FIXME

- IPFS lab does not really share the files
  - Test out with a VM
- Permacoin lab might be interesting

CS 168: Blockchain and Cryptocurrencies

Storage and the Blockchain

Prof. Tom Austin
San José State University

Storage and the Blockchain

- Storage for consensus
- Storage as a byproduct
- Dropbox on the blockchain
- Off-chain storage

Dimensions of Storage Proving Schemes

- Publicly verifiable
- Retrievable
- Zero-knowledge
- Useful
- Dynamically updateable

Verifying Storage

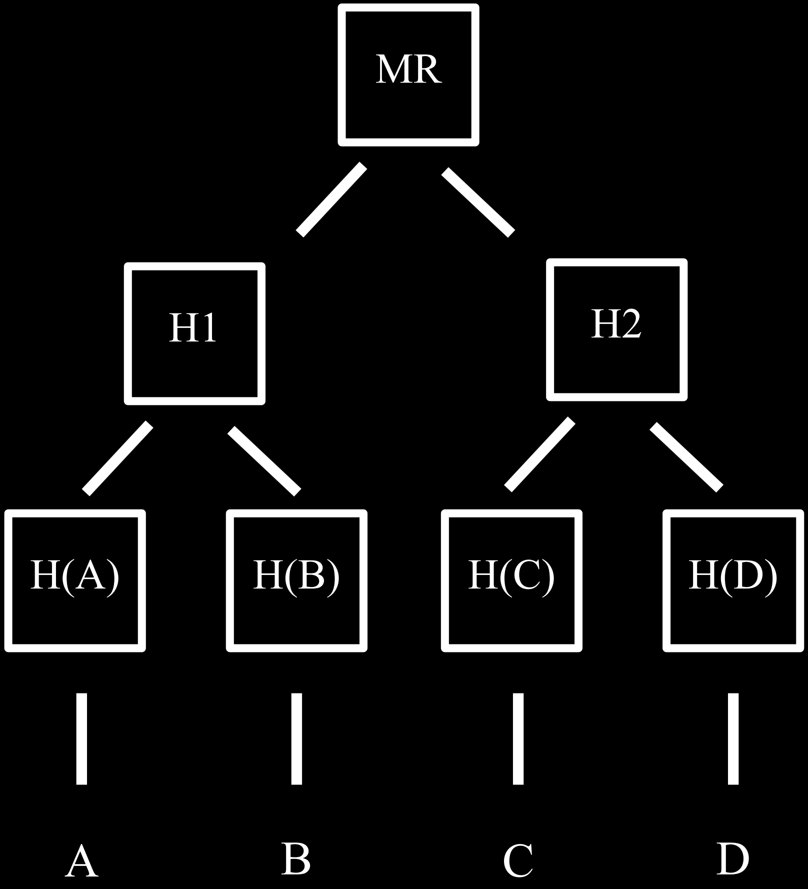
- What knowledge is needed?
- Who can we trust?
  - –Miners?
  - –Storage providers?
  - –Clients?

Review: Merkle Trees

H1 = H(H(A),H(B))

H2 = H(H(C),H(D))

MR = H(H1,H2)
(Merkle root)



