FRIDA

Data Availability Sampling from FRI

Mathias Hall-Andersen

Mark Simkin

Benedikt Wagner

Data Availability Sampling

Data Availability Sampling from FRI

Data Availability Sampling

Data Availability Sampling from FRI

Data Availability Sampling

Data Availability Sampling

Ethereum Community

Central For Roadmap

Vague Idea / Concept

Few Constructions

Data Availability Sampling

Ethereum Community

Cryptography Community

Central For Roadmap

Vague Idea / Concept

Few Constructions

Data Availability Sampling

Ethereum Community

Cryptography Community

Central For Roadmap

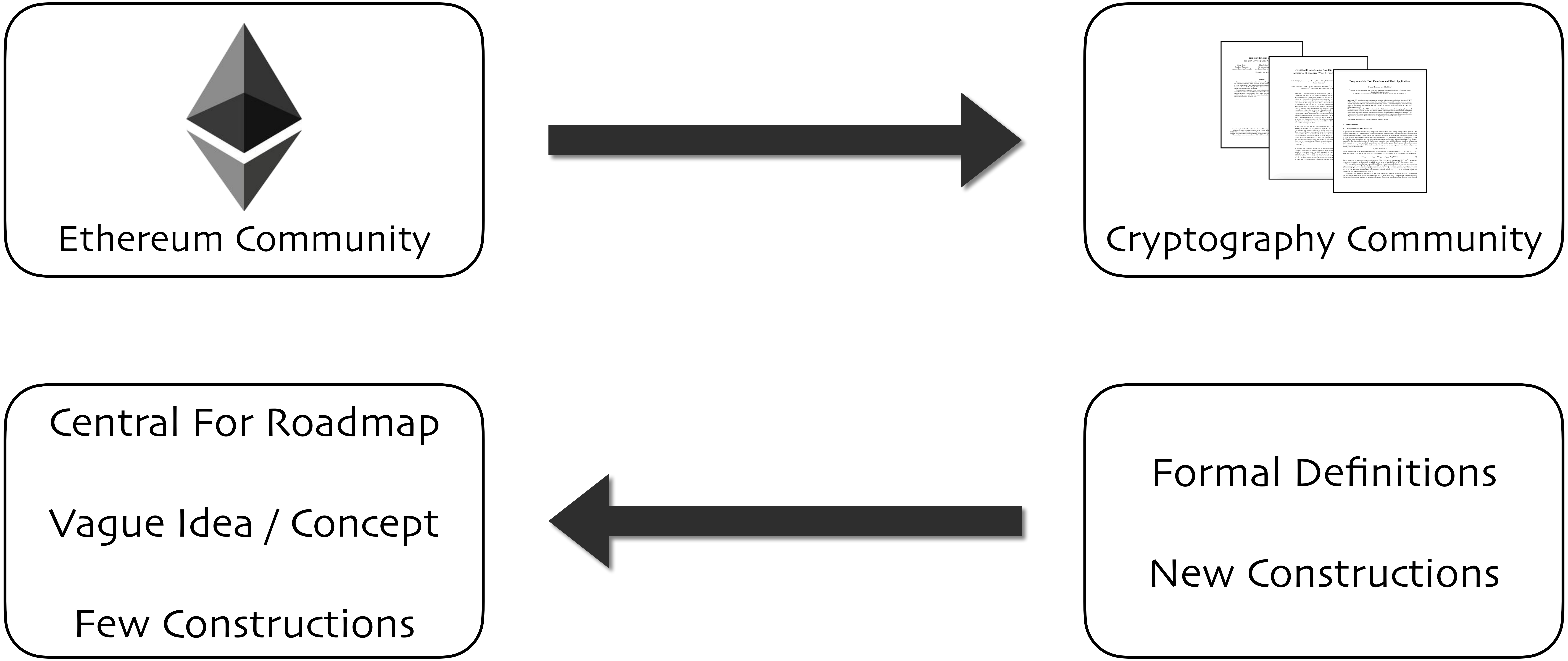
Vague Idea / Concept

Few Constructions

Formal Definitions

New Constructions

Data Availability Sampling



Ethereum Community

Cryptography Community

Central For Roadmap

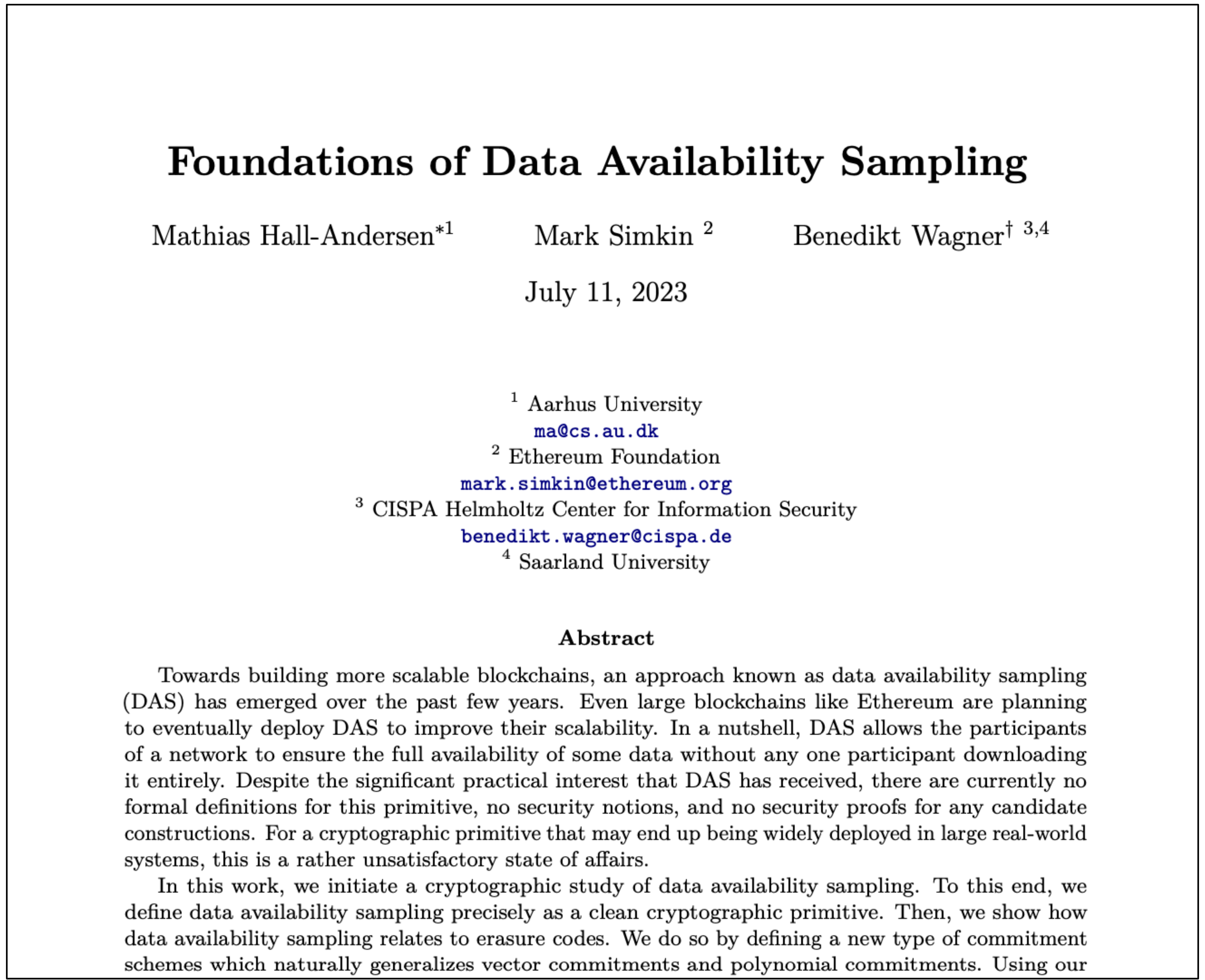
Vague Idea / Concept

Few Constructions

Formal Definitions

New Constructions

Foundations of DAS



Foundations of Data Availability Sampling

Mathias Hall-Andersen[*1] Mark Simkin[2] Benedikt Wagner[†][3,4]

July 11, 2023

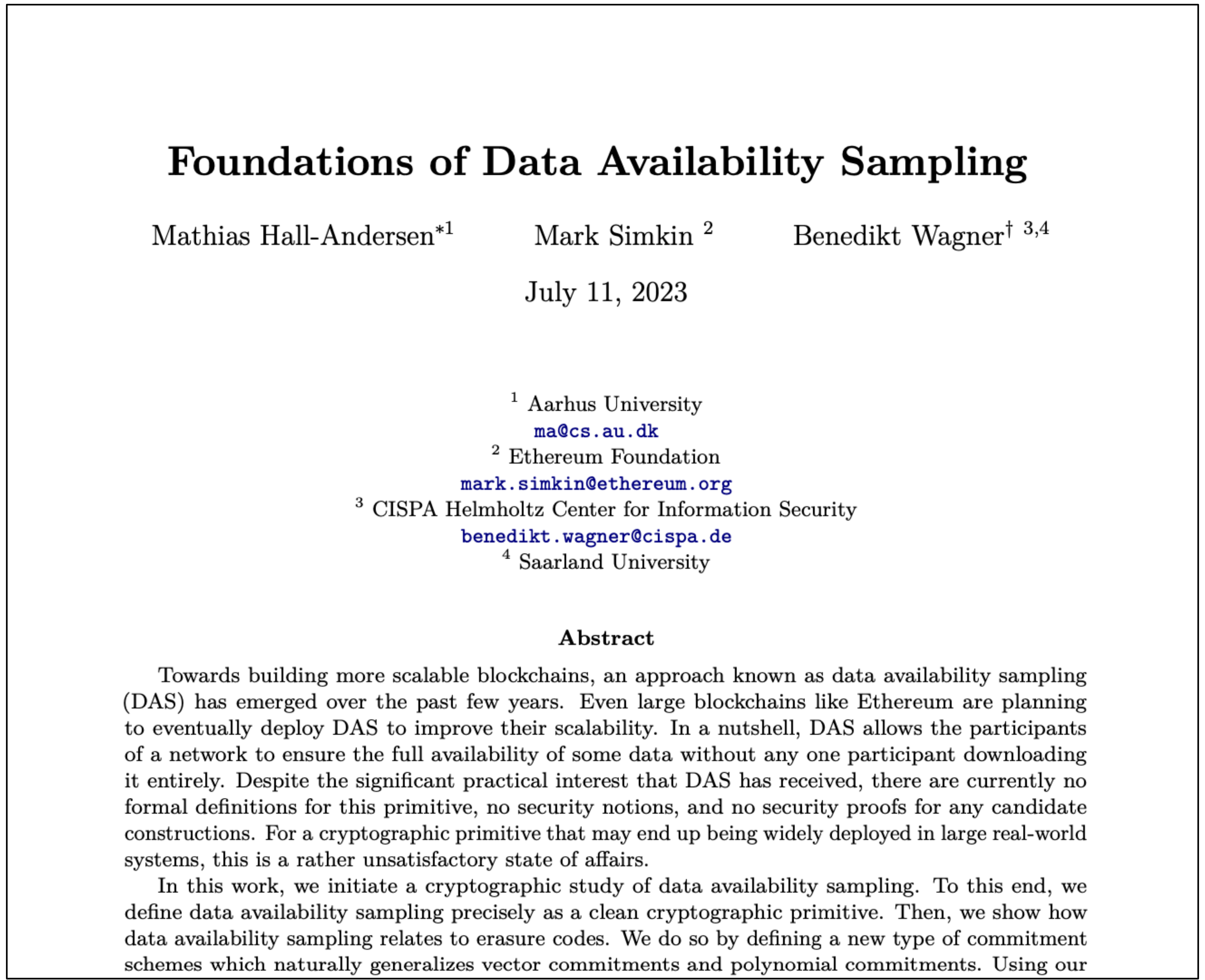
[1] Aarhus University
ma@cs.au.dk
[2] Ethereum Foundation
mark.simkin@ethereum.org
[3] CISPA Helmholtz Center for Information Security
benedikt.wagner@cispa.de
[4] Saarland University

Abstract

Towards building more scalable blockchains, an approach known as data availability sampling (DAS) has emerged over the past few years. Even large blockchains like Ethereum are planning to eventually deploy DAS to improve their scalability. In a nutshell, DAS allows the participants of a network to ensure the full availability of some data without any one participant downloading it entirely. Despite the significant practical interest that DAS has received, there are currently no formal definitions for this primitive, no security notions, and no security proofs for any candidate constructions. For a cryptographic primitive that may end up being widely deployed in large real-world systems, this is a rather unsatisfactory state of affairs.

In this work, we initiate a cryptographic study of data availability sampling. To this end, we define data availability sampling precisely as a clean cryptographic primitive. Then, we show how data availability sampling relates to erasure codes. We do so by defining a new type of commitment schemes which naturally generalizes vector commitments and polynomial commitments. Using our

Foundations of DAS



Foundations of Data Availability Sampling

Mathias Hall-Andersen*[1] Mark Simkin [2] Benedikt Wagner† [3,4]

July 11, 2023

[1] Aarhus University
ma@cs.au.dk
[2] Ethereum Foundation
mark.simkin@ethereum.org
[3] CISPA Helmholtz Center for Information Security
benedikt.wagner@cispa.de
[4] Saarland University

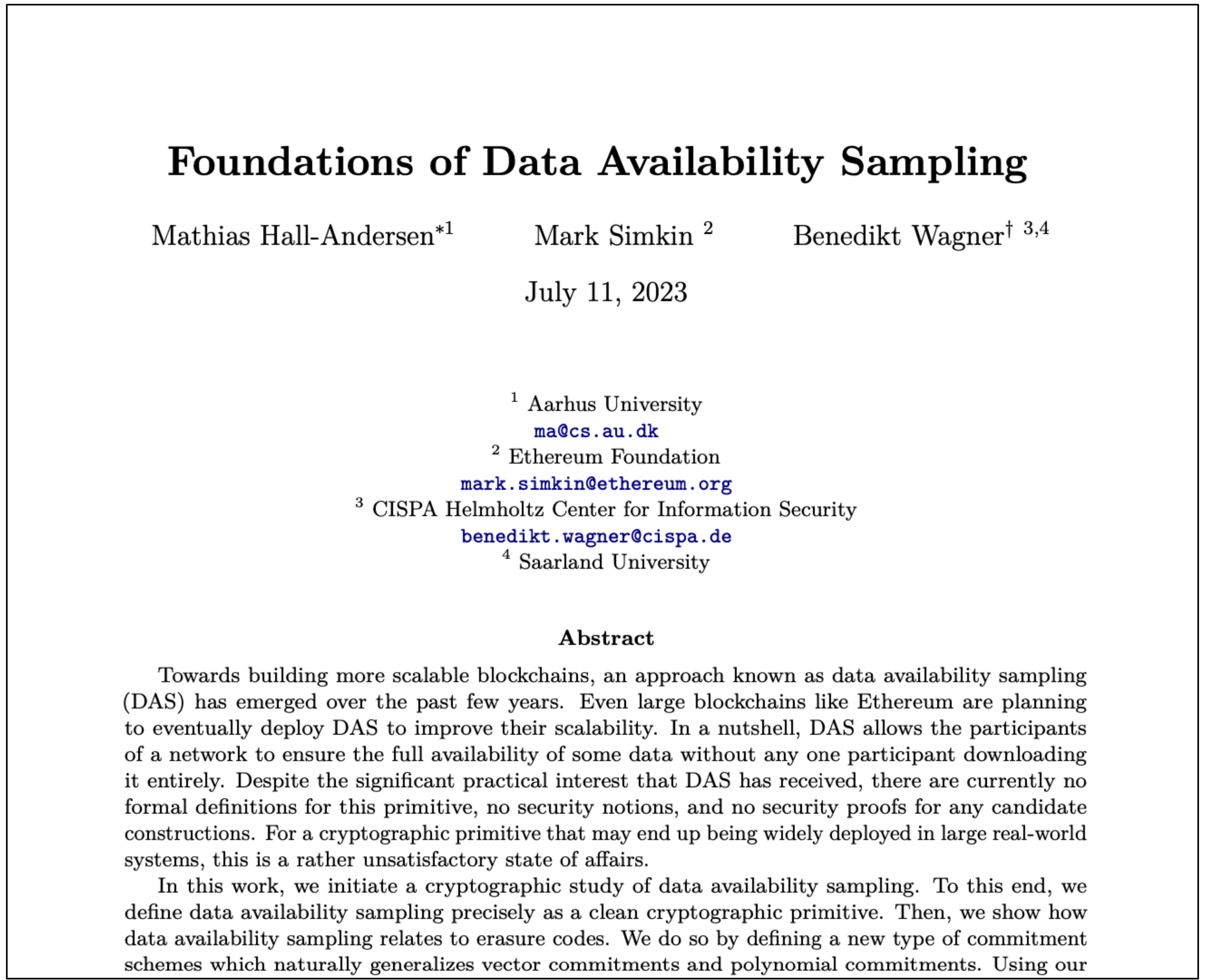
Abstract

Towards building more scalable blockchains, an approach known as data availability sampling (DAS) has emerged over the past few years. Even large blockchains like Ethereum are planning to eventually deploy DAS to improve their scalability. In a nutshell, DAS allows the participants of a network to ensure the full availability of some data without any one participant downloading it entirely. Despite the significant practical interest that DAS has received, there are currently no formal definitions for this primitive, no security notions, and no security proofs for any candidate constructions. For a cryptographic primitive that may end up being widely deployed in large real-world systems, this is a rather unsatisfactory state of affairs.

In this work, we initiate a cryptographic study of data availability sampling. To this end, we define data availability sampling precisely as a clean cryptographic primitive. Then, we show how data availability sampling relates to erasure codes. We do so by defining a new type of commitment schemes which naturally generalizes vector commitments and polynomial commitments. Using our

Definitions

Foundations of DAS



Foundations of Data Availability Sampling

Mathias Hall-Andersen*[1] Mark Simkin [2] Benedikt Wagner† [3,4]

July 11, 2023

[1] Aarhus University
ma@cs.au.dk
[2] Ethereum Foundation
mark.simkin@ethereum.org
[3] CISPA Helmholtz Center for Information Security
benedikt.wagner@cispa.de
[4] Saarland University

Abstract

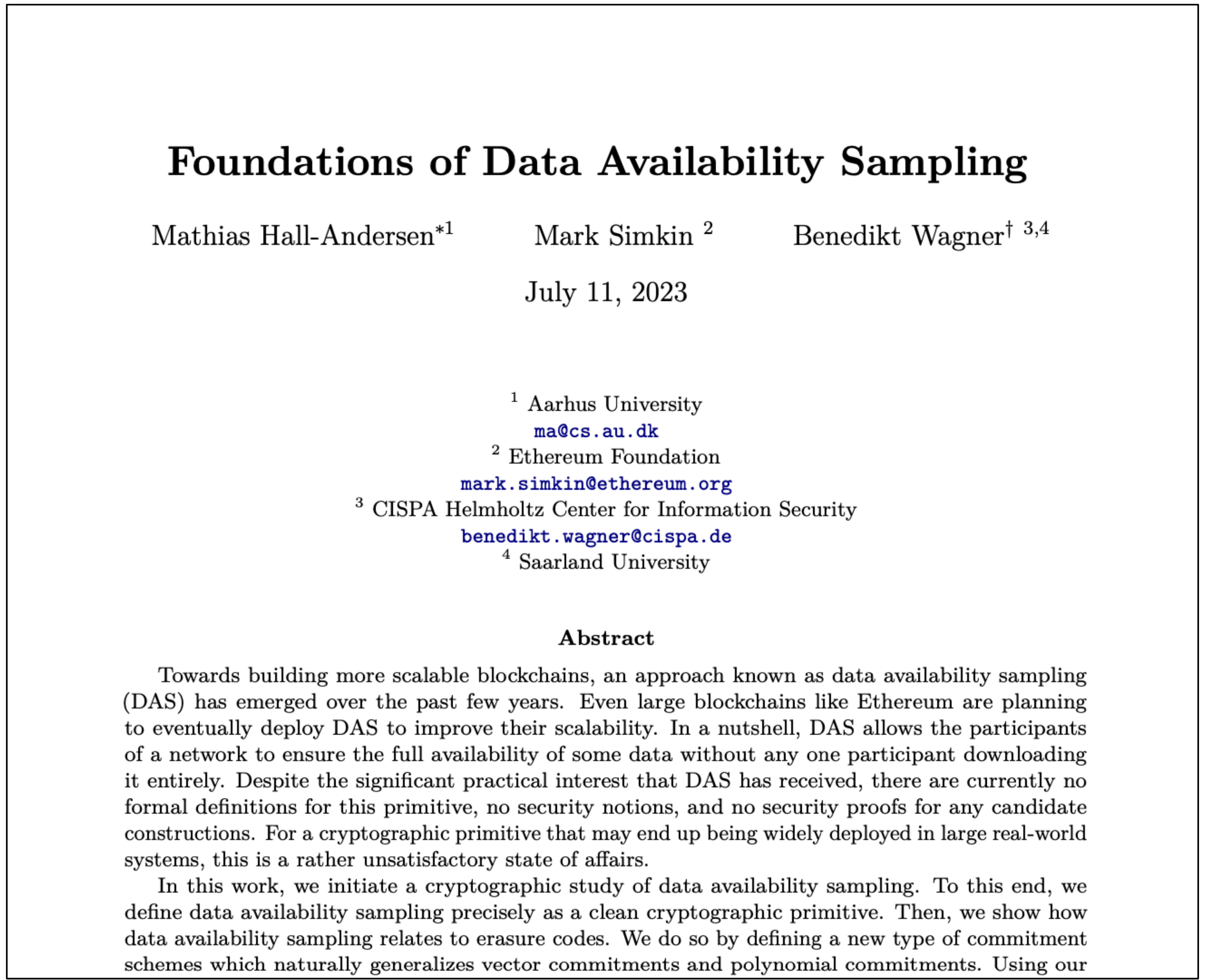
Towards building more scalable blockchains, an approach known as data availability sampling (DAS) has emerged over the past few years. Even large blockchains like Ethereum are planning to eventually deploy DAS to improve their scalability. In a nutshell, DAS allows the participants of a network to ensure the full availability of some data without any one participant downloading it entirely. Despite the significant practical interest that DAS has received, there are currently no formal definitions for this primitive, no security notions, and no security proofs for any candidate constructions. For a cryptographic primitive that may end up being widely deployed in large real-world systems, this is a rather unsatisfactory state of affairs.

