A
    - Stored Dataset / transaction
    - Entry number B
    - Stored Dataset / transaction
    - …
- Entry 2
  - Block number Y
  - Timestamp of Block Y
    - Entry number C
    - Stored Dataset / transaction
    - …
- …

DELETING INFORMATION ON REQUEST

- 특정 정보를 지우기 위해서
  - 참여자가 삭제 요청을 보내는 새 기능이 필요함
  - 이를 위해 사용자가 삭제 트랜잭션을 등록함
    - 일반적인 트랜잭션과 동일한 형태
  - 참조된 블록 번호의 해당하는 엔트리 번호의 데이터가 삭제될 것
- 요약 블록에 포함된 데이터에도 해당됨

DELETING INFORMATION ON REQUEST

- 블록이 숫자에 의해 직접 참조되기 때문에
  - 이 절차는 복잡하지도 않고
  - 비용이 매우 적게 소요될 것

DELETING INFORMATION ON REQUEST

- 세 가지 관점에서 기능을 살펴볼 것
  - 1) 권한
  - 2) 지연된 삭제
  - 3) 임시 엔트리

DELETING INFORMATION ON REQUEST

- 1) 권한
  - 사용자가 정보를 지울 수 있는 권한이 있는지 확인하기 위해
  - 삭제 요청은 (일반적인 경우처럼) 서명을 요구함
- 여러 방법이 있겠지만, 가령
  - 쿼럼의 (신뢰 받는) 앵커 노드가 관리 권한을 전적으로 위임받을 수 있음
  - 마스터 서명

DELETING INFORMATION ON REQUEST

- 이는 응용/목적에 따라 다를 것
  - 또 다른 방법으로, 투표 기반으로 이루어질 수도 있고
  - 룰 기반으로 이루어질 수도 있고
  - 자신의 데이터는 오직 자신만 지우도록 할 수 있음
    - 지우고자 하는 데이터의 서명과 비교
  - 등
- 아무튼, 여러 방법을 통해 삭제 요청이 블록체인에 통합됨

DELETING INFORMATION ON REQUEST

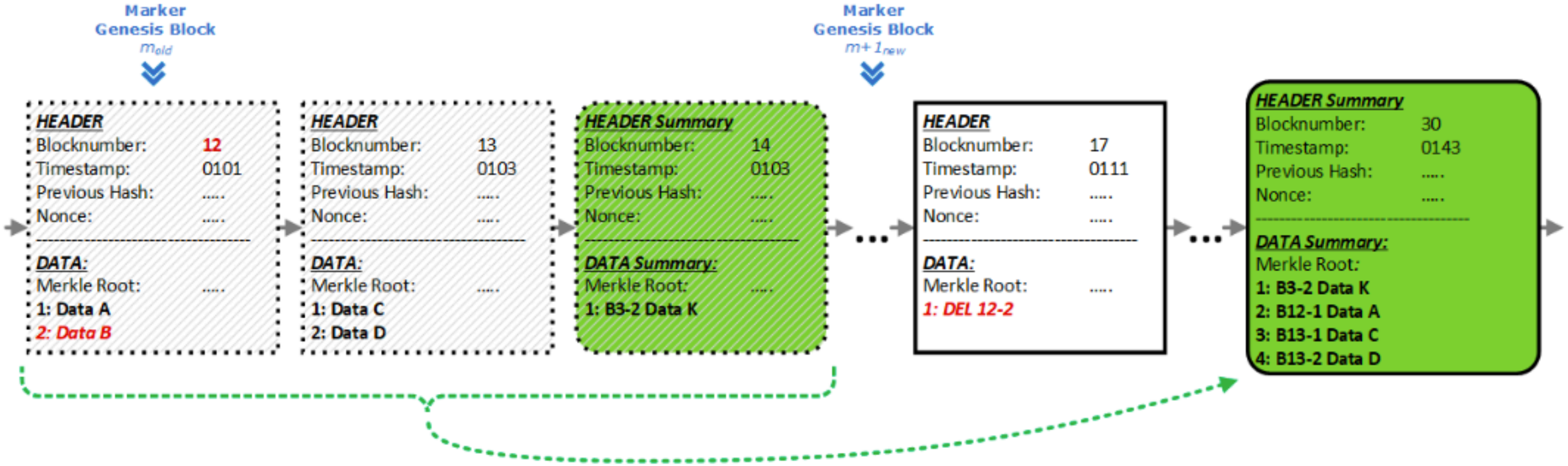
- 2) 지연된 삭제
  - 이 경우, 정보의 삭제는 즉각적이지 않음
  - 특정 데이터가 미래에 삭제될 것으로 표기됨
  - 특정 시점 이후, 이 데이터에 접근하는 트랜잭션의 수행은 허가되지 않음
  - 표기된 데이터는 시퀀스가 오래되면 삭제되게 됨
- 삭제 요청을 계속 유지할 필요가 없으므로
  - 요약 블록에 삭제 트랜잭션을 포함되지 않을 수 있음

DELETING INFORMATION ON REQUEST

- 3) 임시 엔트리
  - 지연된 삭제 디자인은 임시 엔트리로 확장될 수 있음
  - 특정 타임스탬프나 블록 번호에 도달하면 자동으로 삭제되는 데이터
  - 최대 스토리지 타임에 해당하는 추가적 데이터 필드를 요구함
- 블록체인이 이를 초과할 경우
  - 권한 필요 없이 자동으로 블록체인에서 제거됨

EXAMPLE

- 데이터 B가 삭제됨
  - 삭제 요청에 의해 새 요약 블록에 복제되지 않음



DISCUSSION

DISCUSSION: SIZE OF SUMMARY BLOCKS

- 요약 블록에서 정보가 추가되므로
  - 시간이 지남에 따라 많이 커질 수 있음
    - 복제될 데이터의 양에 따라
- 이는 오프체인 솔루션으로 해결 가능
  - 외부 저장소에 저장하고 참조 해시만 블록체인에 등록
  - Luke’s comment: 솔직히 좋은 방법은 아니지…

DISCUSSION: SIZE OF SUMMARY BLOCKS

- 다른 방법으로는 결과만을 저장하는 것
  - 모든 트랜잭션을 그대로 담지 않고,
  - 결과 상태를 저장하는 방법
  - Luke’s comment: 유효하게 작은 크기일지는 의문

SUMMARY

SUMMARY

- 블록체인에서 엔트리를 제거할 수 있는 콘셉트 소개
- 블록체인이 시퀀스를 단위로 나뉘고
  - 요약 블록을 포함하게 됨
  - 요약 블록은 오래된 정보를 포함하고 있음
- 오래된 시퀀스는 제거되고 새 요약 블록으로 그 정보가 통합되게 됨
  - 이 과정에서 데이터 삭제가 가능해짐
- 제네시스 블록이 시퀀스를 따라 계속 이동하는 격

SUMMARY

- 이러한 방법은 모든 종류의 블록체인에 통합될 수 있음
  - 합의 알고리즘과 상관 없이
  - 비트코인 등
- 요약 블록의 용량 문제가 발생할 수 있음
  - 그렇지만 요약 블록은 전파할 필요는 없기 때문에
  - 네트워크 대역폭 등은 고려하지 않아도 됨

SELECTIVE DELETION

IN A BLOCKCHAIN