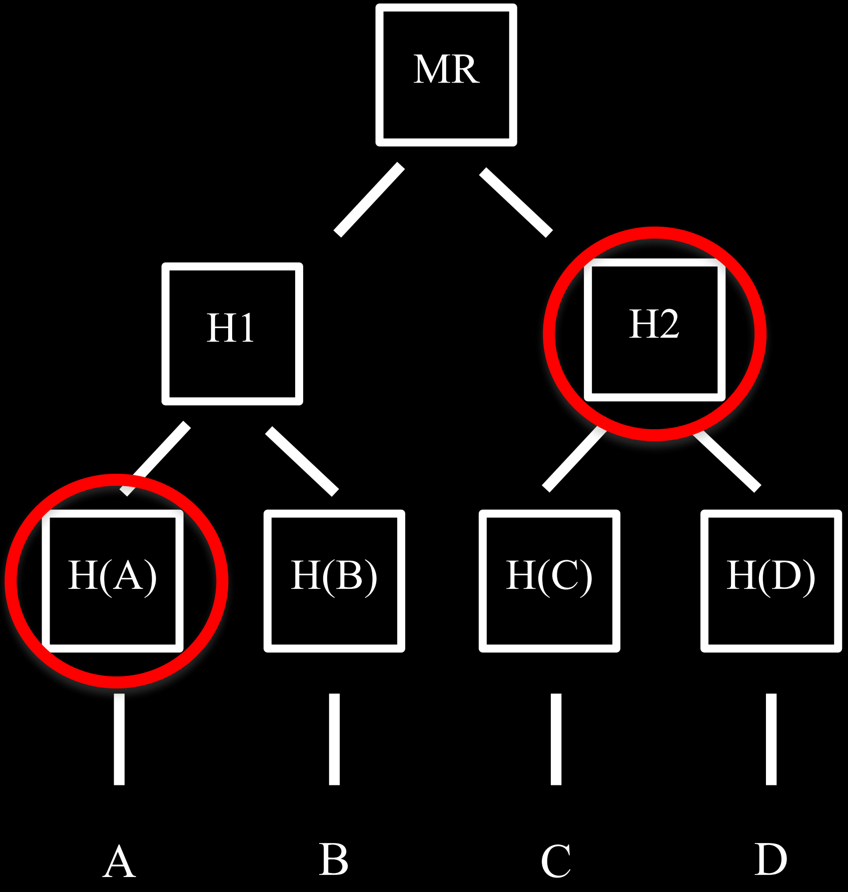
Using Merkle Trees for Storage

- Merkle root of data is known
- Challenger requests specific block(s)
- Attacker provides Merkle Proof
  - Pieces needed to reconstruct Merkle root

Merkle Trees for Storage Proofs

Merkle proof for block B:

- Block B
- H(A)
- H2



Spacemint: Storage for Consensus

- Data only useful for consensus
- Miners invest disk space (PoSpace)
- Motivation
    - Minimal computation
    - Egalitarian