In this work, we initiate a cryptographic study of data availability sampling. To this end, we define data availability sampling precisely as a clean cryptographic primitive. Then, we show how data availability sampling relates to erasure codes. We do so by defining a new type of commitment schemes which naturally generalizes vector commitments and polynomial commitments. Using our

Definitions

Constructions

Foundations of DAS



Foundations of Data Availability Sampling

Mathias Hall-Andersen*[1] Mark Simkin [2] Benedikt Wagner† [3,4]

July 11, 2023

[1] Aarhus University
ma@cs.au.dk
[2] Ethereum Foundation
mark.simkin@ethereum.org
[3] CISPA Helmholtz Center for Information Security
benedikt.wagner@cispa.de
[4] Saarland University

Abstract

Towards building more scalable blockchains, an approach known as data availability sampling (DAS) has emerged over the past few years. Even large blockchains like Ethereum are planning to eventually deploy DAS to improve their scalability. In a nutshell, DAS allows the participants of a network to ensure the full availability of some data without any one participant downloading it entirely. Despite the significant practical interest that DAS has received, there are currently no formal definitions for this primitive, no security notions, and no security proofs for any candidate constructions. For a cryptographic primitive that may end up being widely deployed in large real-world systems, this is a rather unsatisfactory state of affairs.

In this work, we initiate a cryptographic study of data availability sampling. To this end, we define data availability sampling precisely as a clean cryptographic primitive. Then, we show how data availability sampling relates to erasure codes. We do so by defining a new type of commitment schemes which naturally generalizes vector commitments and polynomial commitments. Using our

Long Talk

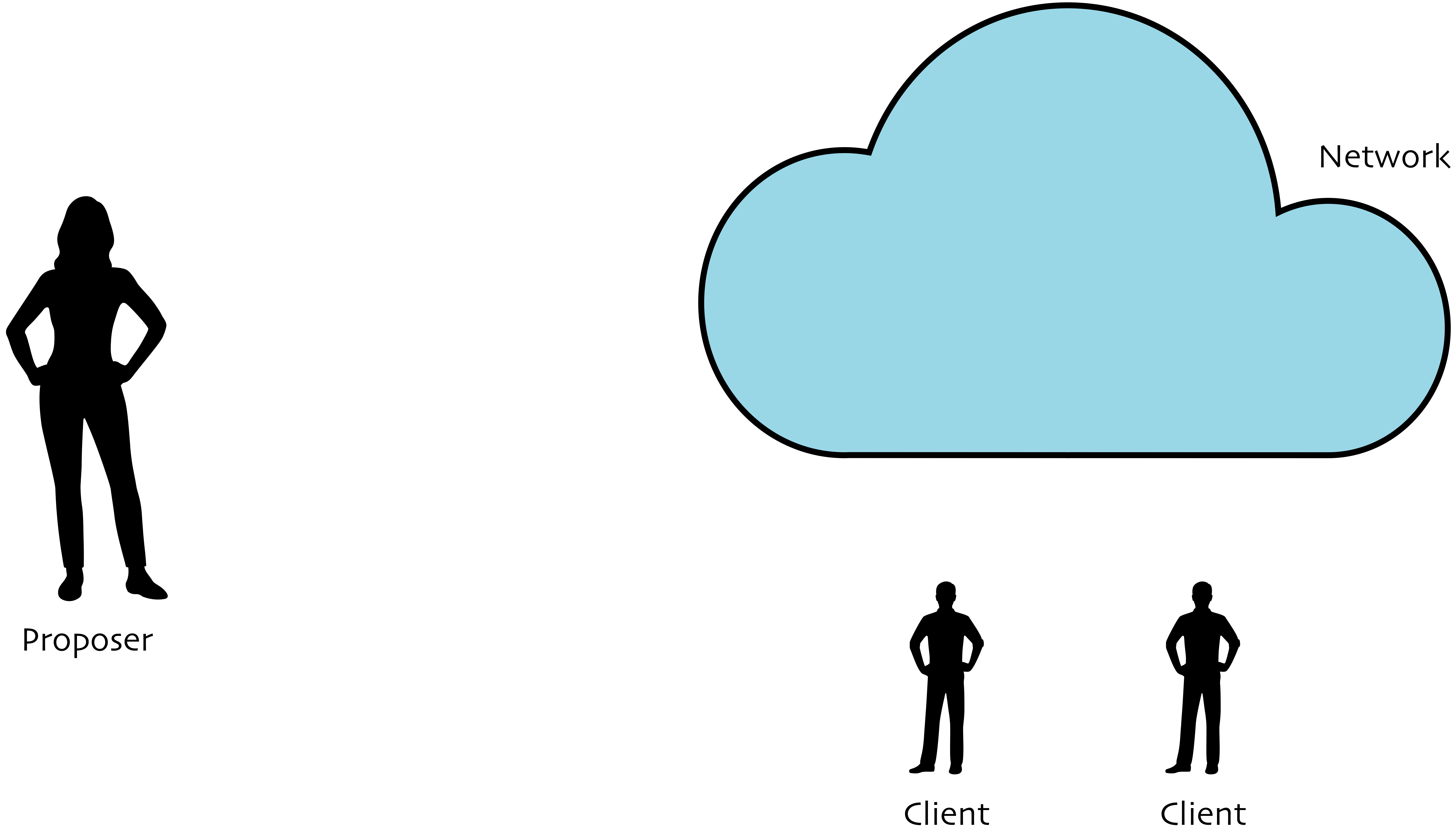
YouTube

Definitions

Constructions

Data Availability Sampling

Data Availability Sampling



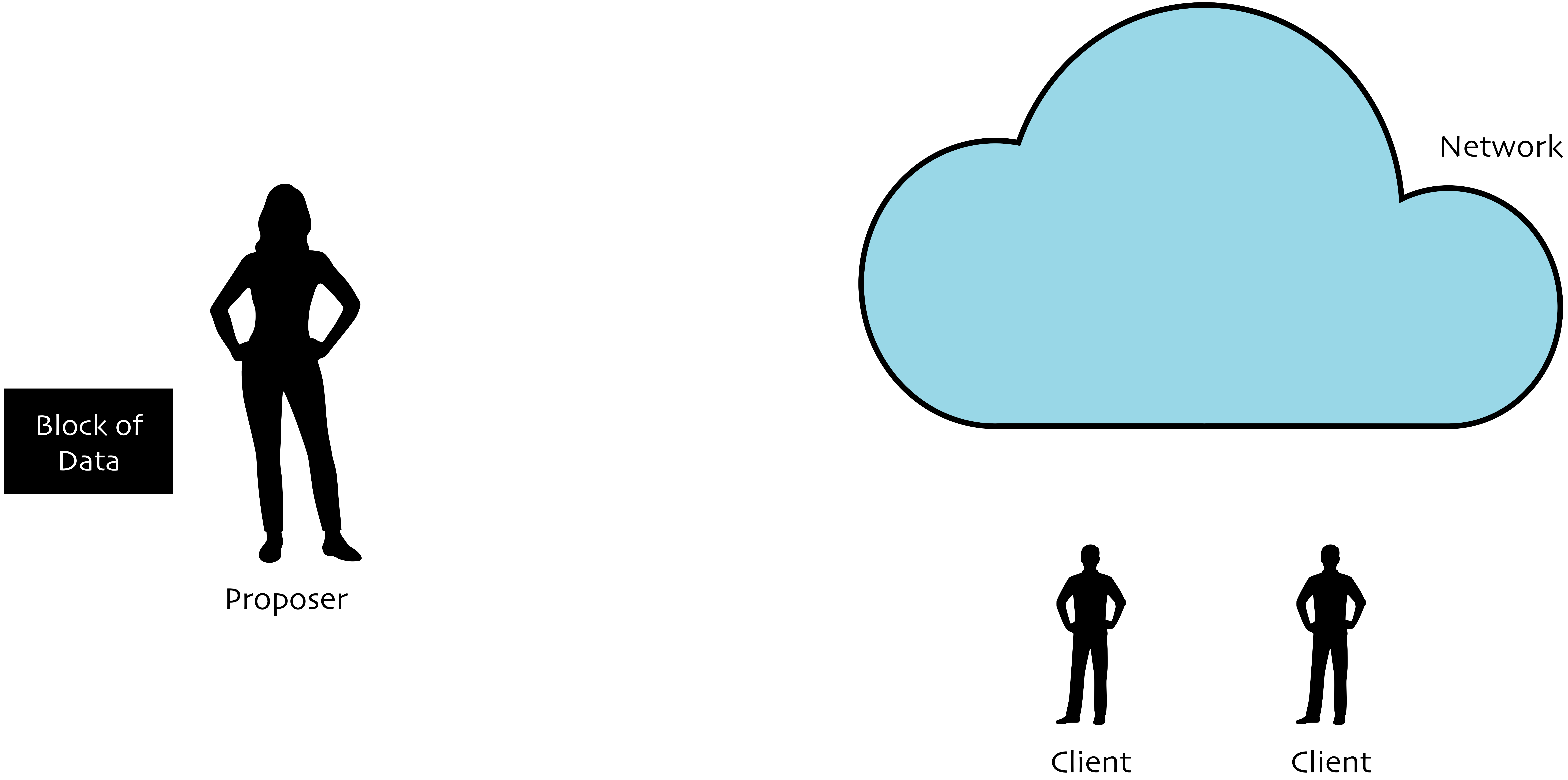
Network

Proposer

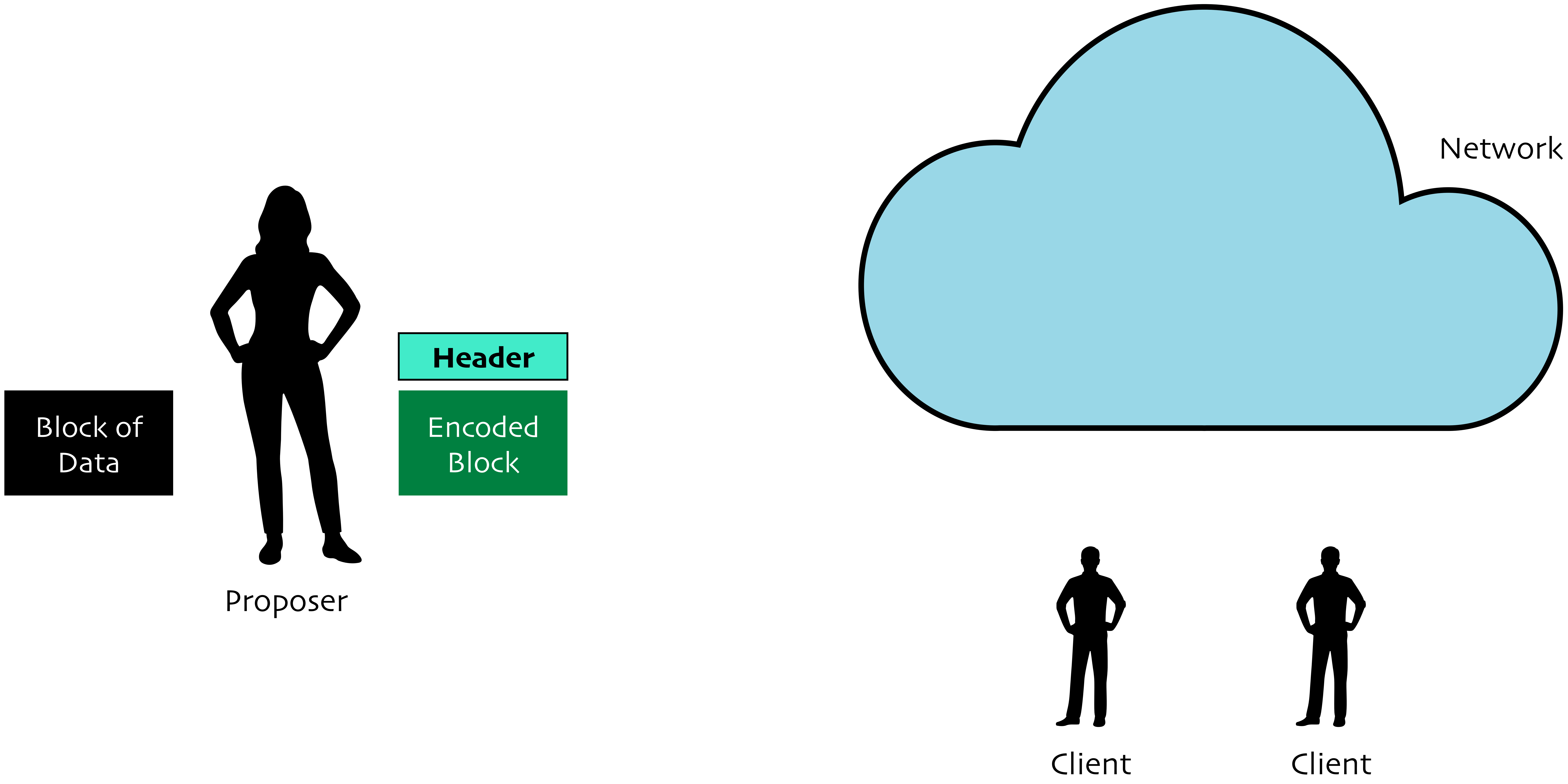
Client

Client

Data Availability Sampling



Data Availability Sampling



Block of Data

Header

Encoded Block

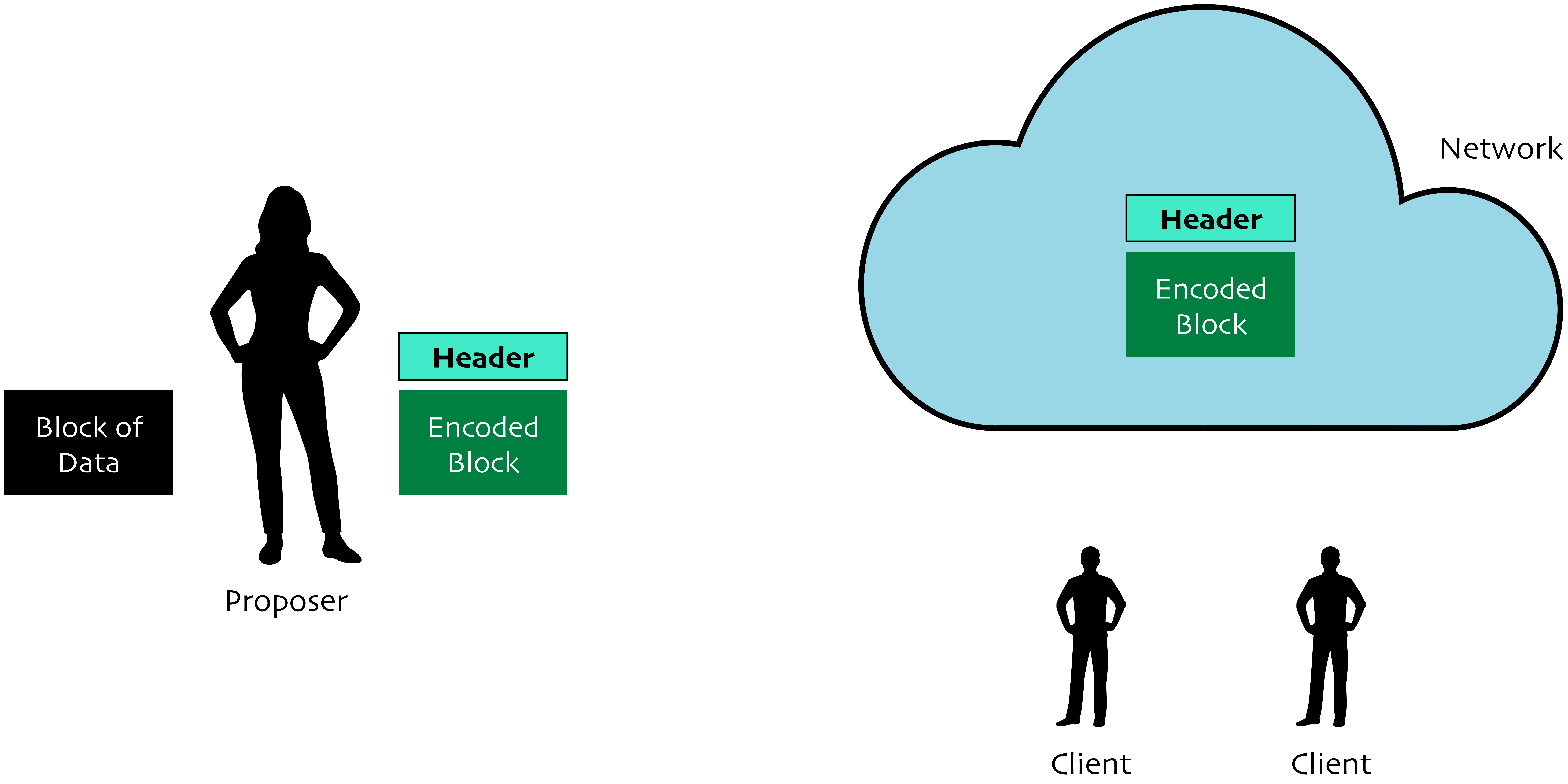
Proposer

Network

Client

Client

Data Availability Sampling



Block of Data

Proposer

Header

Encoded Block

Network

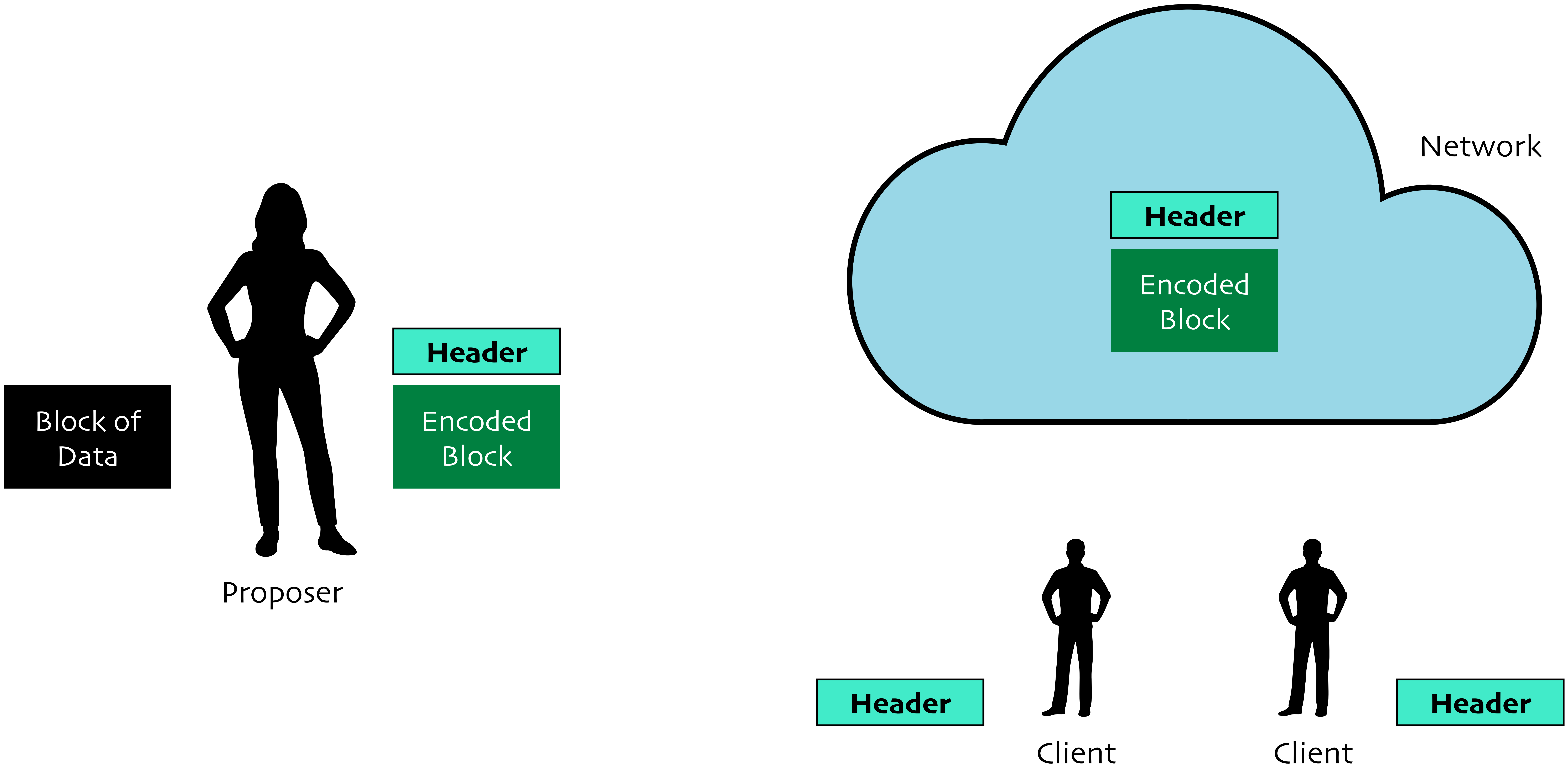
Header

Encoded Block

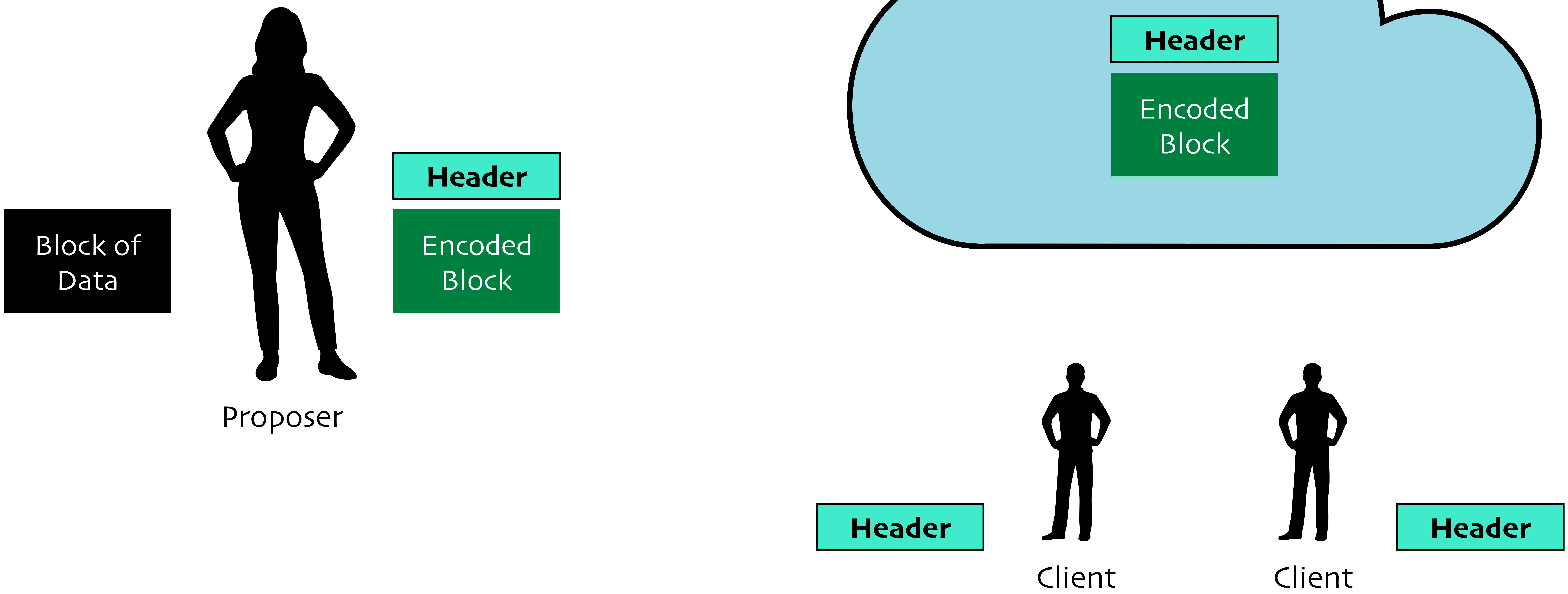
Client

Client

Data Availability Sampling



Data Availability Sampling



Block of Data

Proposer

Header

Encoded Block

Network

Header

Encoded Block

Header

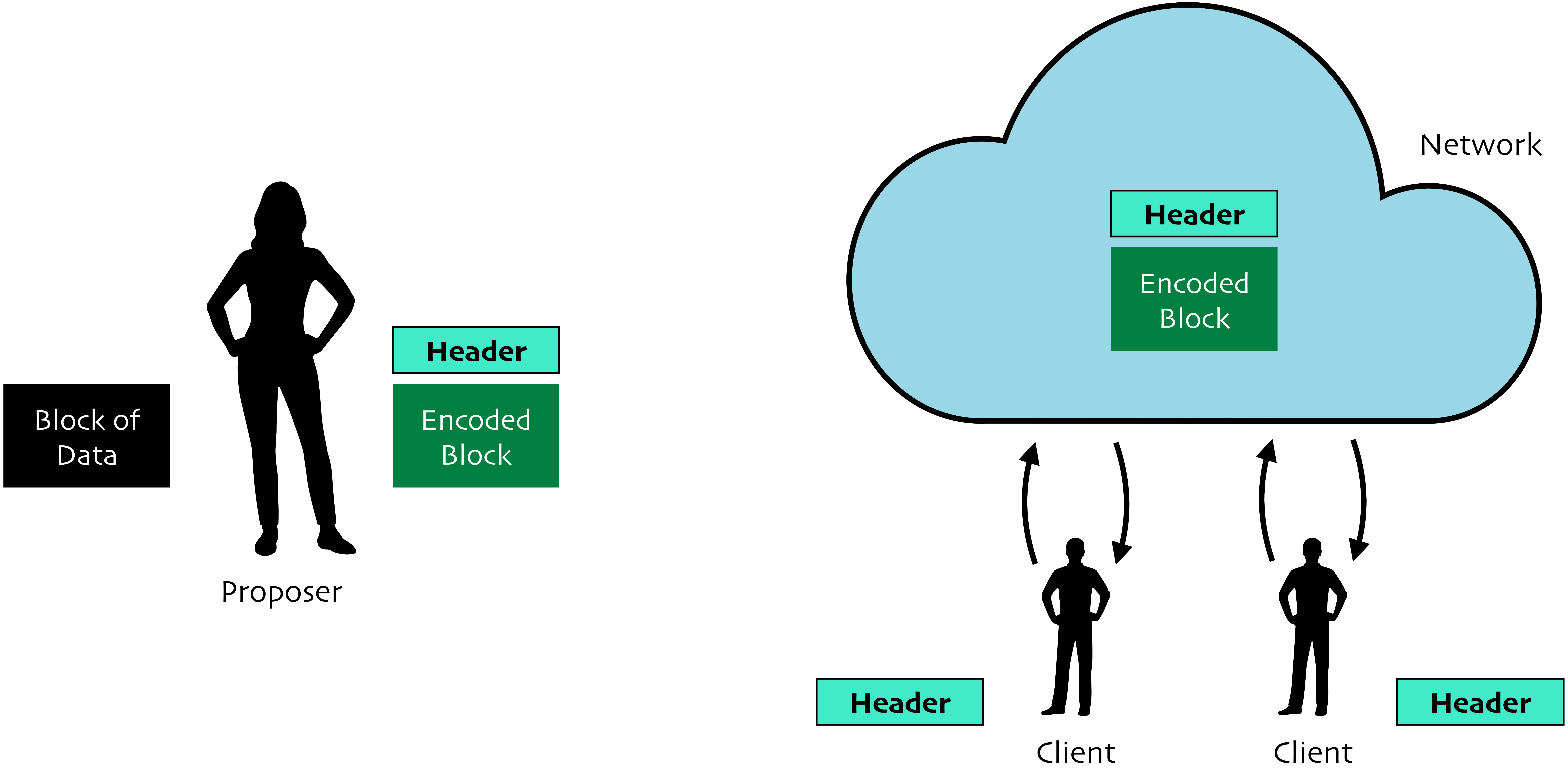
Client

Client

Header

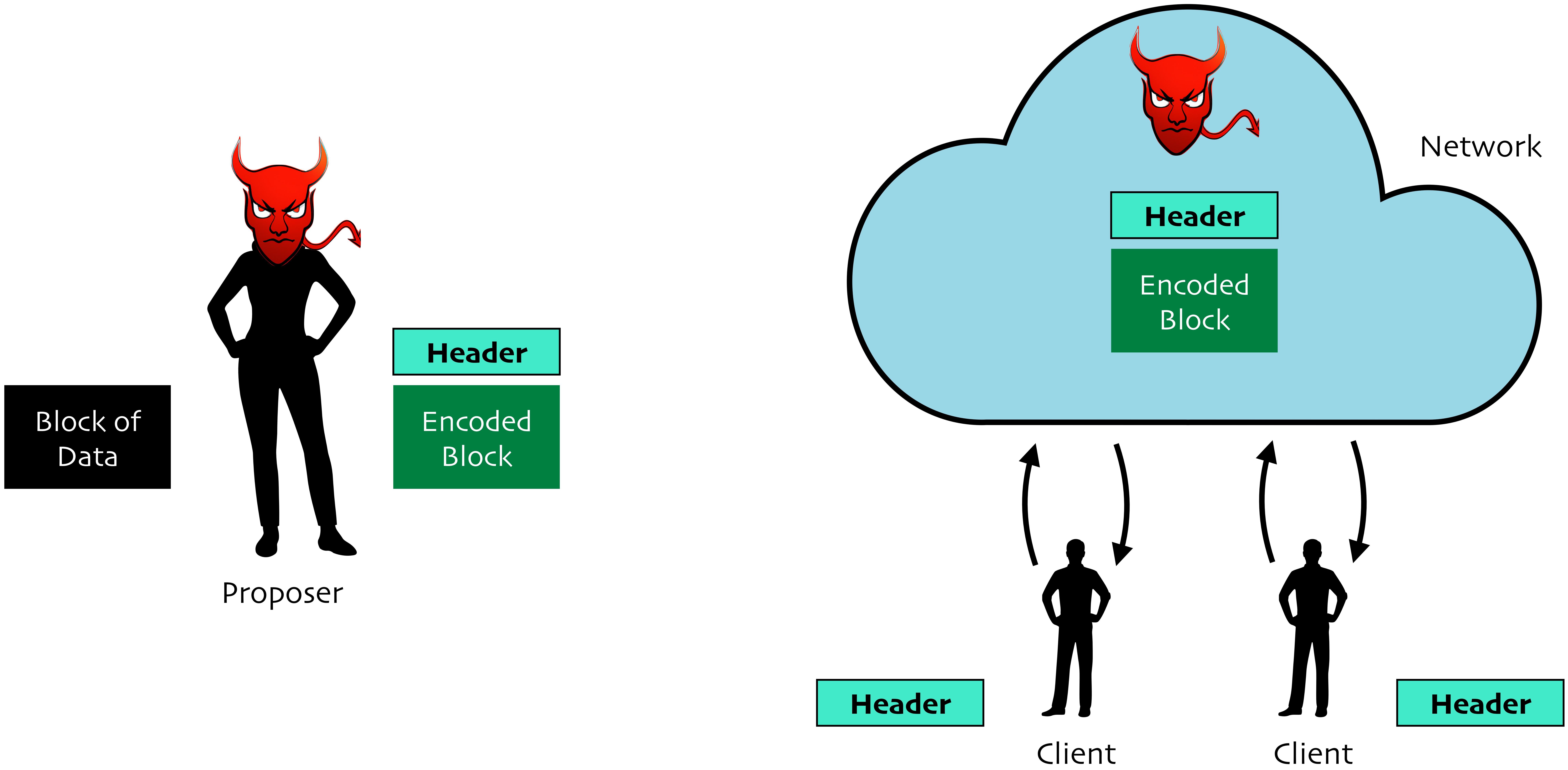
Is data available?

Data Availability Sampling



Is data available?

Data Availability Sampling



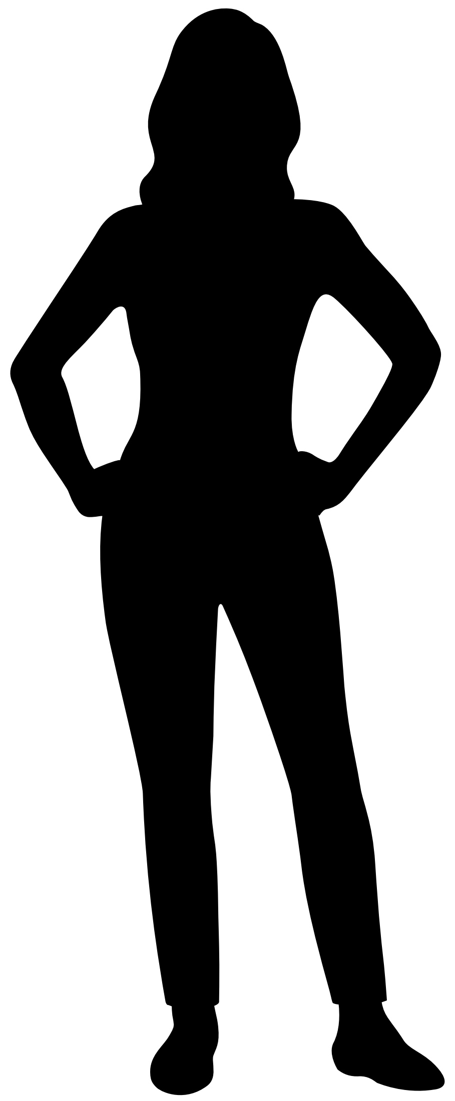
Is data available?