Archival Storage

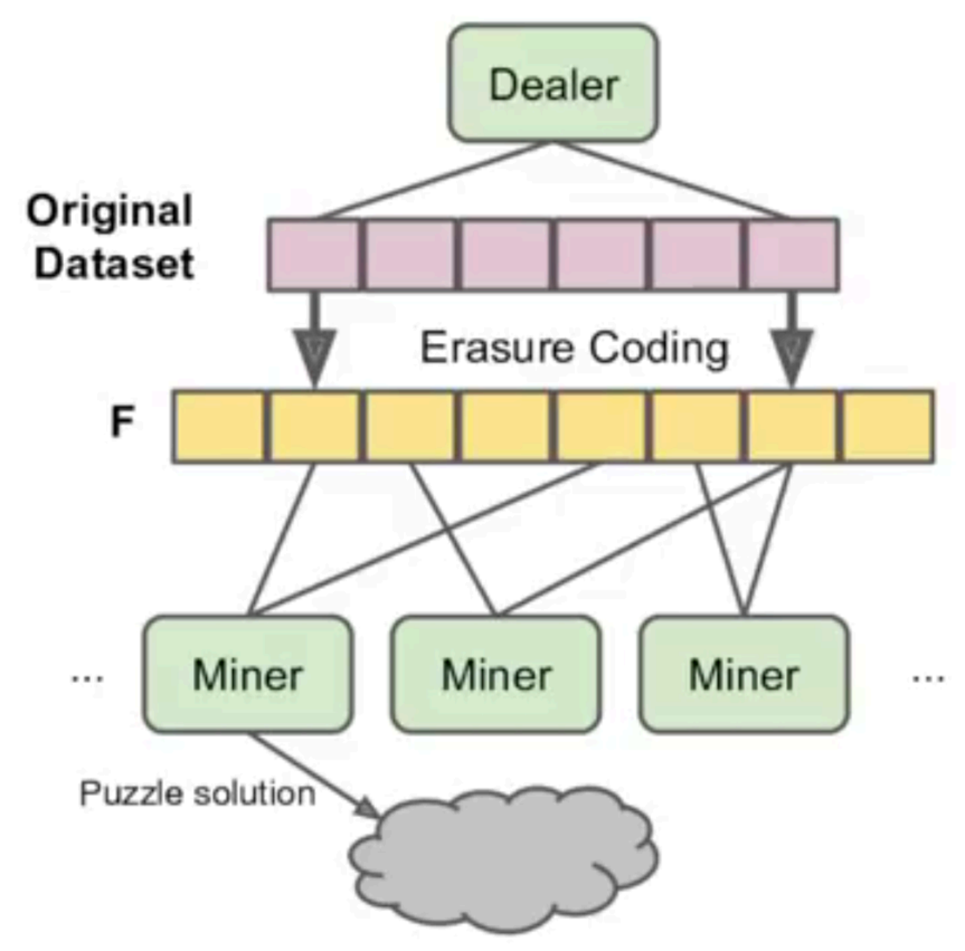
Permacoin: Useful, incidental storage

- Storage of archival data
- Miller et al. 2014
- Proof-of-work (PoW) and proof-of-retrievability (PoRet)
  - Solve proof-of-retrievability
  - Solution feeds into PoW puzzle

Permacoin Process

(taken from https://www.youtube.com/watch?v=gIJim7JKW_M)

1. Setup – archival file is *erasure coded*