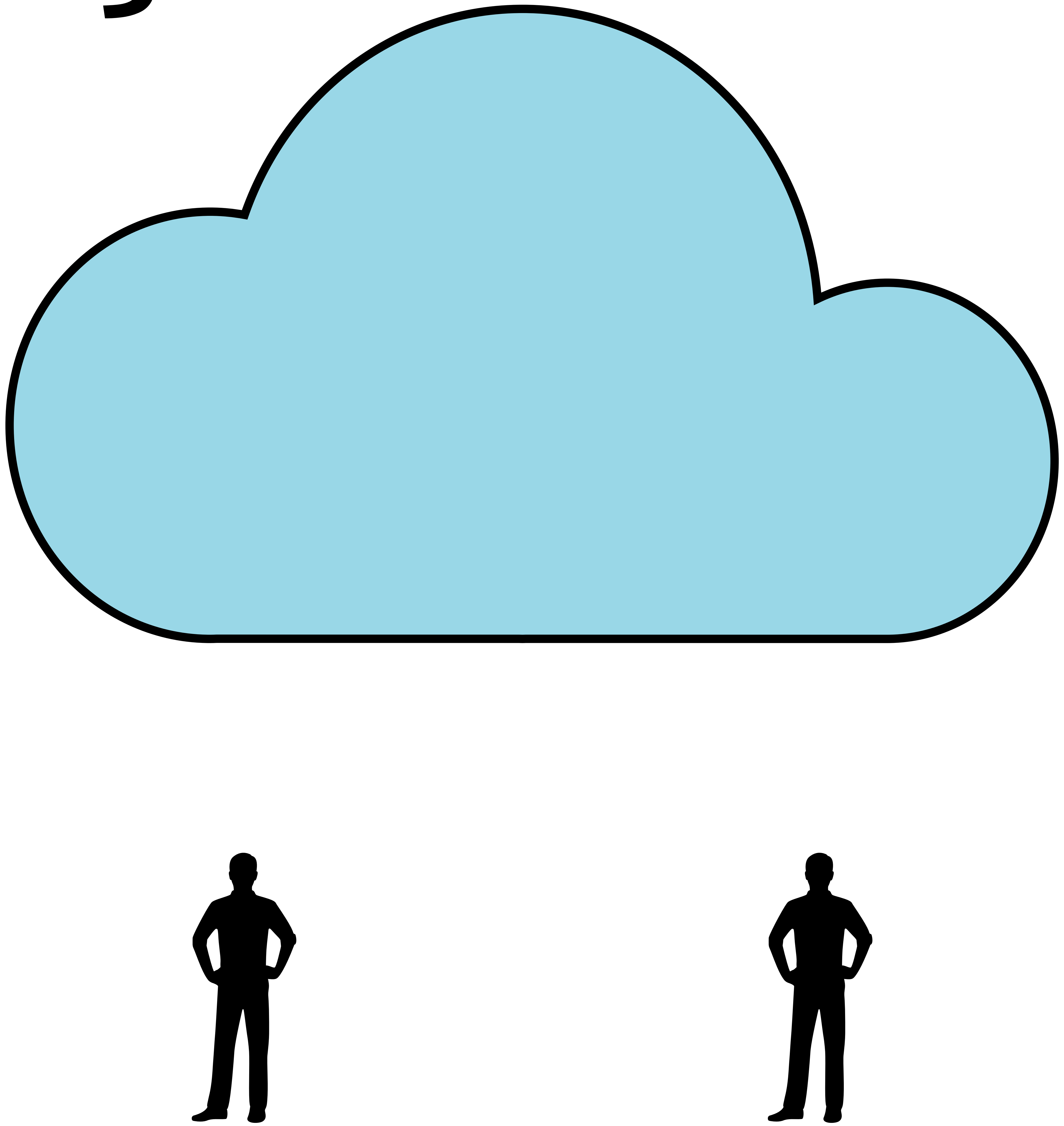
Data Availability Sampling

Naively

Data Availability Sampling

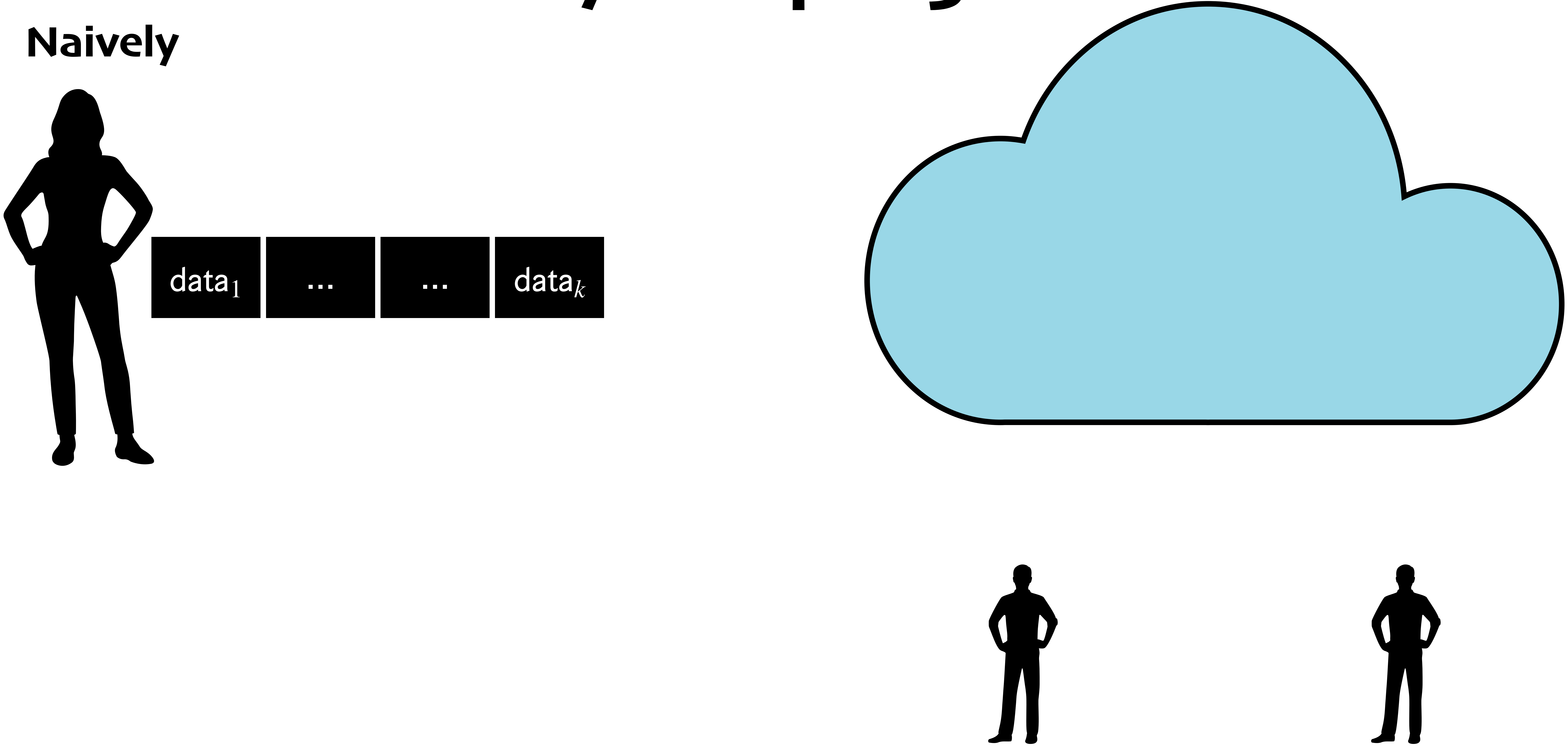
Naively





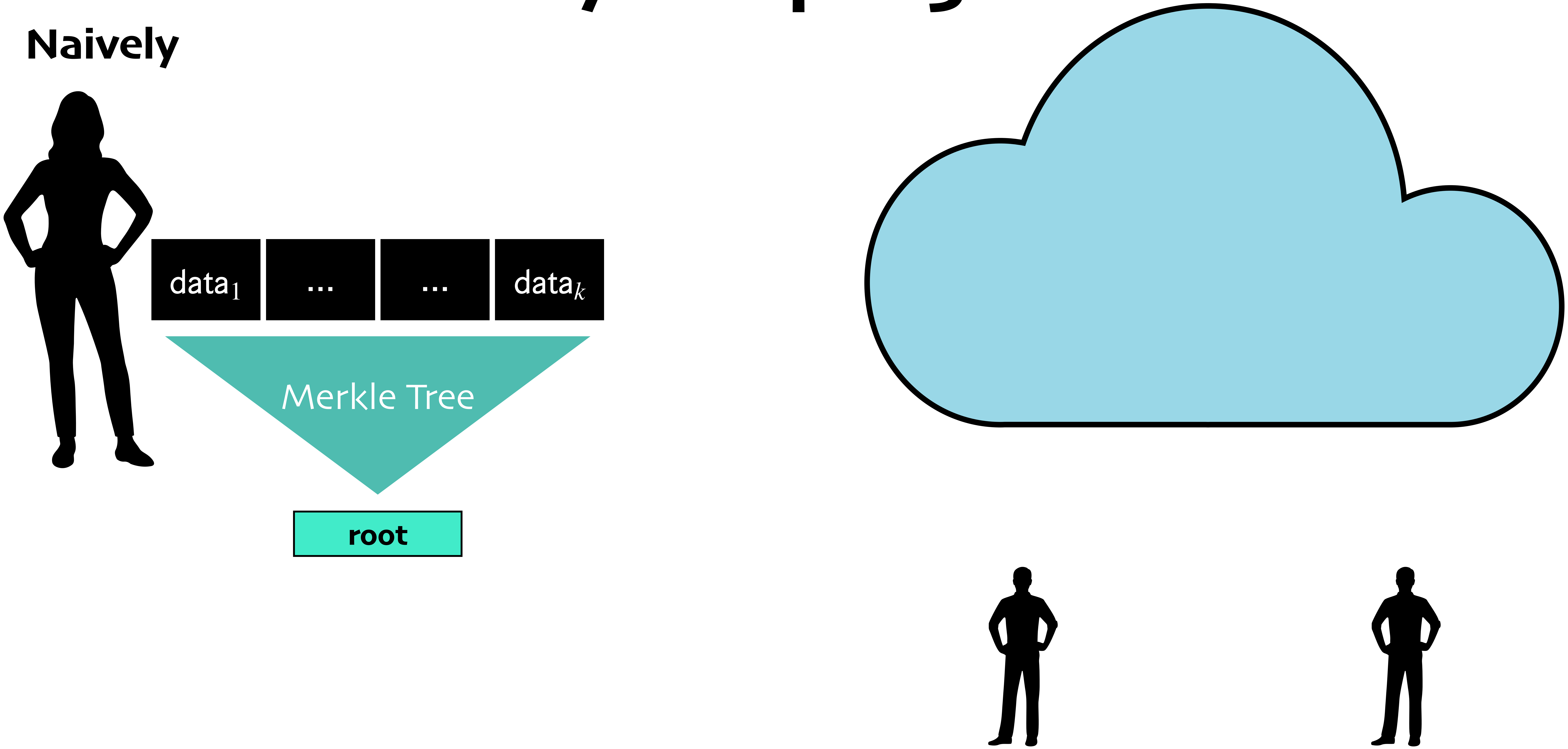
Data Availability Sampling

Naively



Data Availability Sampling

Naively



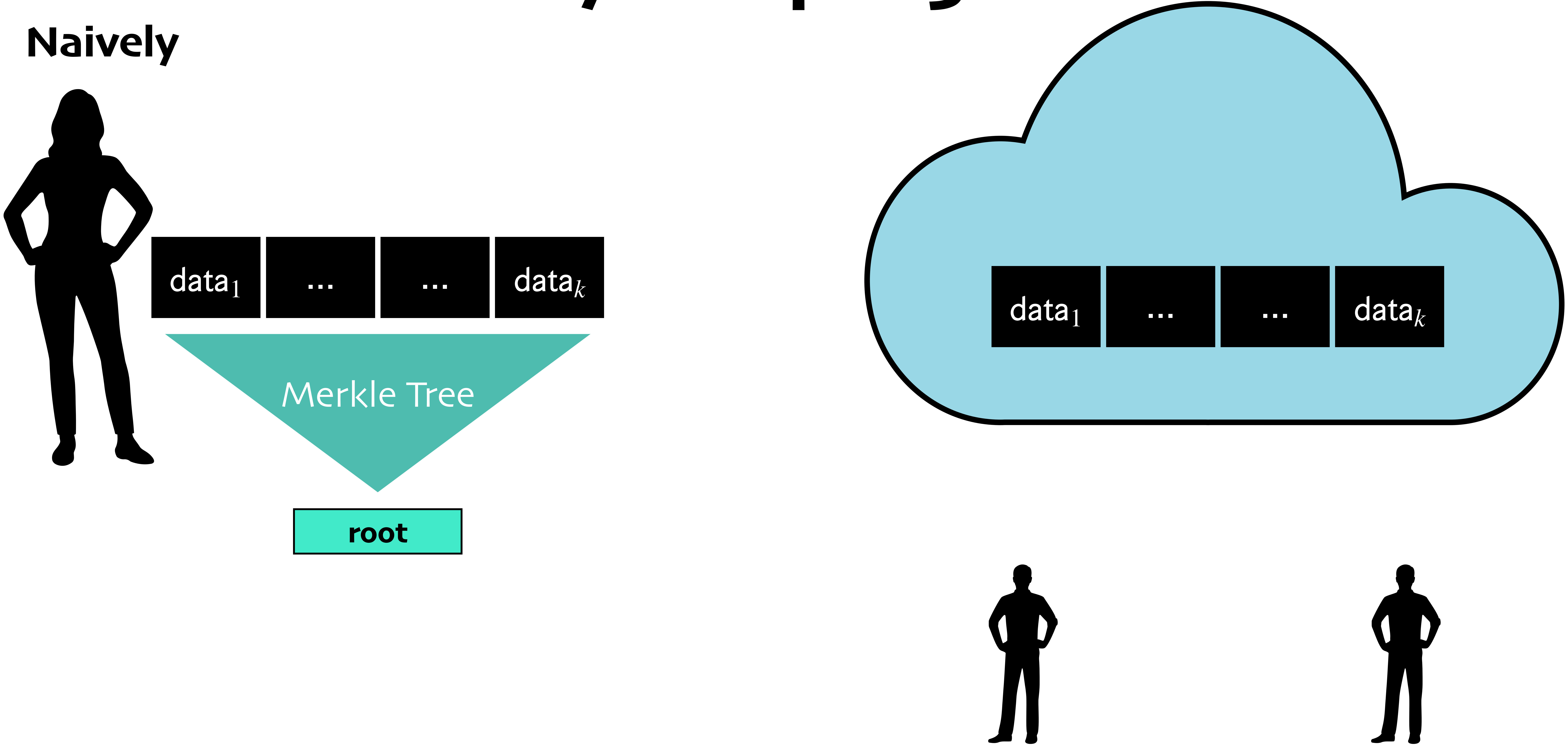
data$_1$ … … data$_k$

Merkle Tree

root

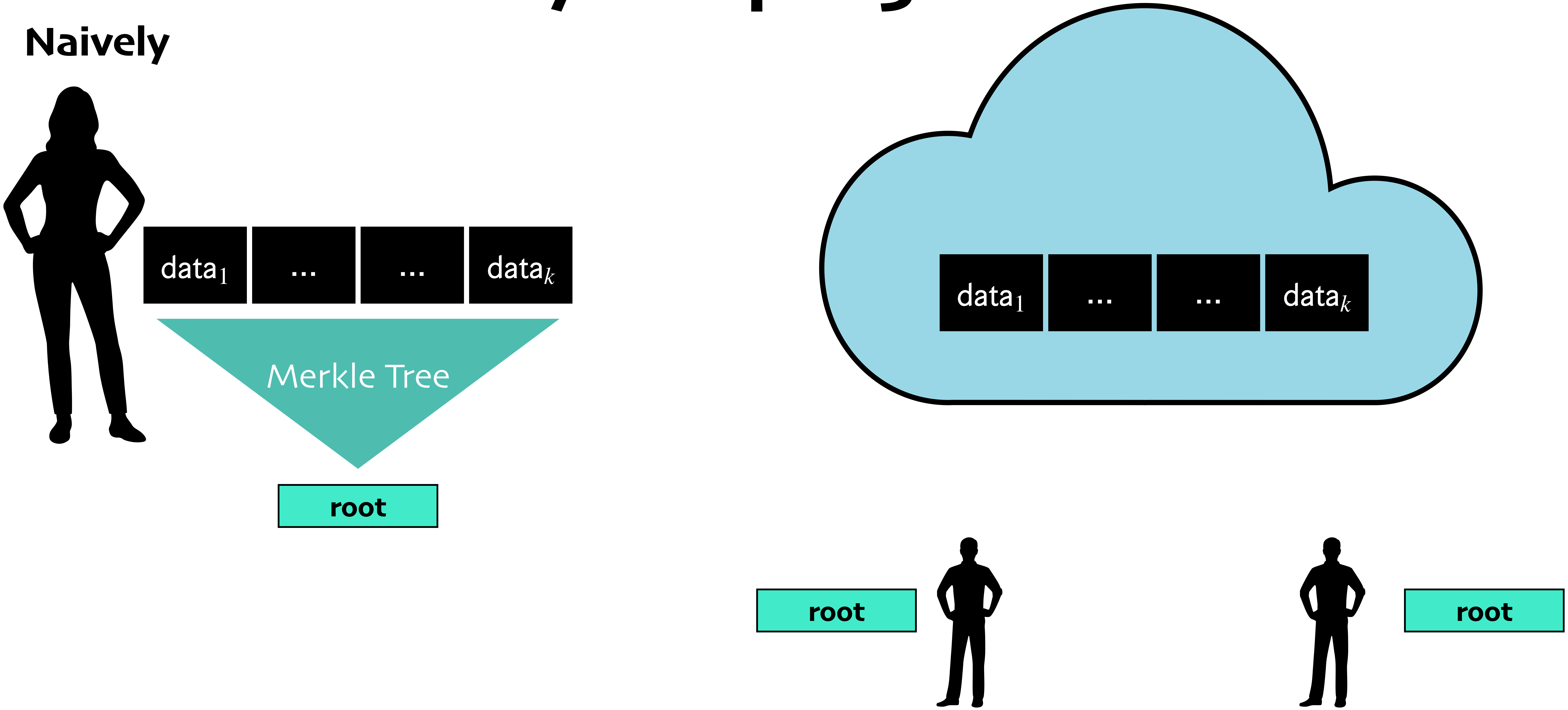
Data Availability Sampling

Naively



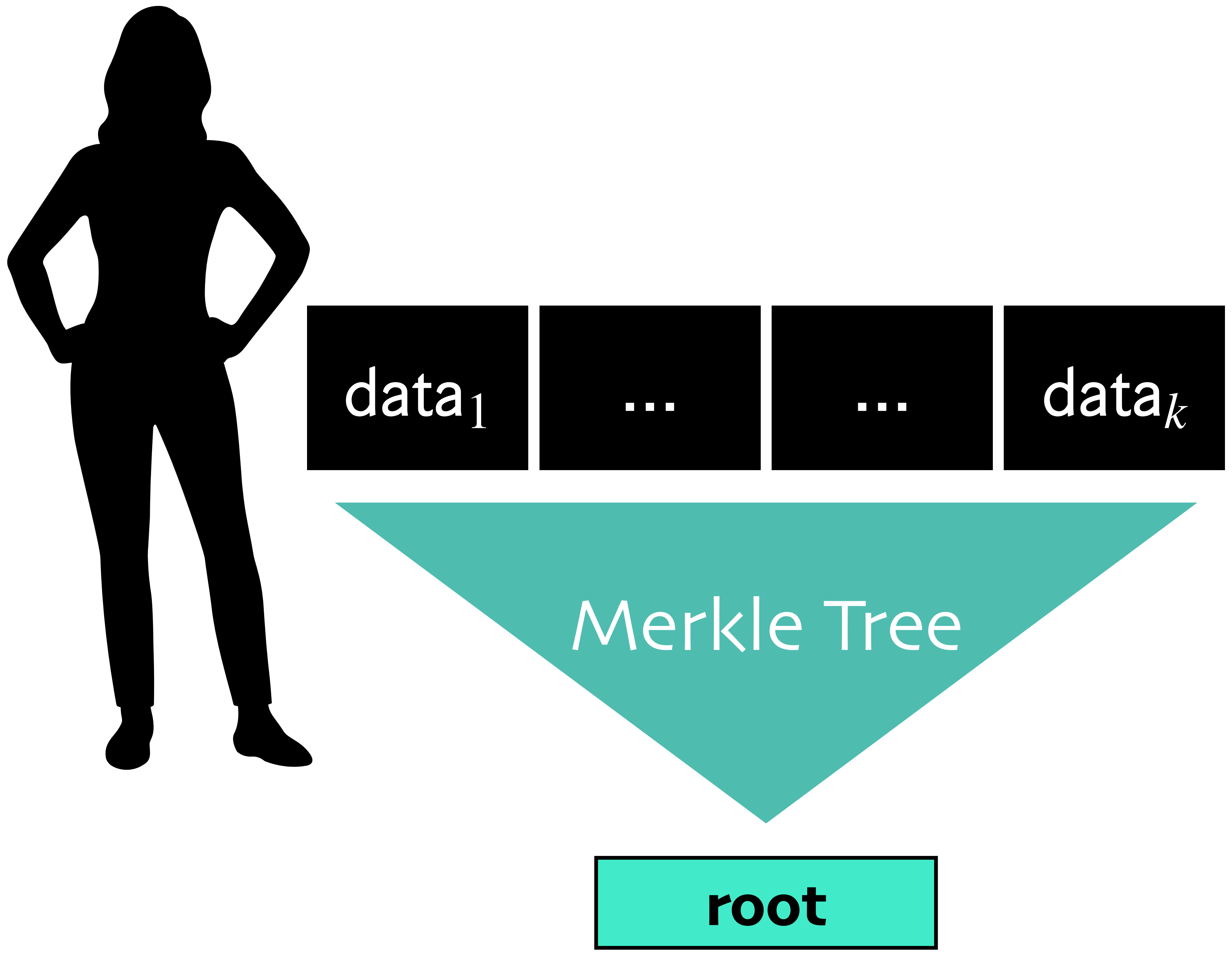
Data Availability Sampling

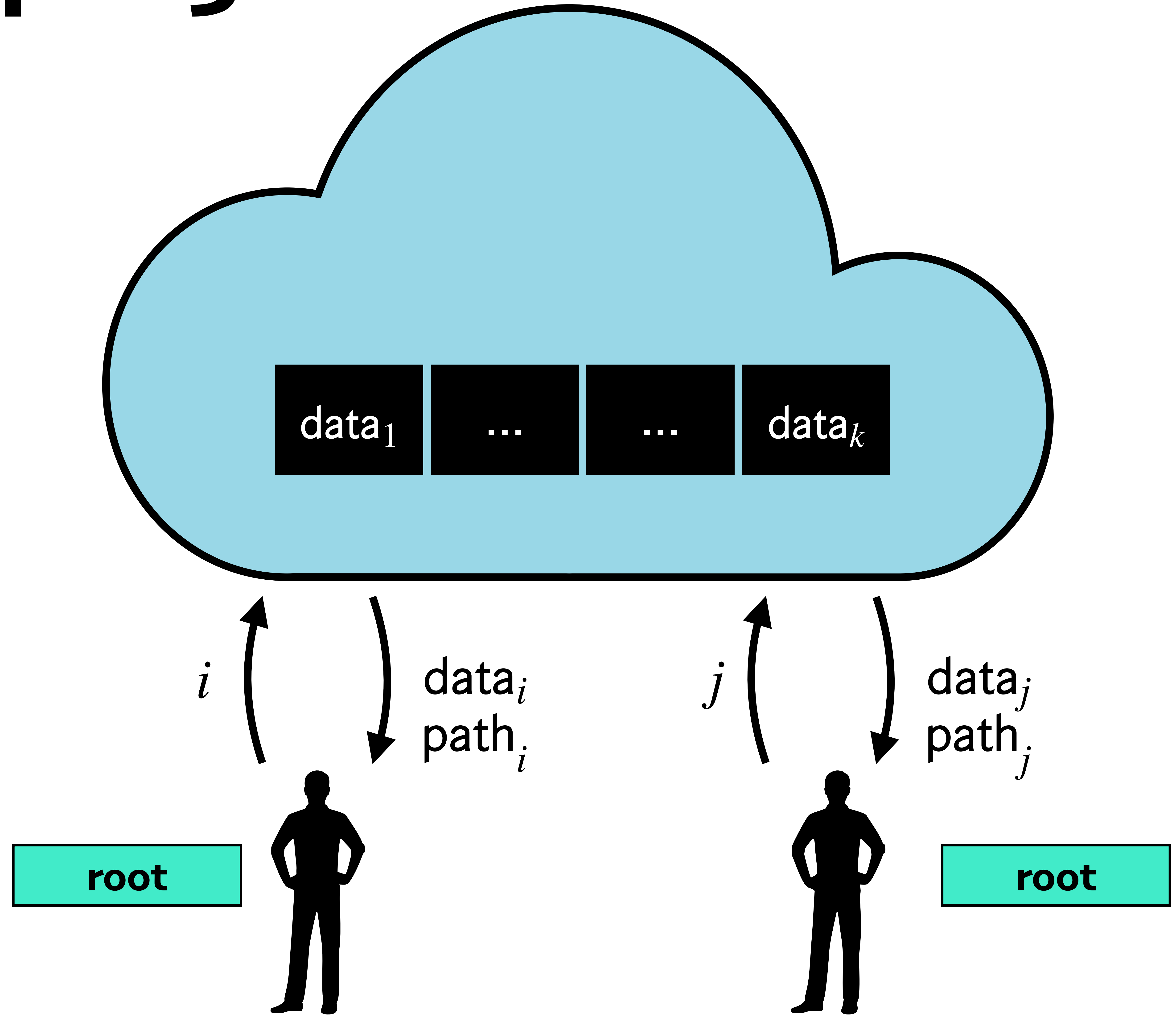
Naively



Data Availability Sampling

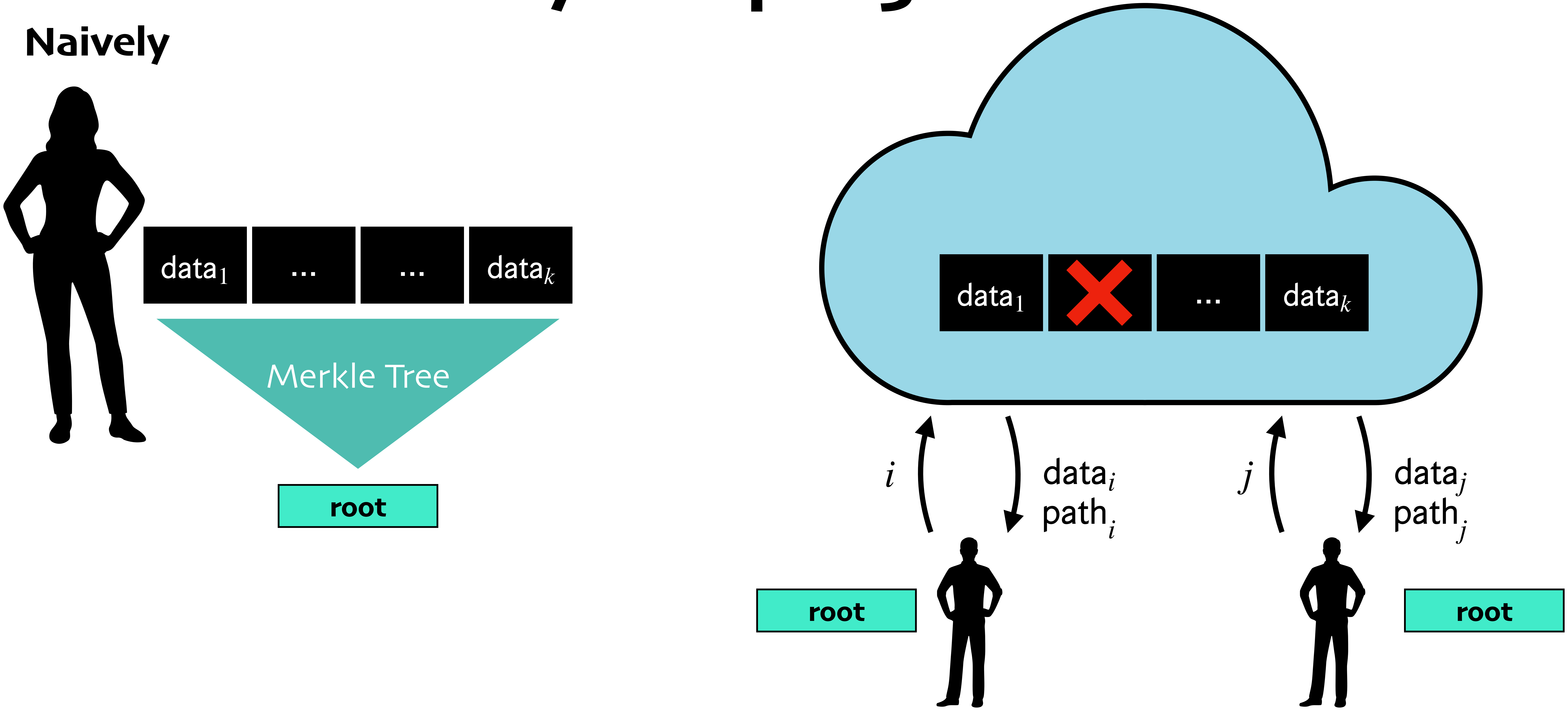
Naively





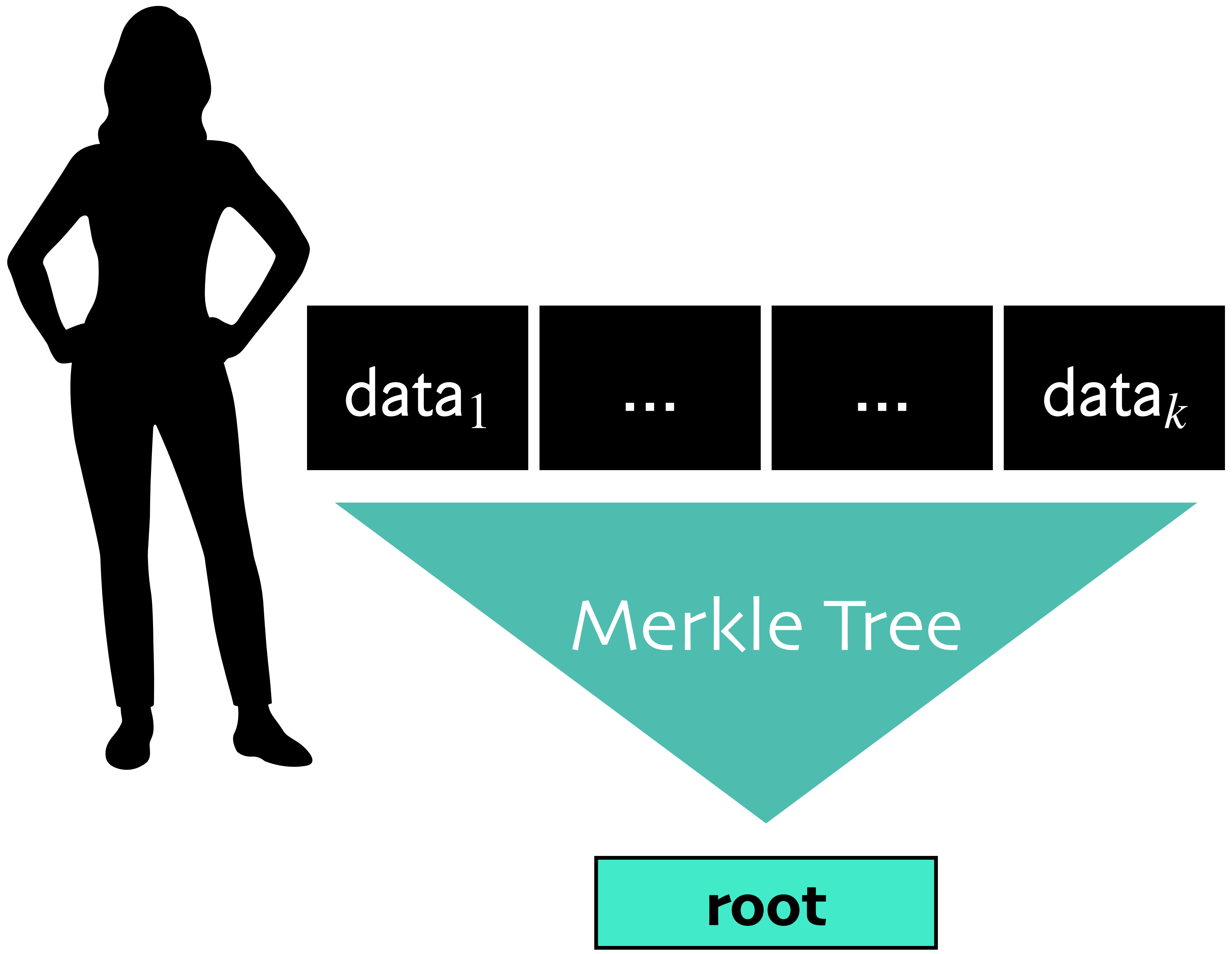
Data Availability Sampling

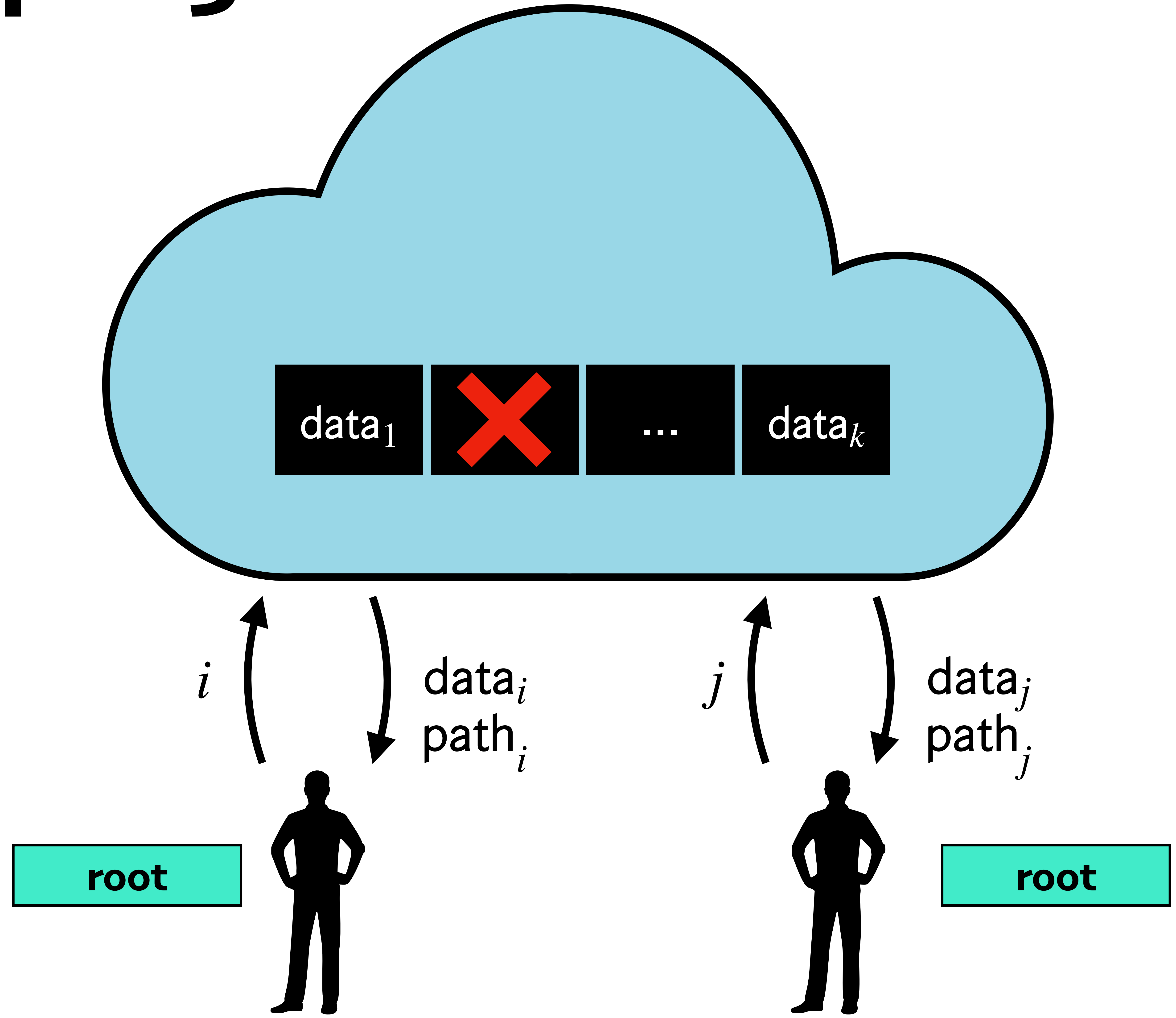
Naively



Data Availability Sampling

Naively

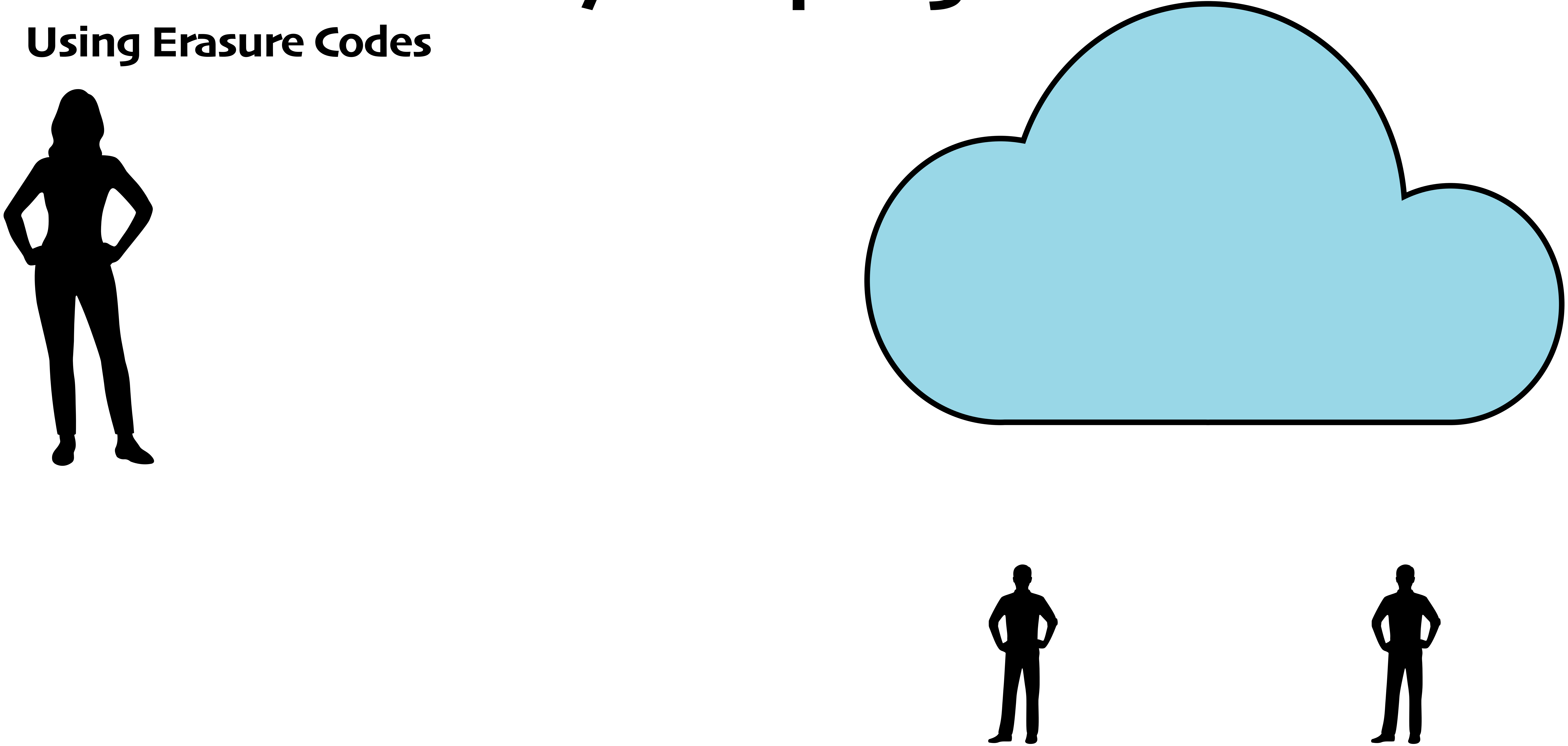




Bad Soundness

Data Availability Sampling

Using Erasure Codes



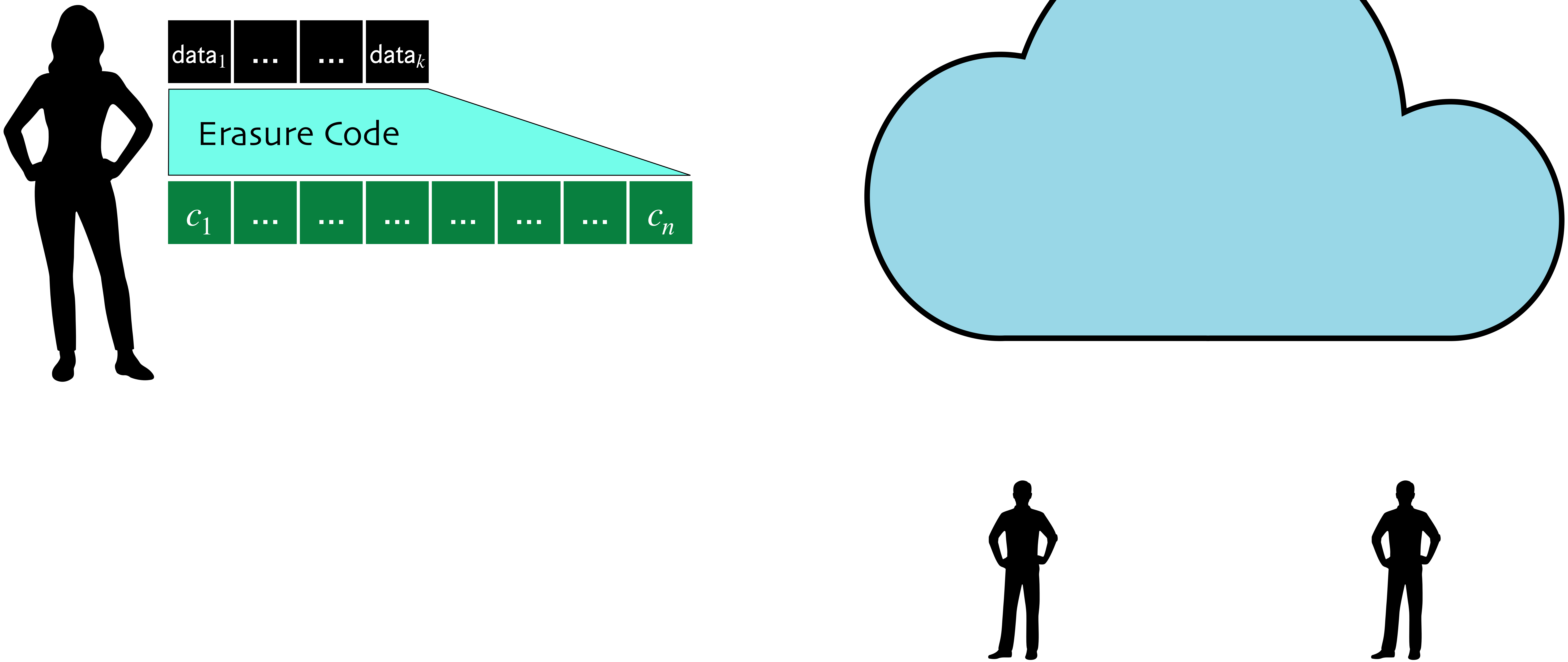
Data Availability Sampling

Using Erasure Codes

data$_1$ ... ... data$_k$

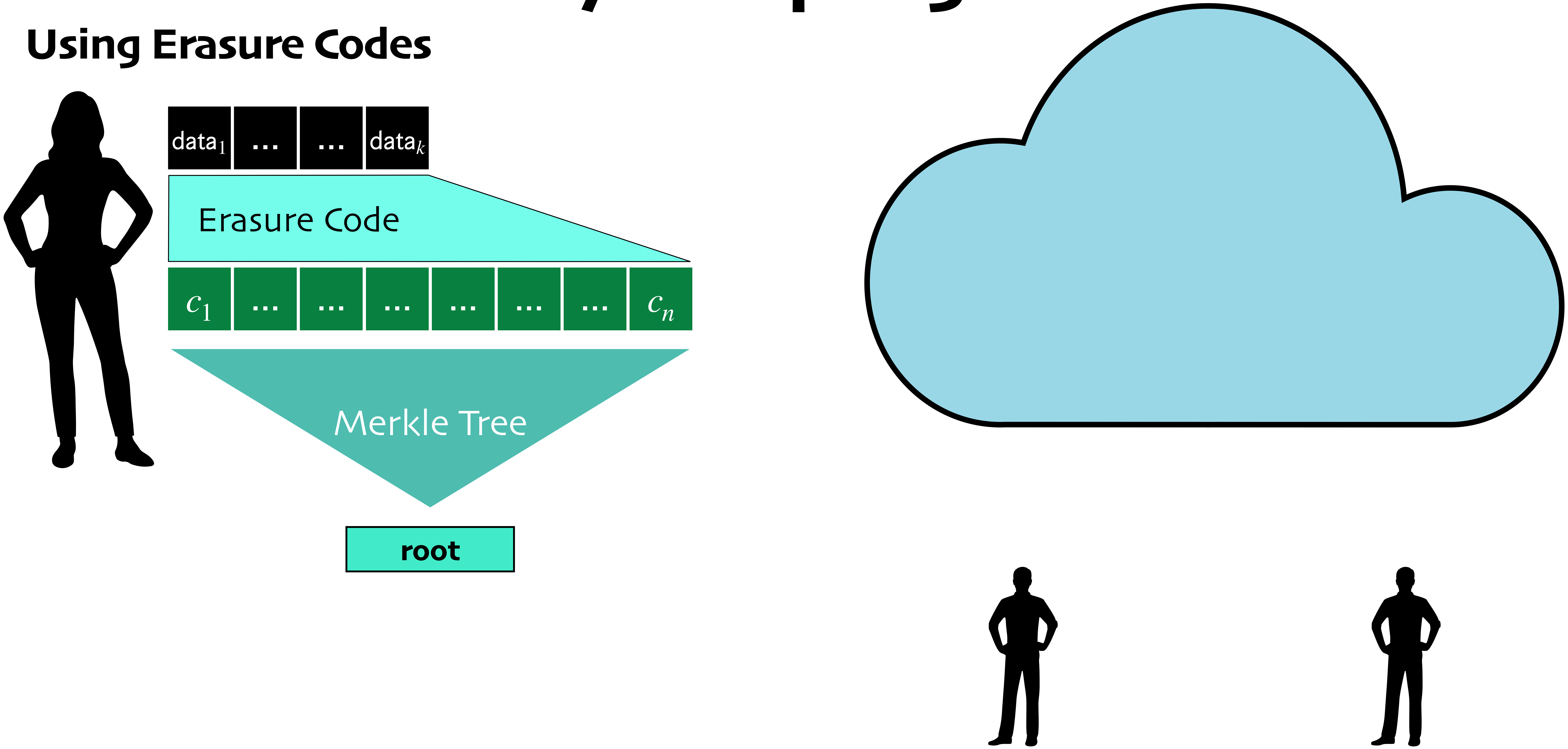
Data Availability Sampling

Using Erasure Codes



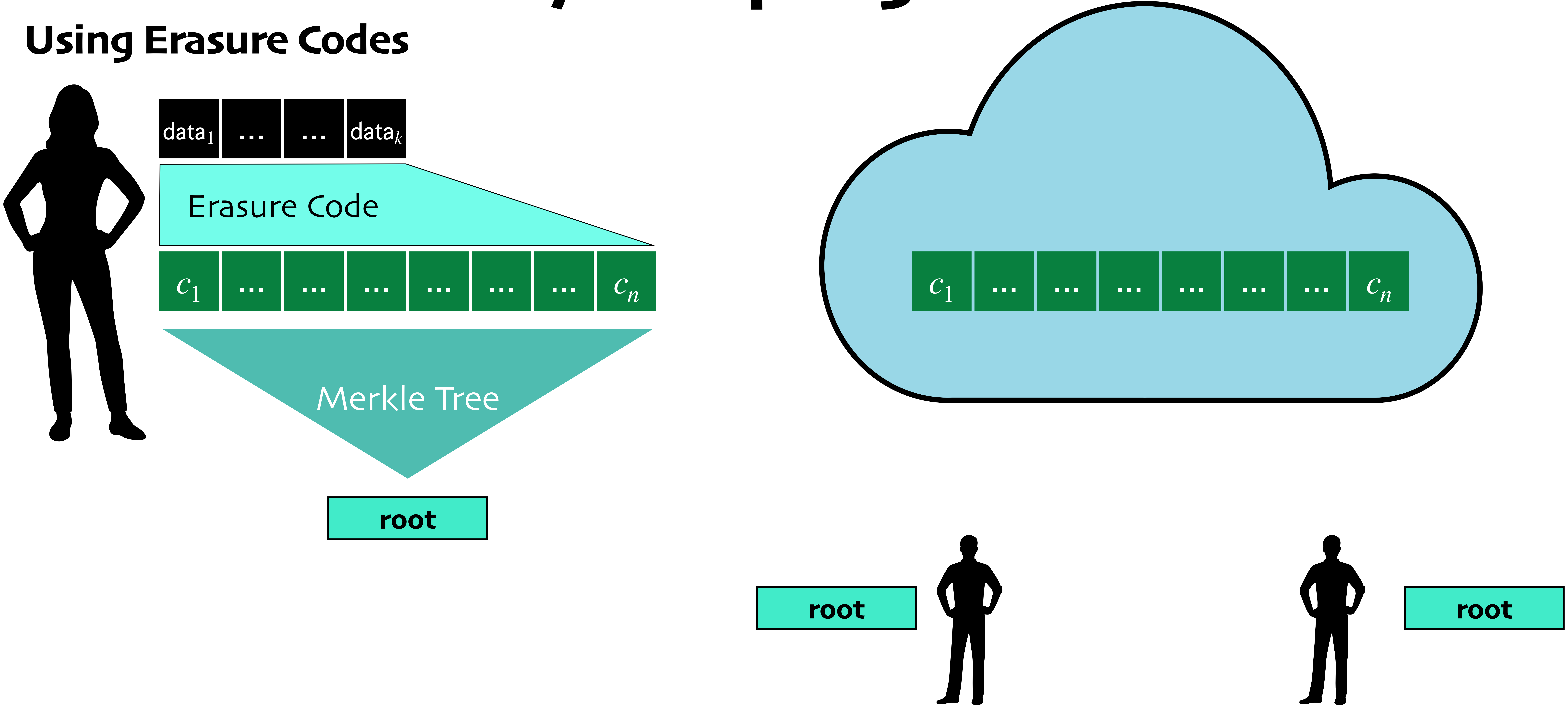
Data Availability Sampling

Using Erasure Codes



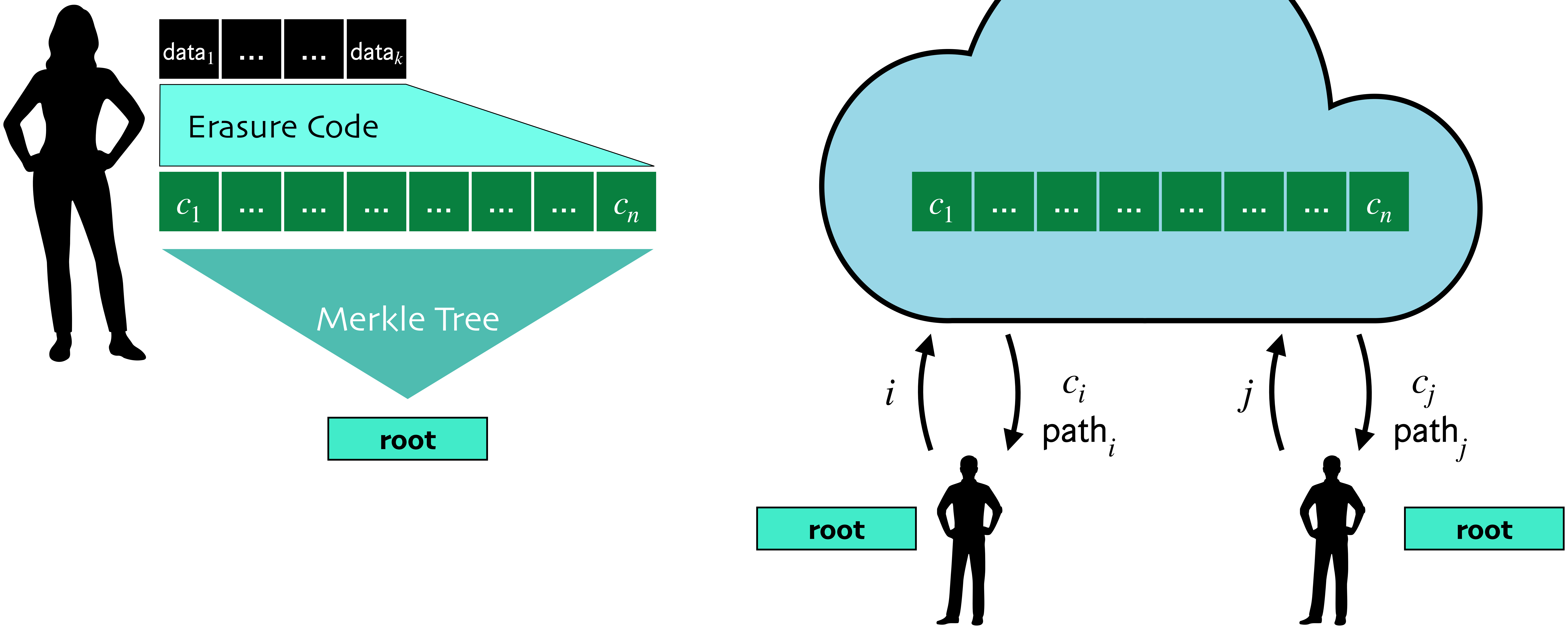
Data Availability Sampling

Using Erasure Codes



Data Availability Sampling

Using Erasure Codes



data$_1$ … … data$_k$

Erasure Code

c_1 … … … … … … c_n

Merkle Tree

root

c_1 … … … … … … c_n

i

c_i
path$_i$

j

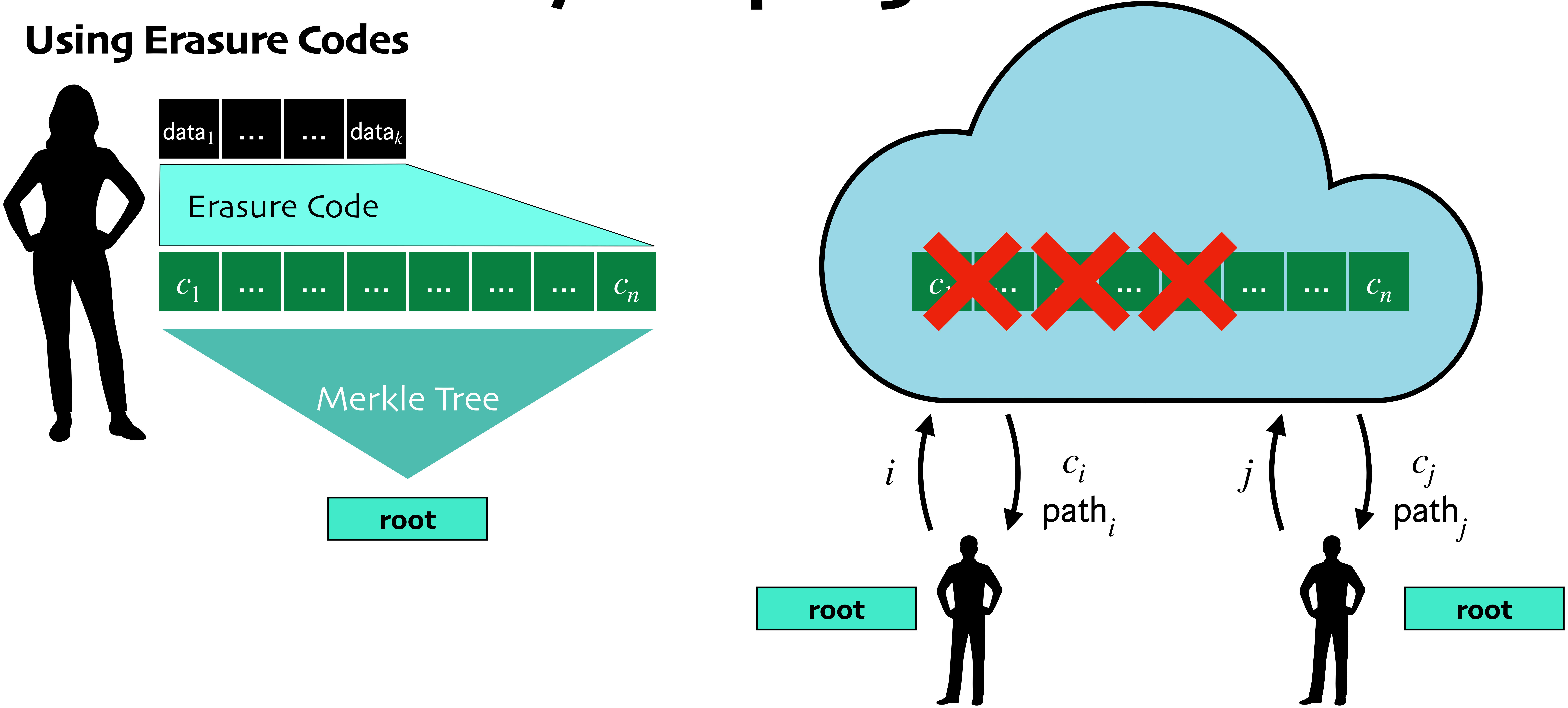
c_j
path$_j$