2. Users generate keypairs
3. Miners look for solutions
   - Requires locally storing data



"Puzzle Solving"

Bitcoin puzzle solving:

- `H(puz||pk||r) < target`

Permacoin solves 2 puzzles (in sequence):

1. `H(puz||pk||r)` selects blocks to reveal
2. `H(puz||pk||r||dataBlks) < target`

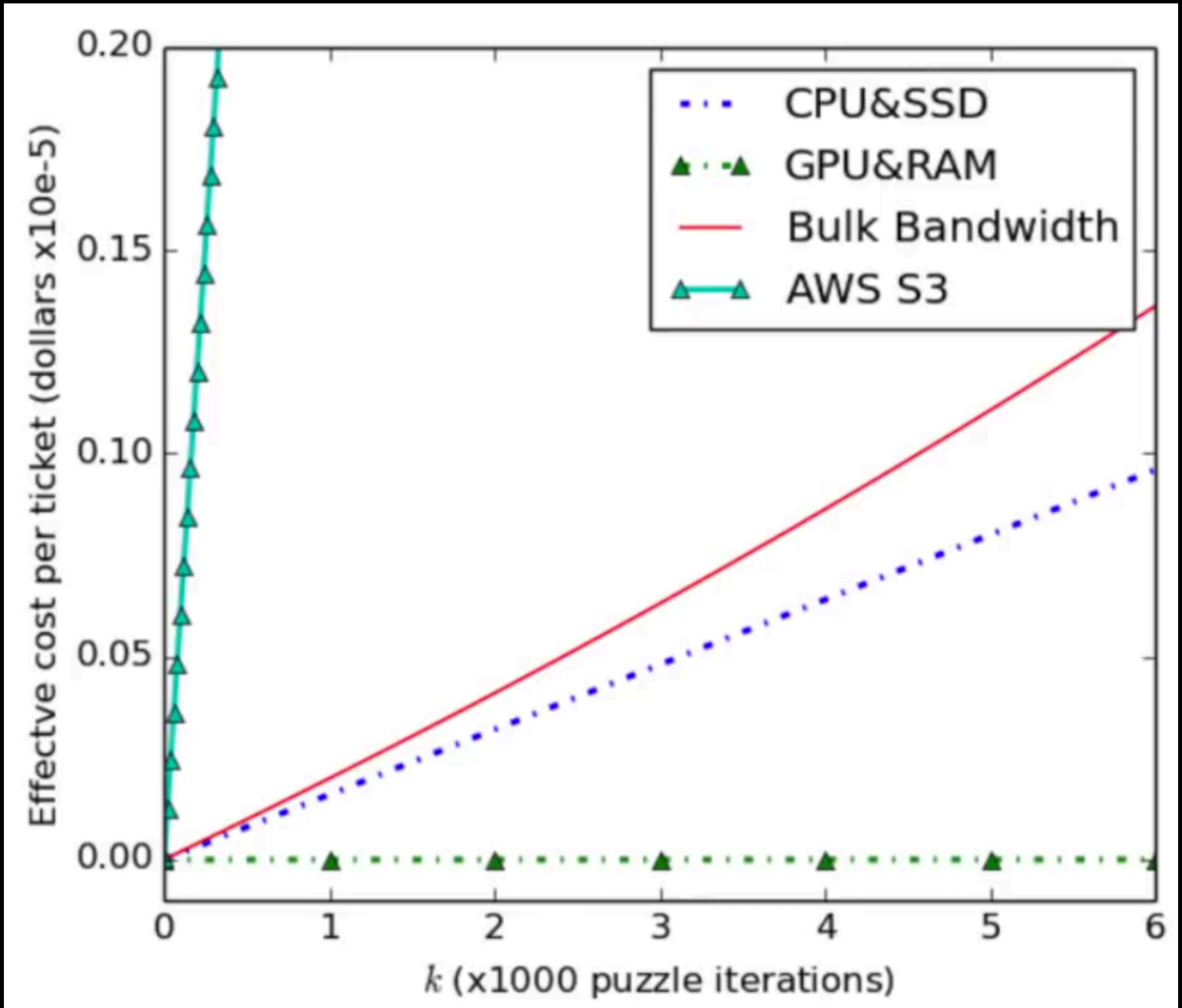
If data is not stored, 1st solution found is useless.