root

root

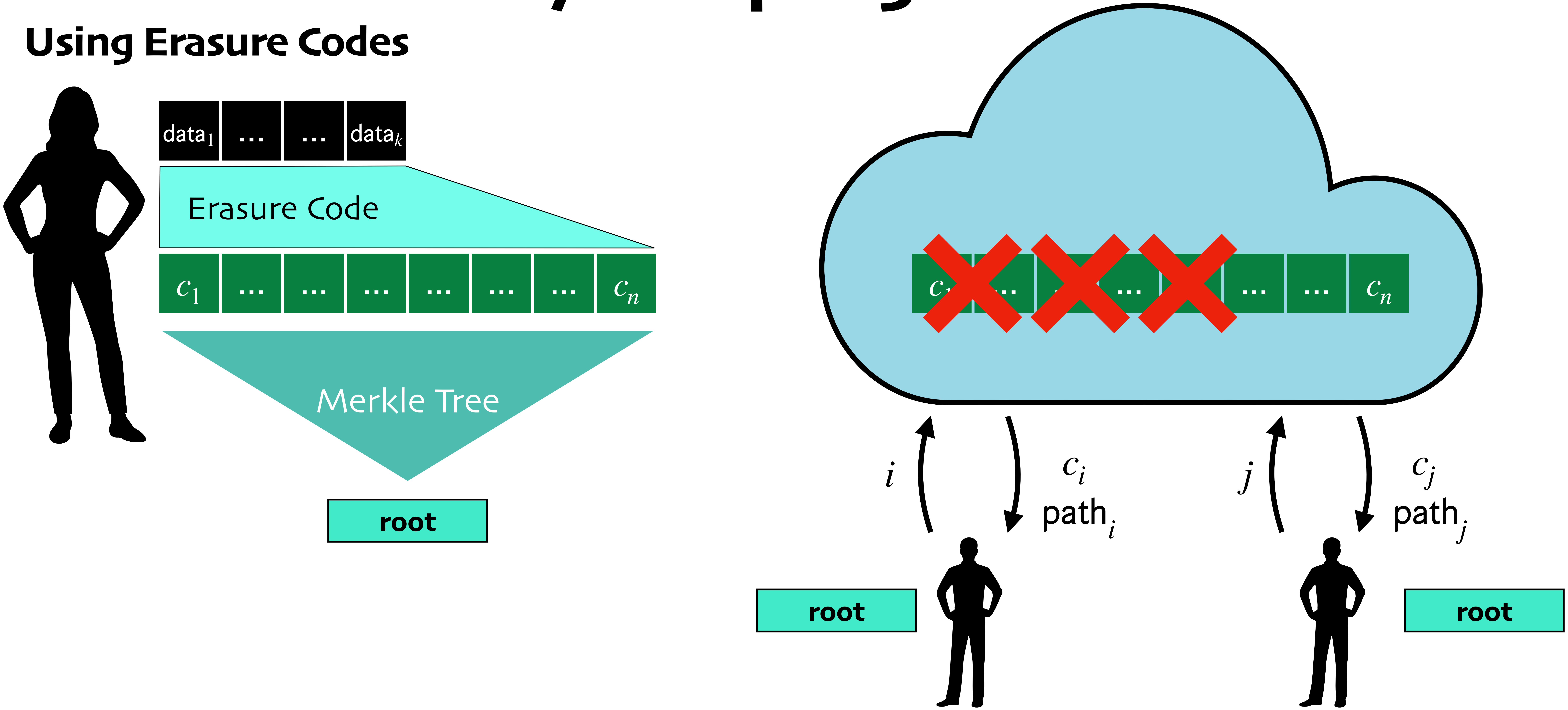
Data Availability Sampling

Using Erasure Codes



Data Availability Sampling

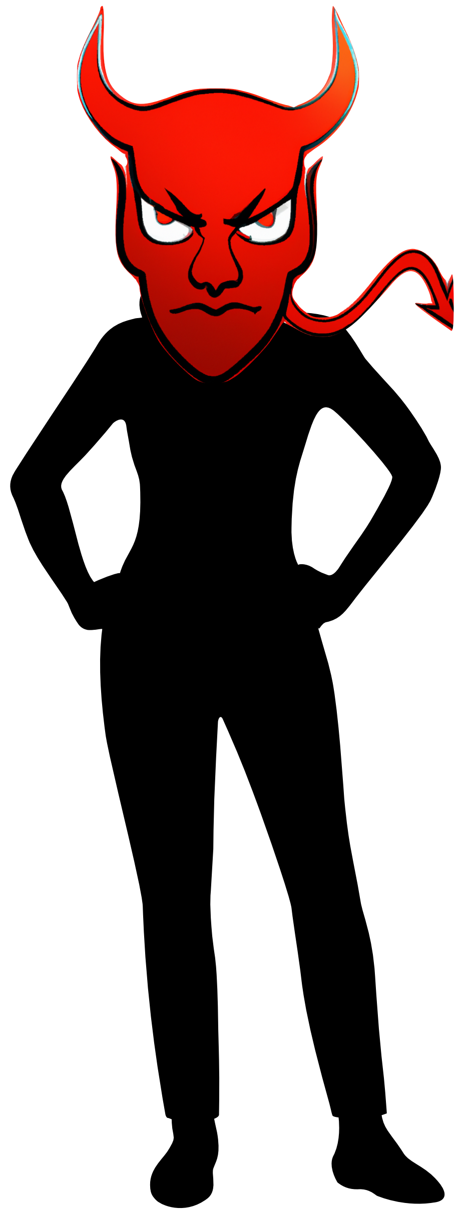
Using Erasure Codes

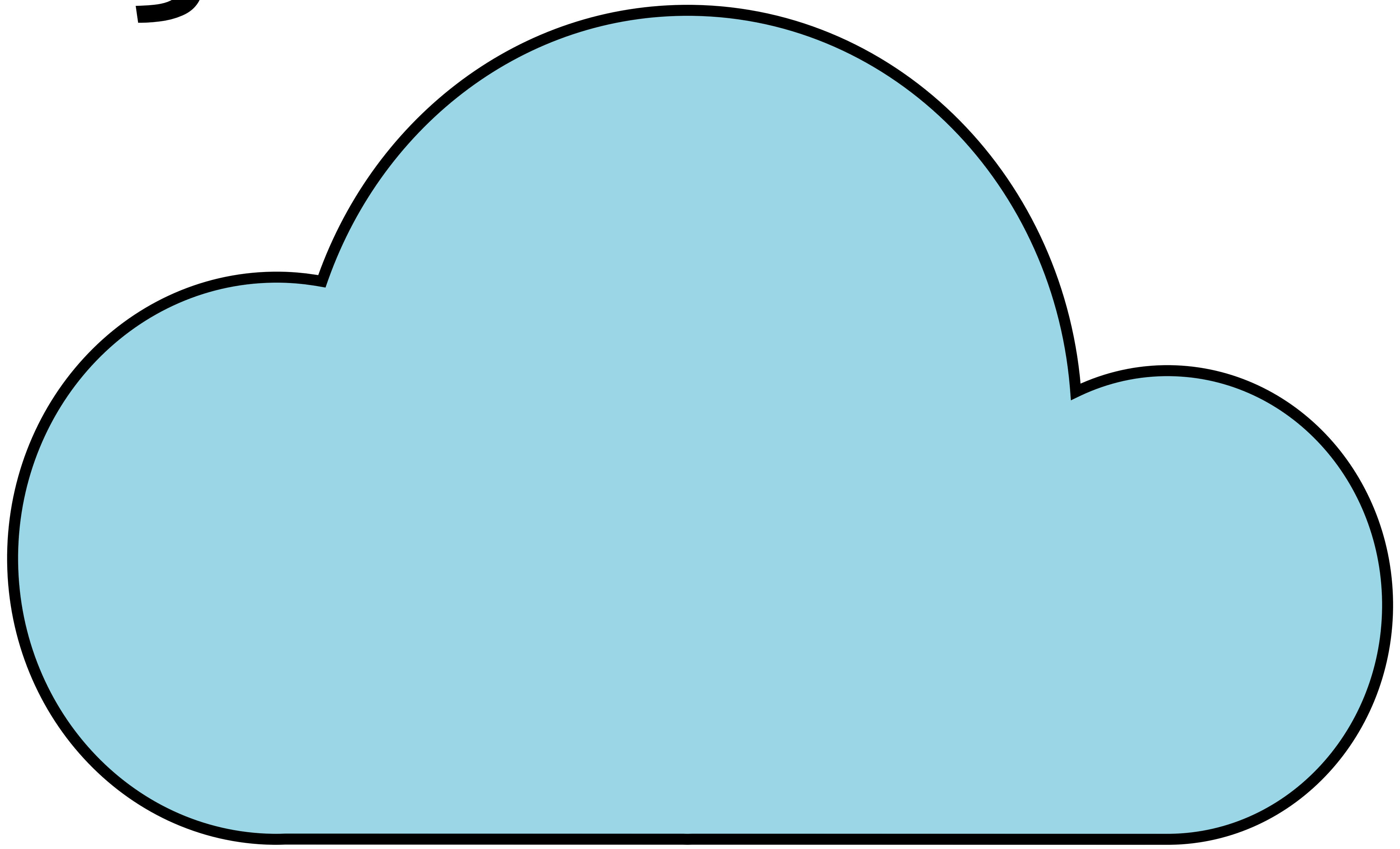


Better Soundness

Data Availability Sampling

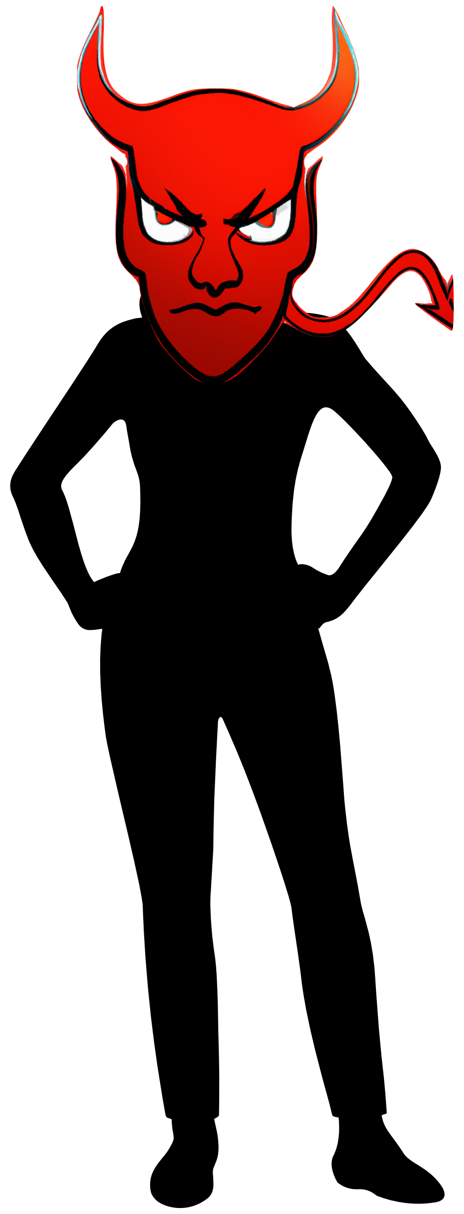
Using Erasure Codes - Inconsistency





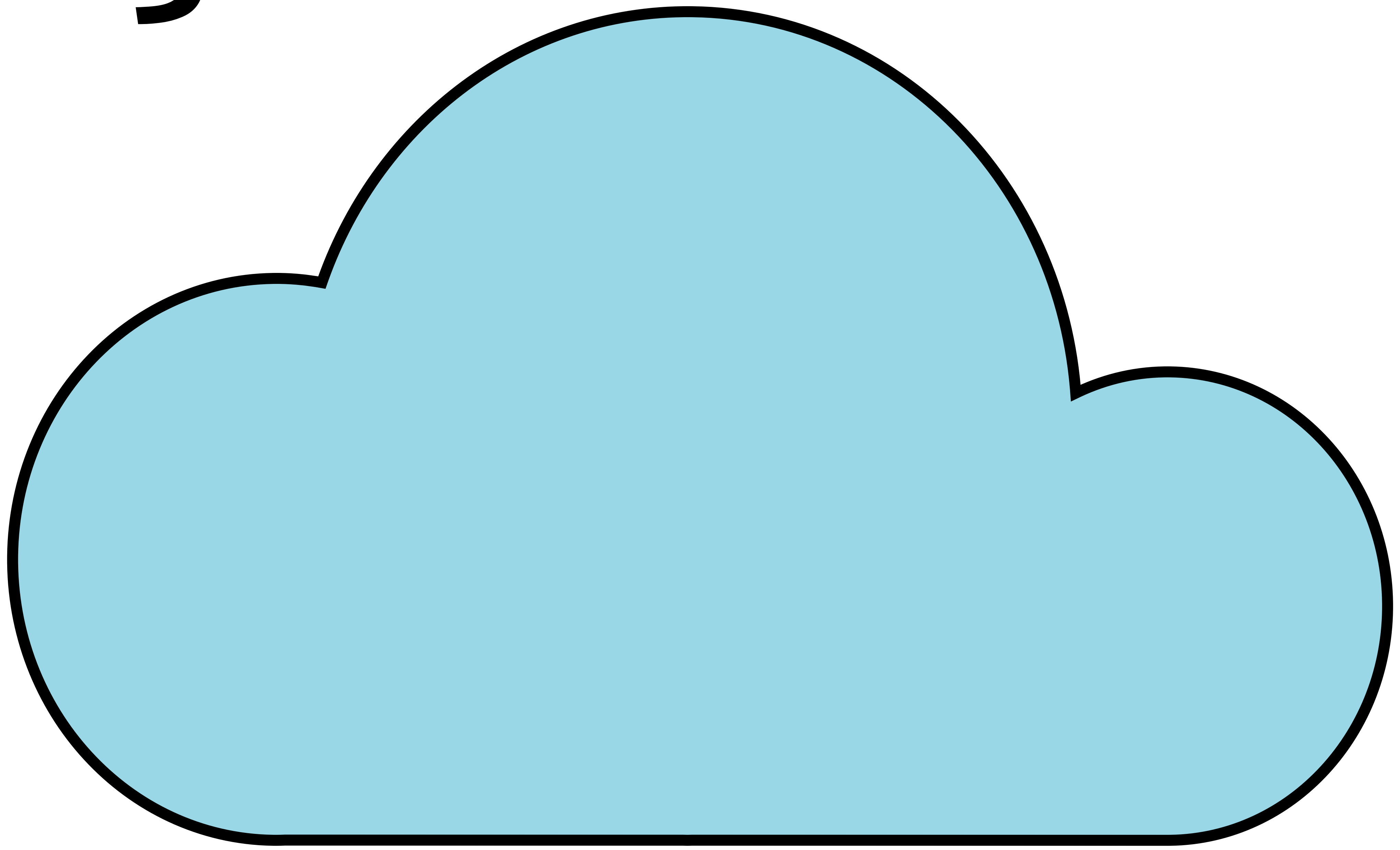
Data Availability Sampling

Using Erasure Codes - Inconsistency



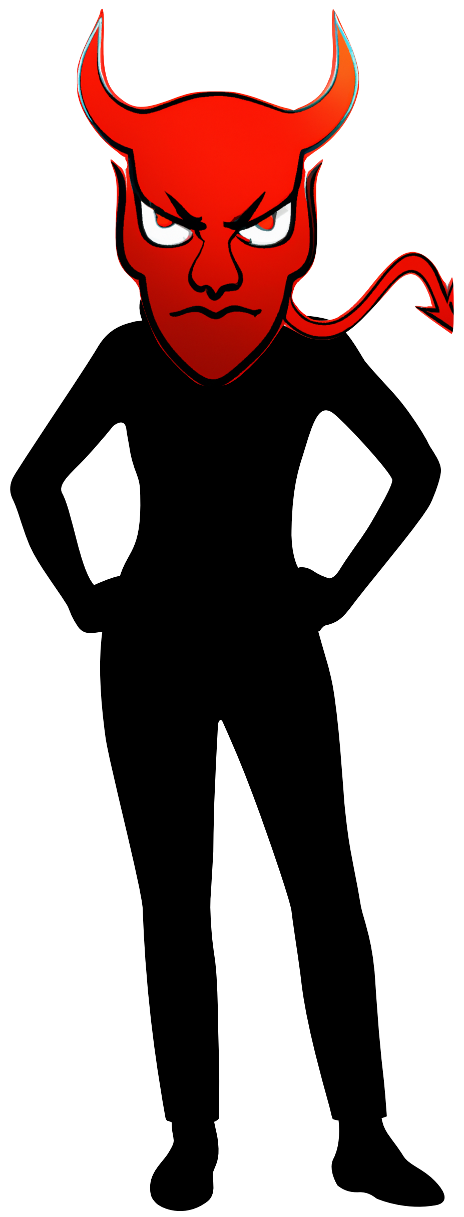
data_1

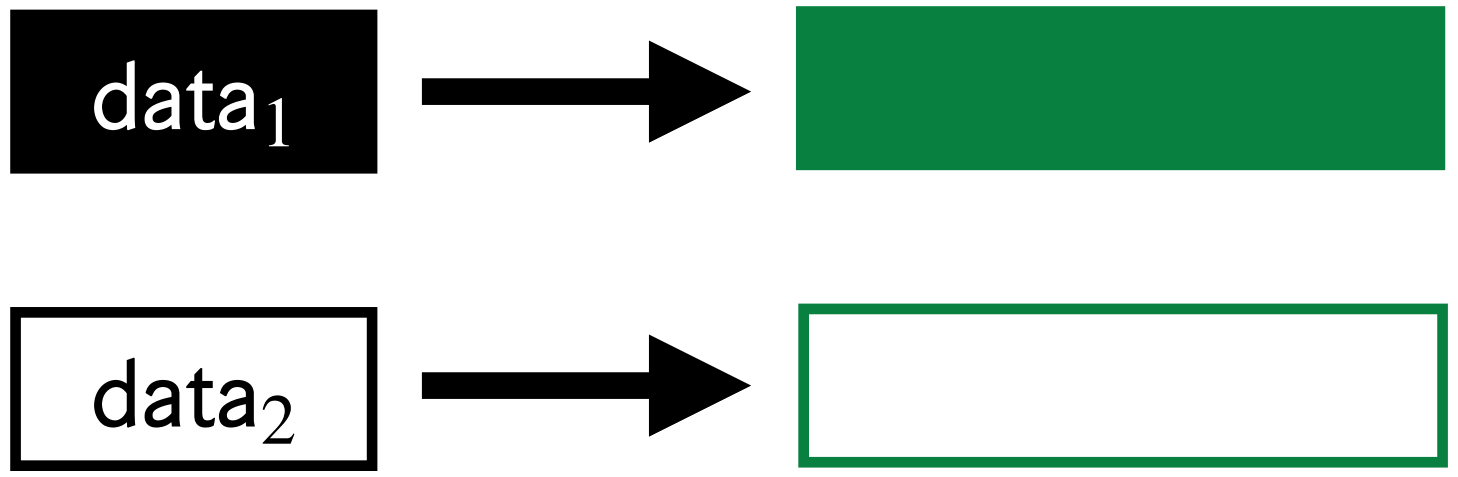
data_2



Data Availability Sampling

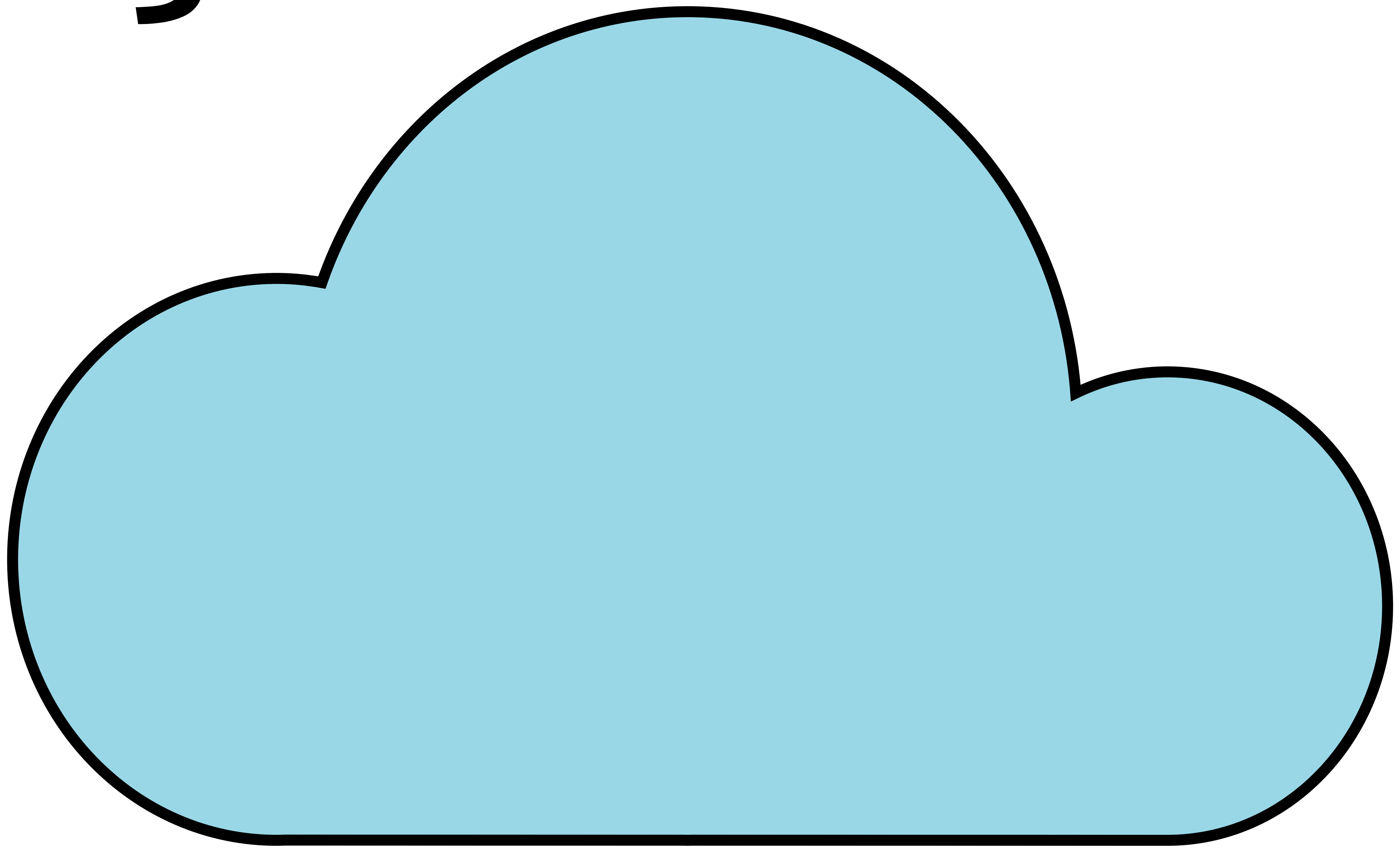
Using Erasure Codes - Inconsistency





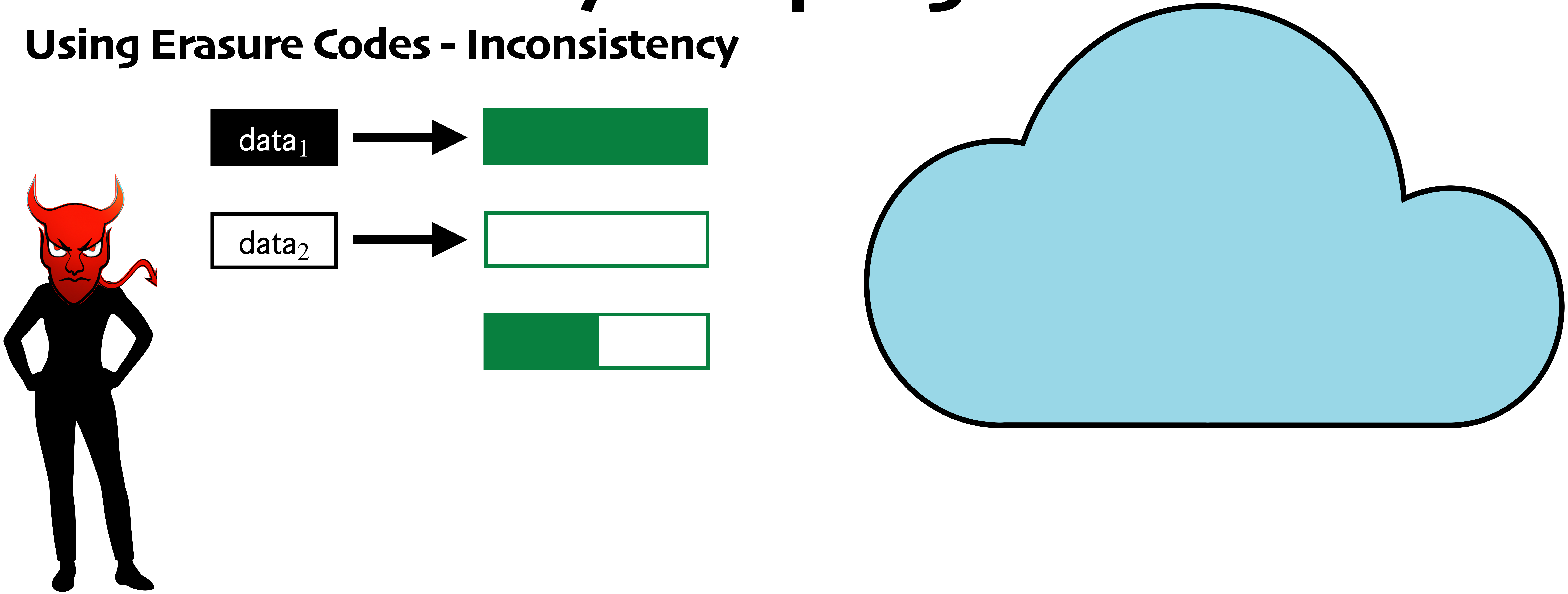
data_1

data_2



Data Availability Sampling