Forcing Local Storage

- Goal: prevent outsourcing of storage.
- Solution: modify previous approach to include a signing step.
  - Related to non-outsourceable puzzles.
- Miner then must choose:
  - Share data and keys with the 3rd party
    - Keys could be stolen
  - Store data remotely, but keys locally
  - Store data and keys locally

Economics of Permacoin Mining

(taken from https://www.youtube.com/watch?v=gIJim7JKW_M)



Dropbox on the Blockchain

Filecoin

sia

STORJ

MaidSafe

InterPlanetary File System (IPFS)

- Content-addressable storage
  - Hash of data serves as its ID
- Peer-to-peer
- Used in Catalan independence referendum
- No real guarantees data will be stored long term

Filecoin

- Incentive layer for IPFS (next slide)
- Storage market
  - Guarantees data is stored
  - Very slow, by design
- Retrieval market
  - Caches frequently requested data
  - Offers CDN functionality
  - (Details a little murky)

Attacks

- Outsourcing
- Generation
- Sybils (or collusion)

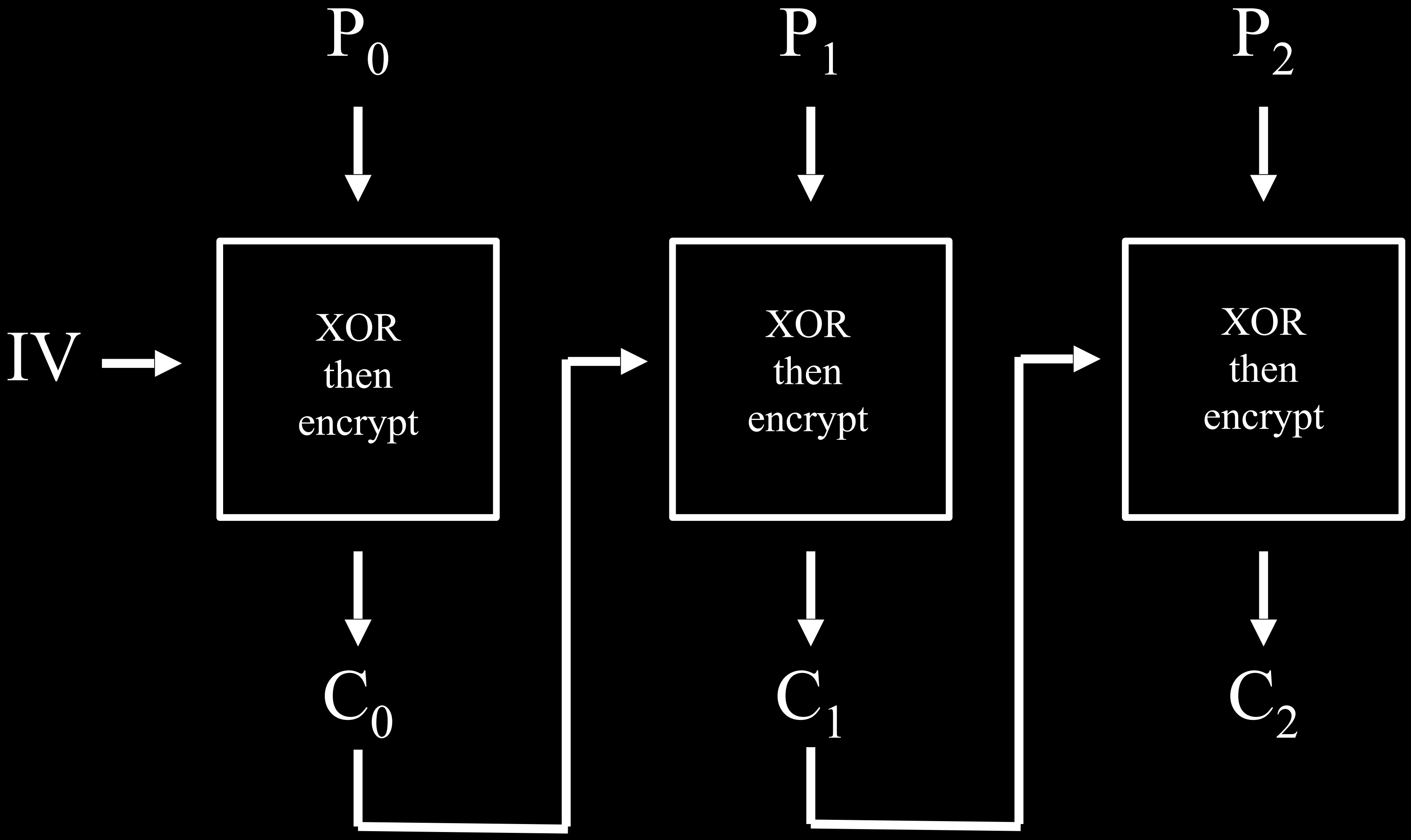
Review: Cipher Block Chaining (CBC)

- Block – data chunk cipher encrypts
  - No relation to blockchain blocks
- $C_0 = E(IV \oplus P_0, K)$

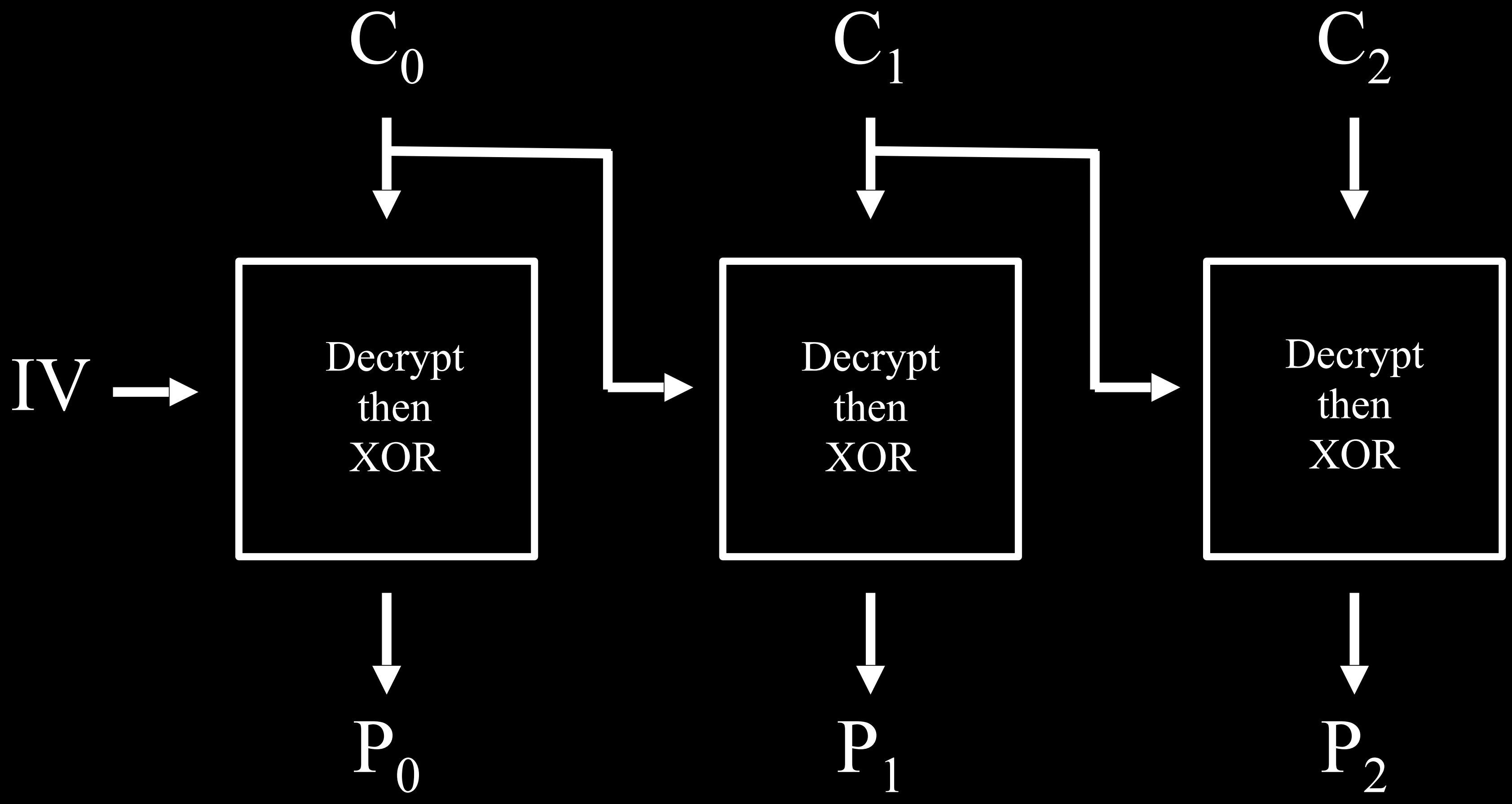
  $C_i = E(C_{i-1} \oplus P_i, K)\ \forall i.\ i>0$
- $P_0 = IV \oplus D(C_0, K)$

  $P_i = C_{i-1} \oplus D(C_i, K)\ \forall i.\ i>0$

CBC Encryption



CBC Decryption



Can encryption be parallelized?

Can decryption be parallelized?