Using Erasure Codes - Inconsistency

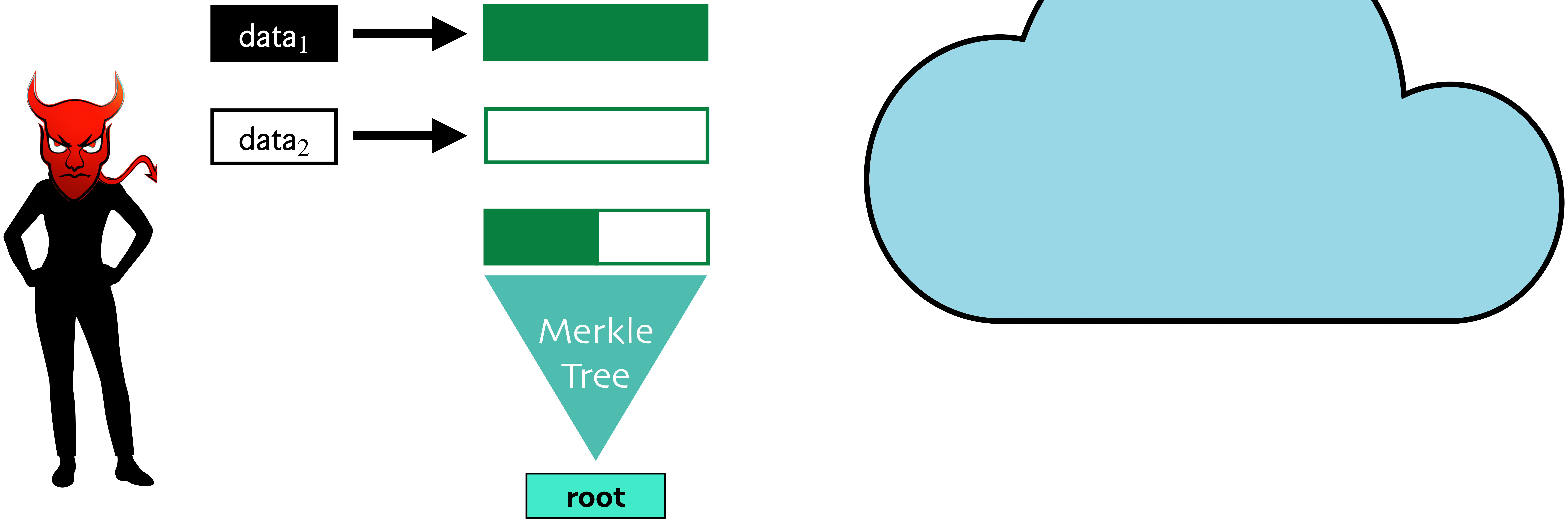


data_1

data_2

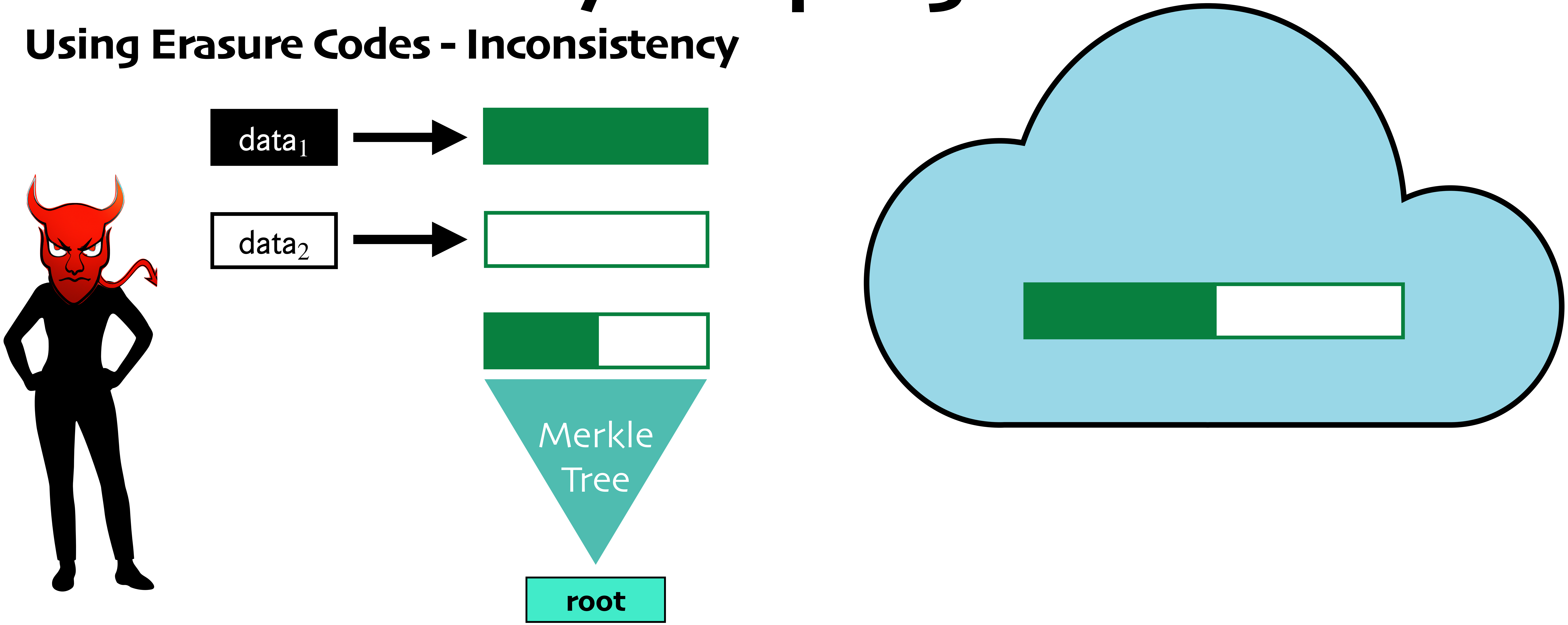
Data Availability Sampling

Using Erasure Codes - Inconsistency



Data Availability Sampling

Using Erasure Codes - Inconsistency



data$_1$

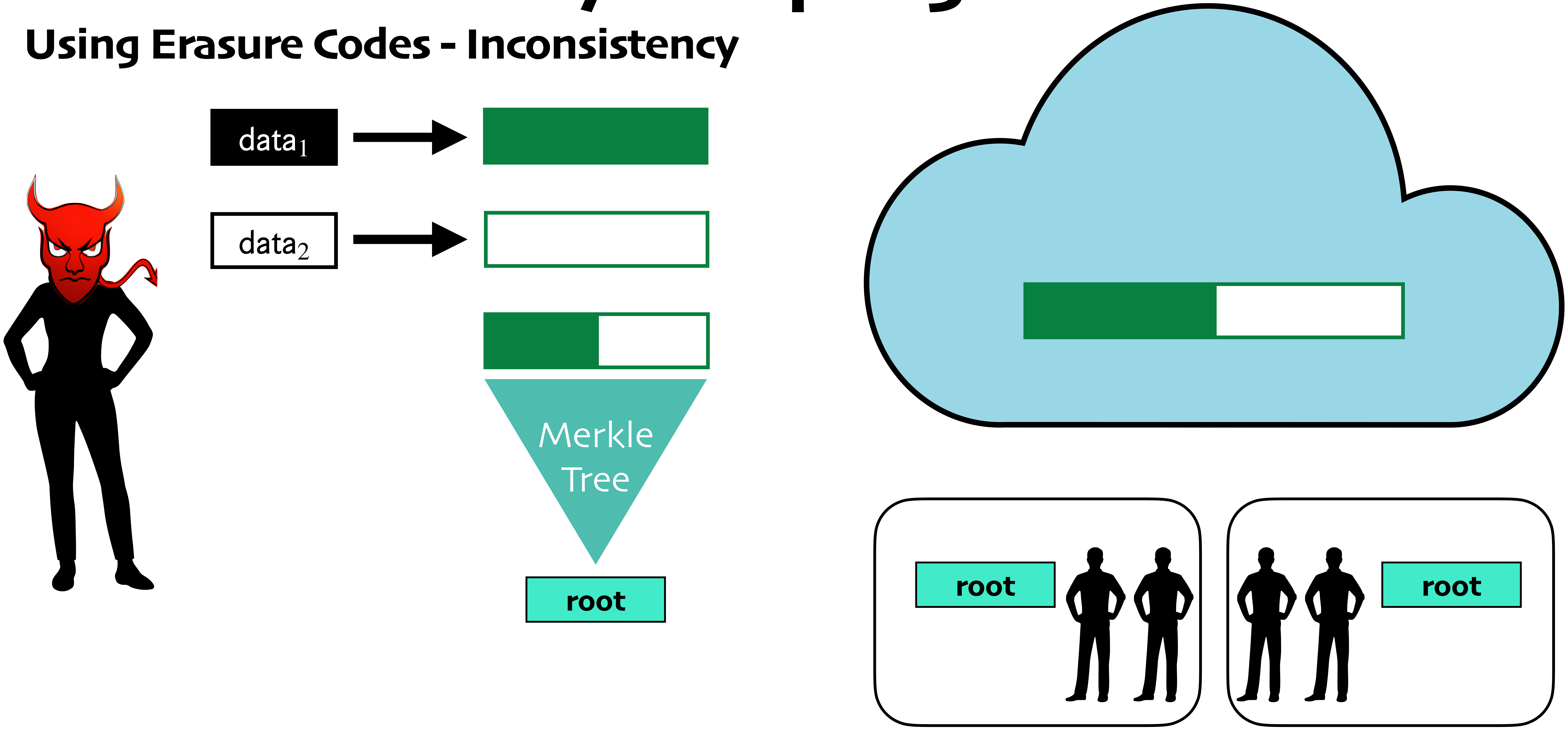
data$_2$

Merkle Tree

root

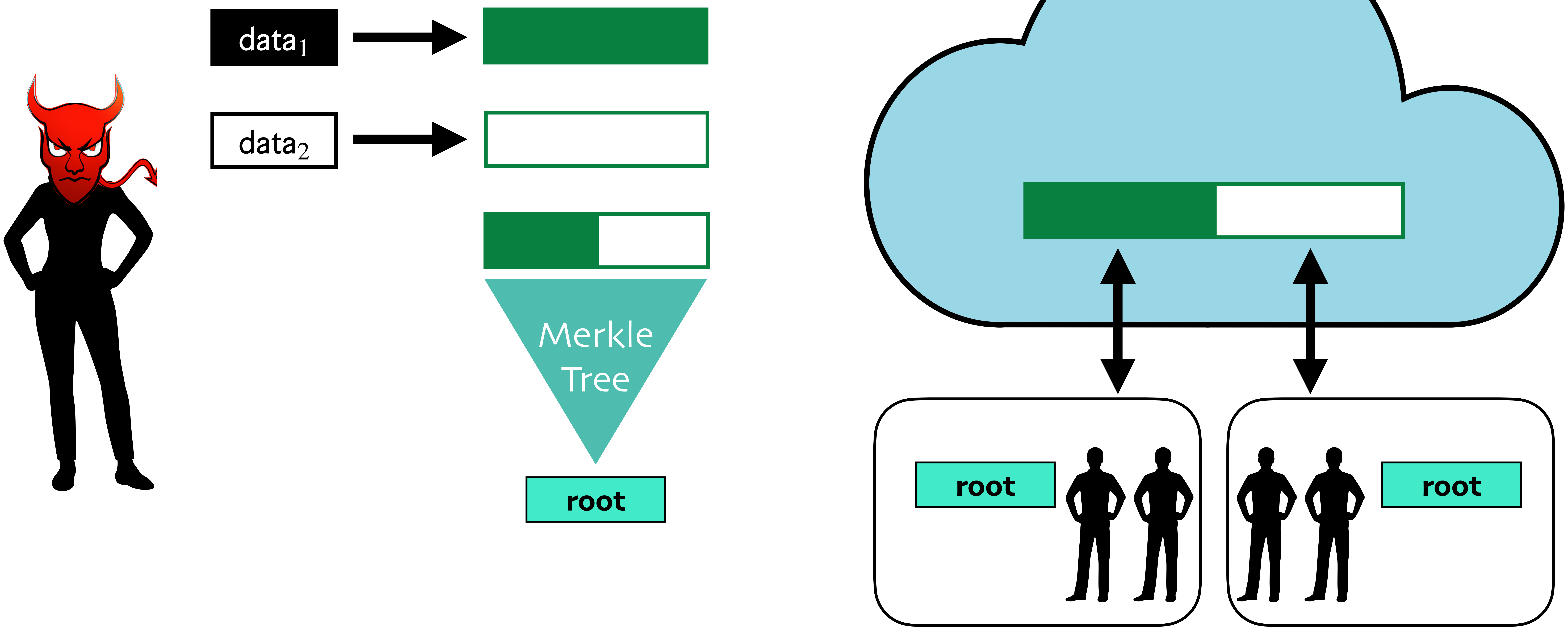
Data Availability Sampling

Using Erasure Codes - Inconsistency



Data Availability Sampling

Using Erasure Codes - Inconsistency



data_1

data_2

Merkle Tree

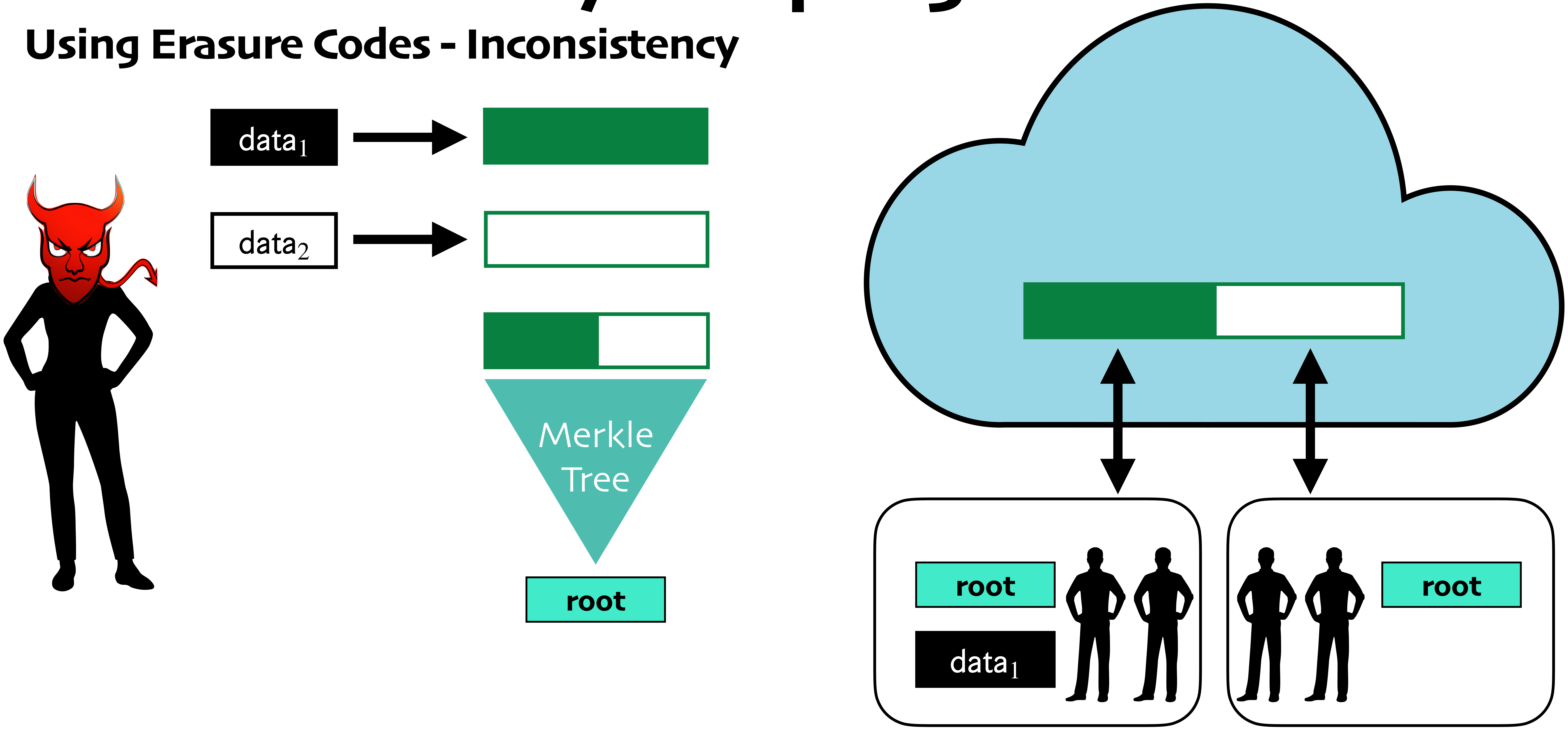
root

root

root

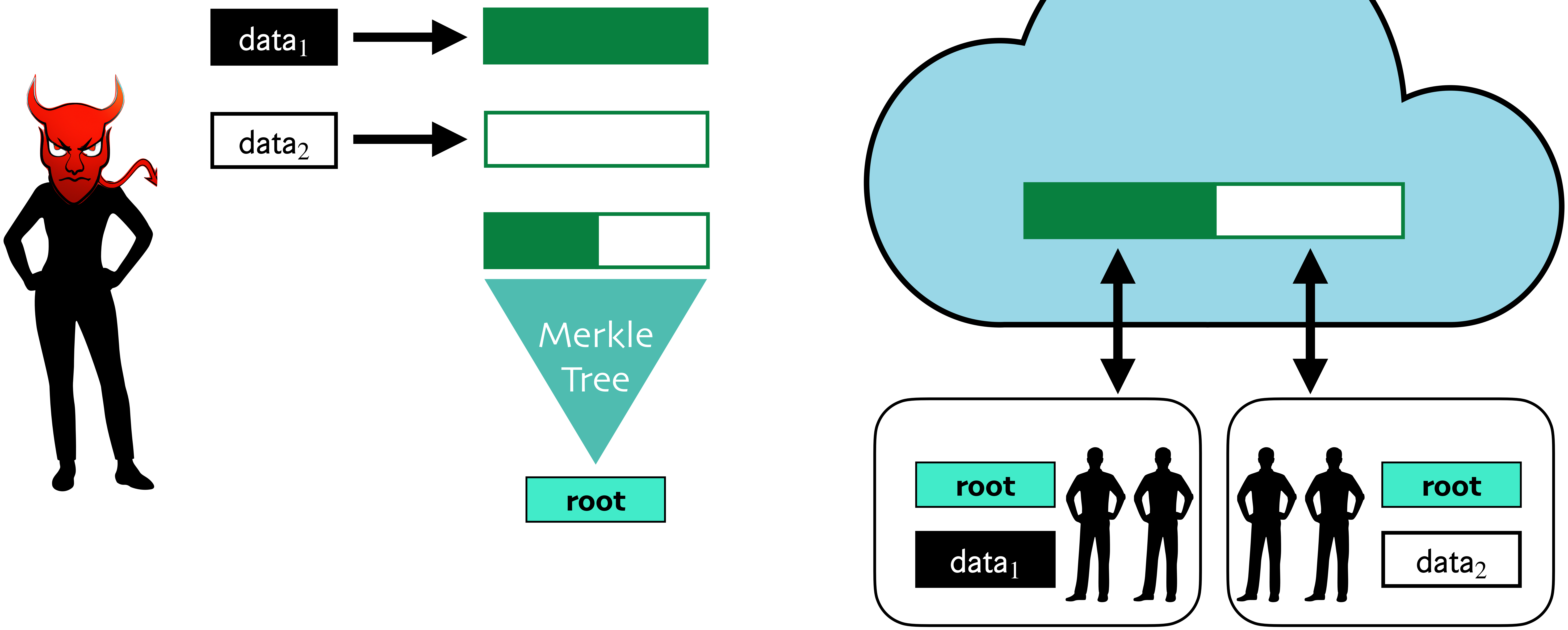
Data Availability Sampling

Using Erasure Codes - Inconsistency



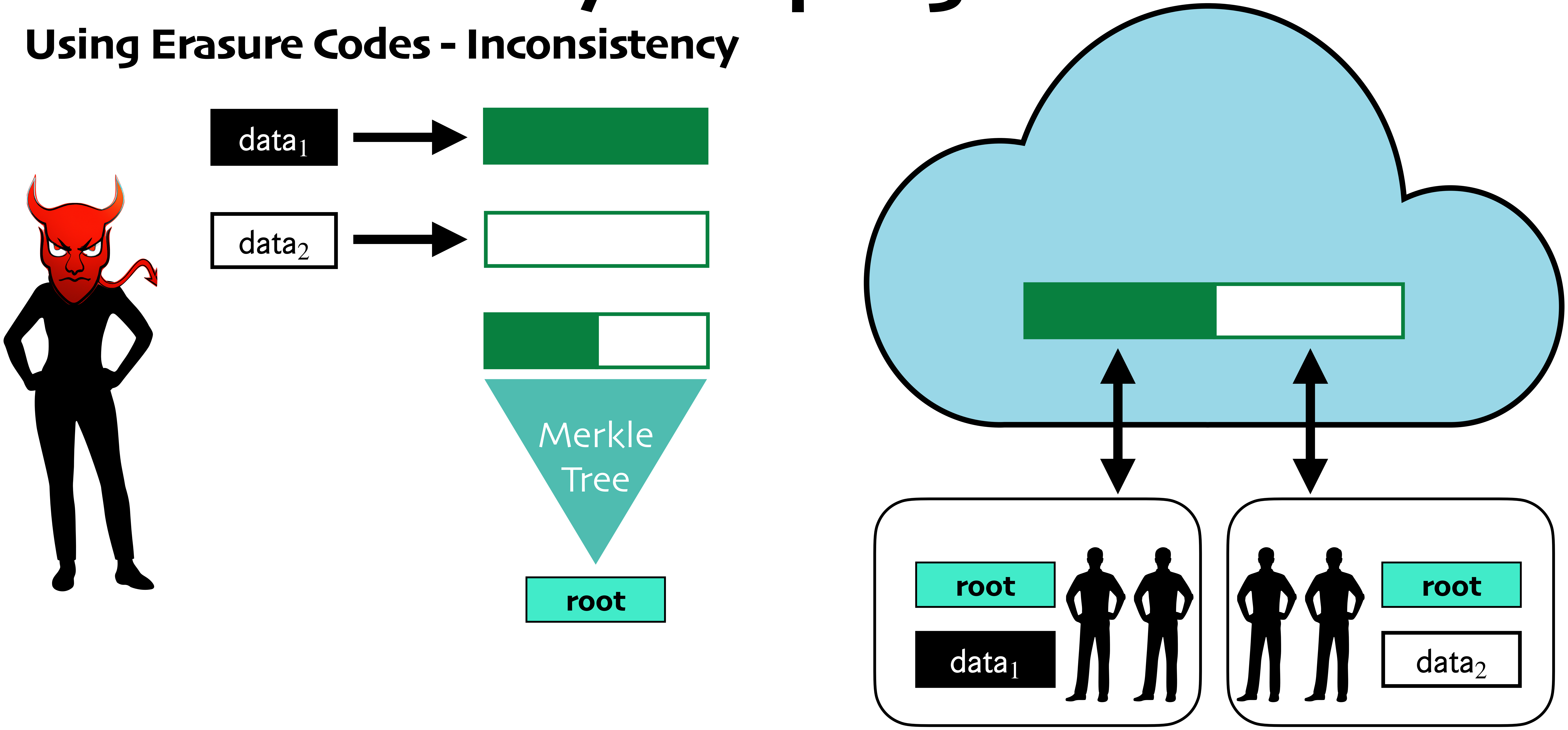