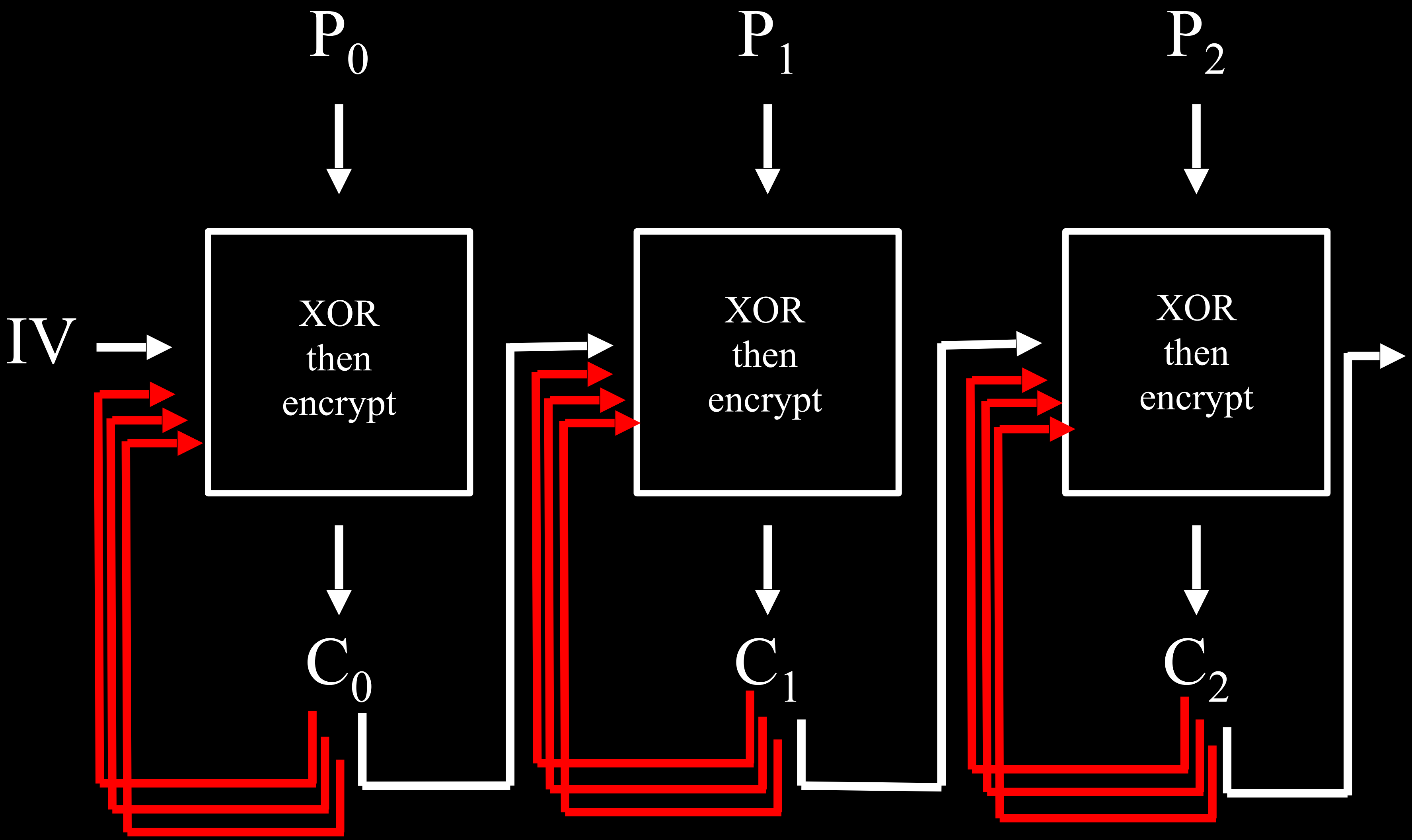
Proof-of-replication

- Ensure that miner is storing **as many copies of a file as they claim**.
- Each copy of data must be unique
  - Ensured by *sealing key*
- Miner must provide data within time limit
- Uses modified versions of CBC mode
  - Slows down encryption

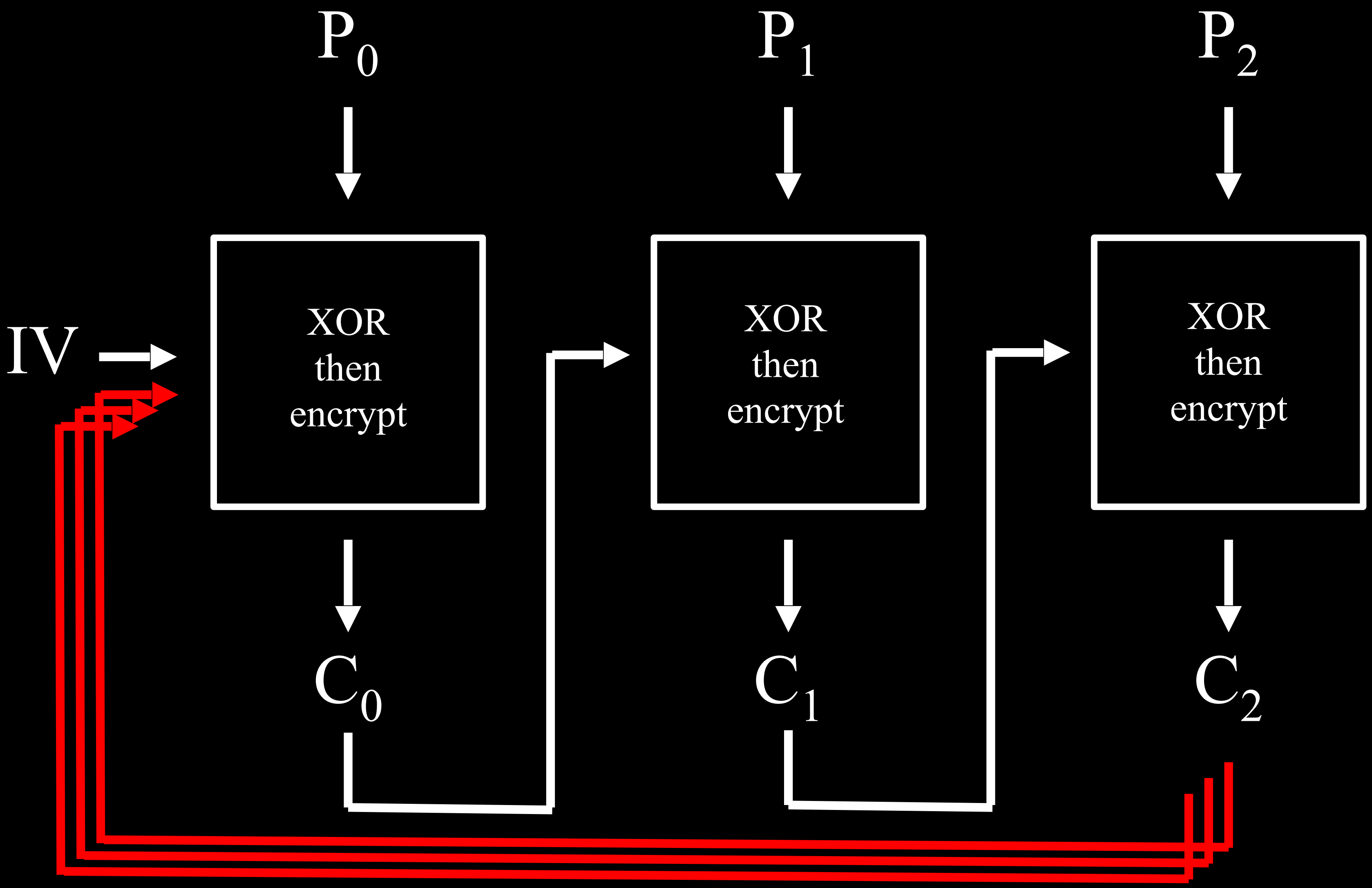
Modified CBC Modes

- Shuffling
  - Data spread across many blocks
- Streaming
  - Each block chained to itself N times
- Layering
  - The last block is chained to the first block M times.

CBC Encryption, Streaming Mode



CBC Encryption, Layering Mode



Proof-of-spacetime

- Filecoin miners can also prove that they are continually storing their data.
- Proof-of-replication determines next round of challenge.
- Miners write these proofs to the blockchain to get paid.