Data Availability Sampling

Using Erasure Codes - Inconsistency



Data Availability Sampling

Using Erasure Codes - Inconsistency



Inconsistency

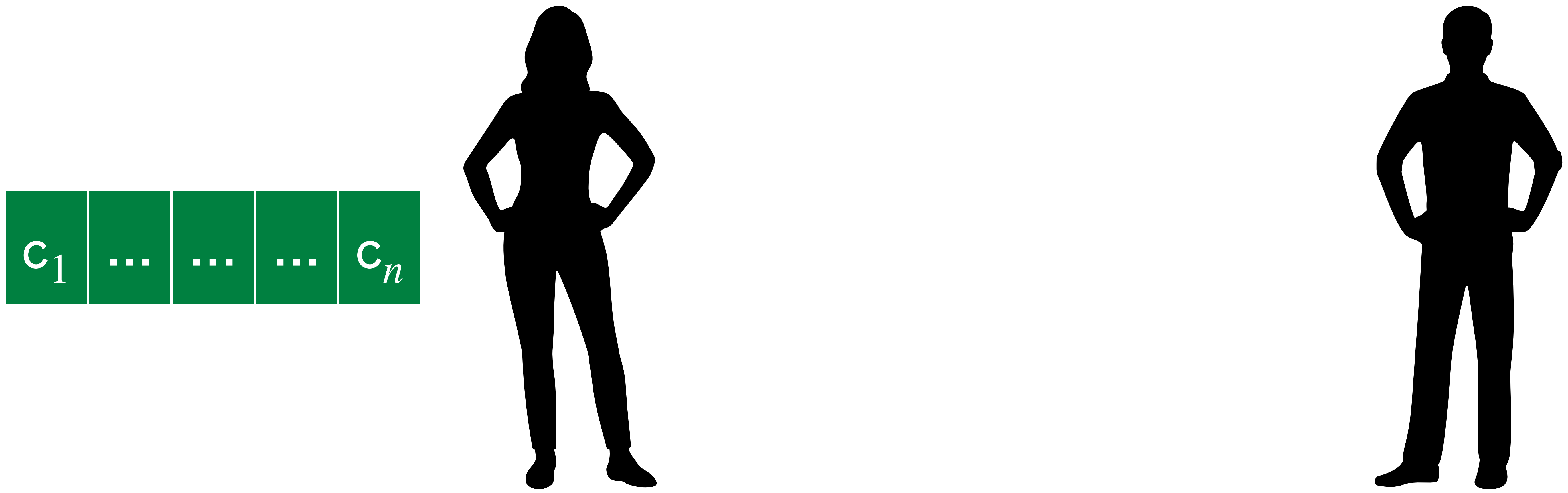
Data Availability Sampling

Data Availability Sampling

Erasure Code
Commitment

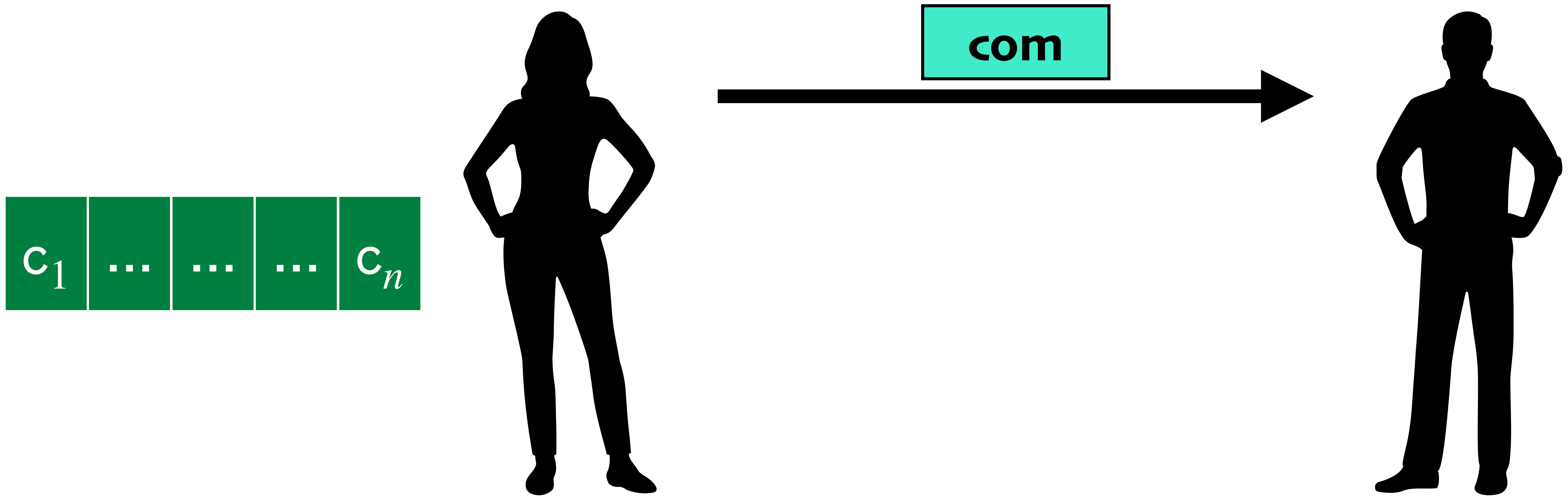
Data Availability Sampling

Erasure Code
Commitment



Data Availability Sampling

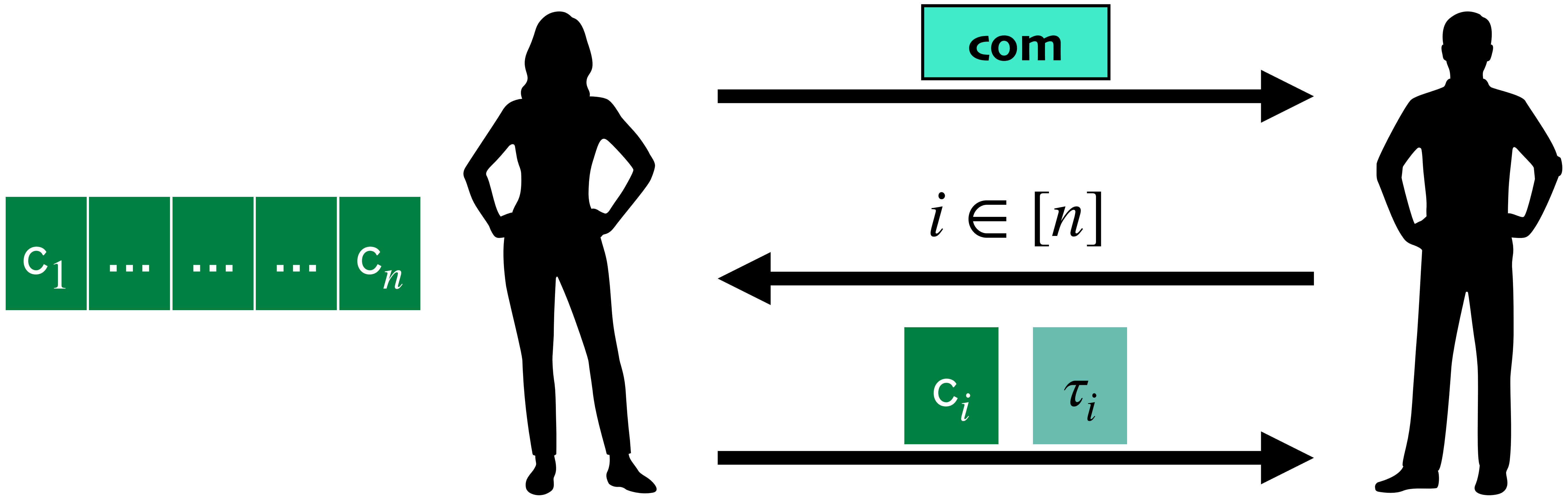
Erasure Code Commitment



c_1	…	…	…	c_n

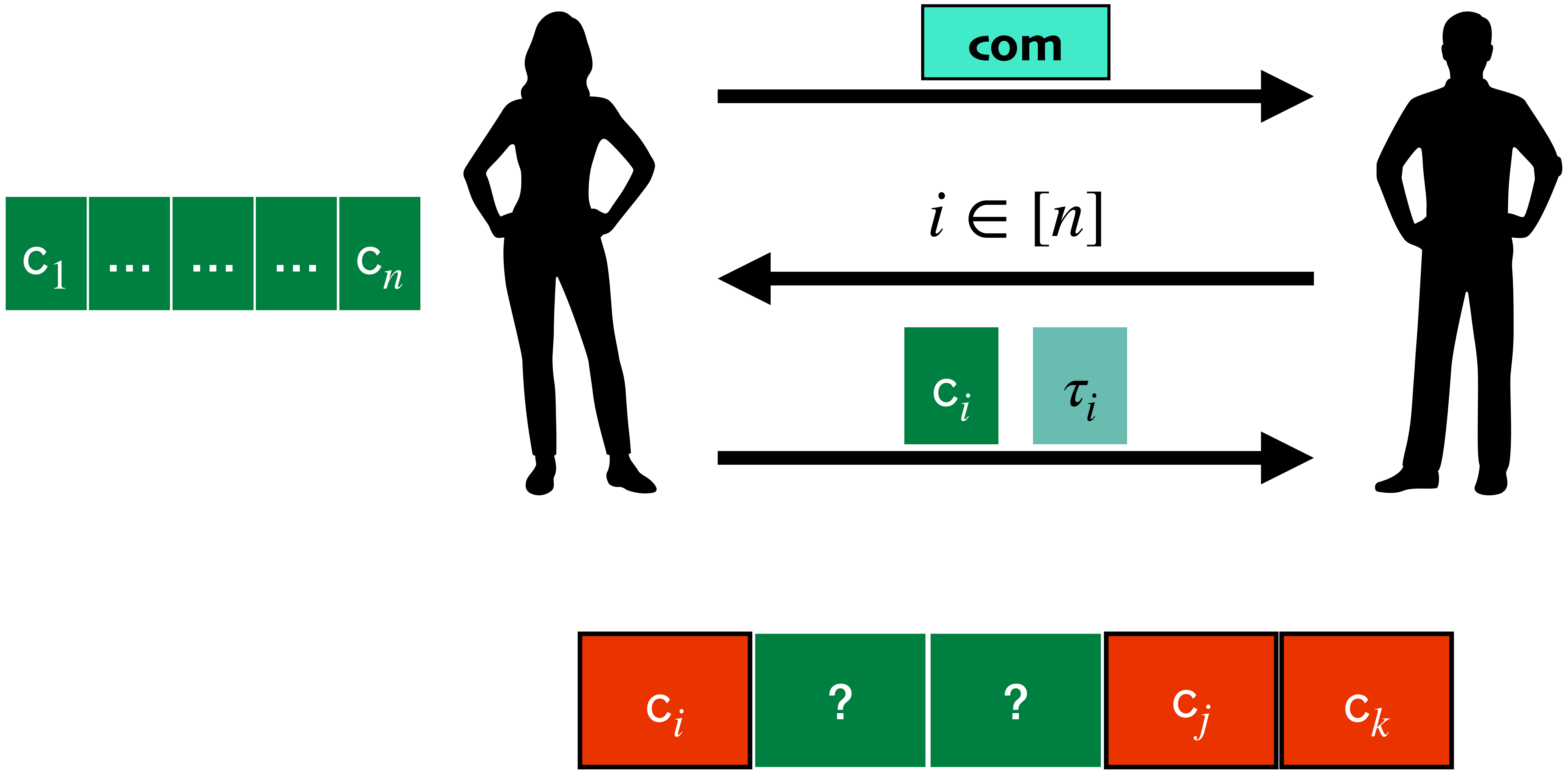
Data Availability Sampling

Erasure Code Commitment



Data Availability Sampling

Erasure Code Commitment

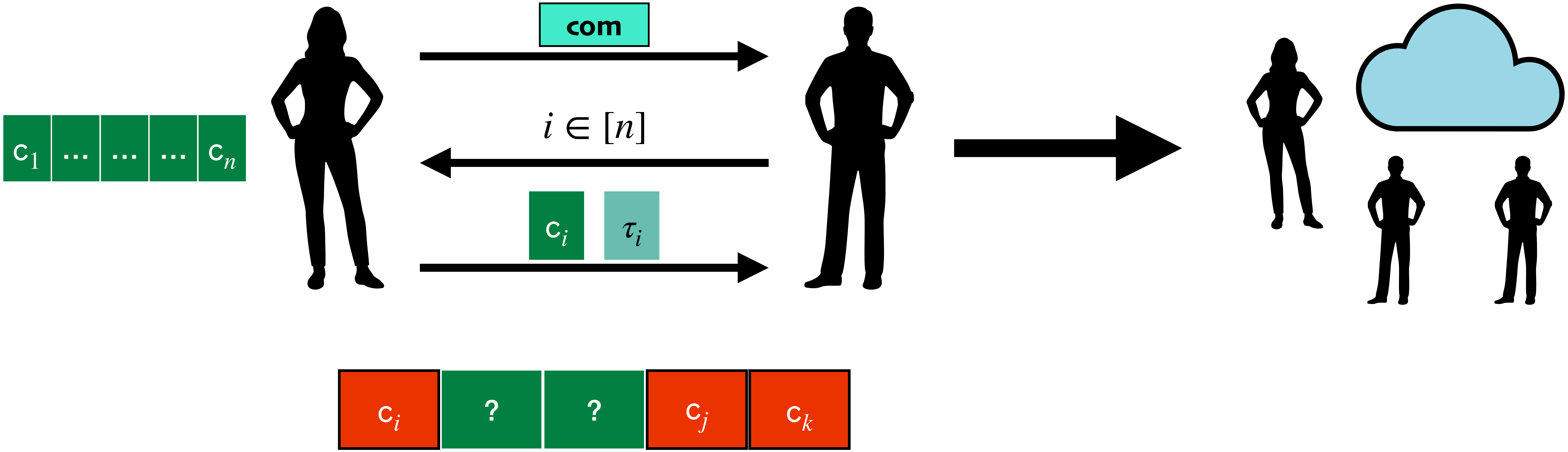


Always Consistent with a Codeword

Data Availability Sampling

Erasure Code Commitment

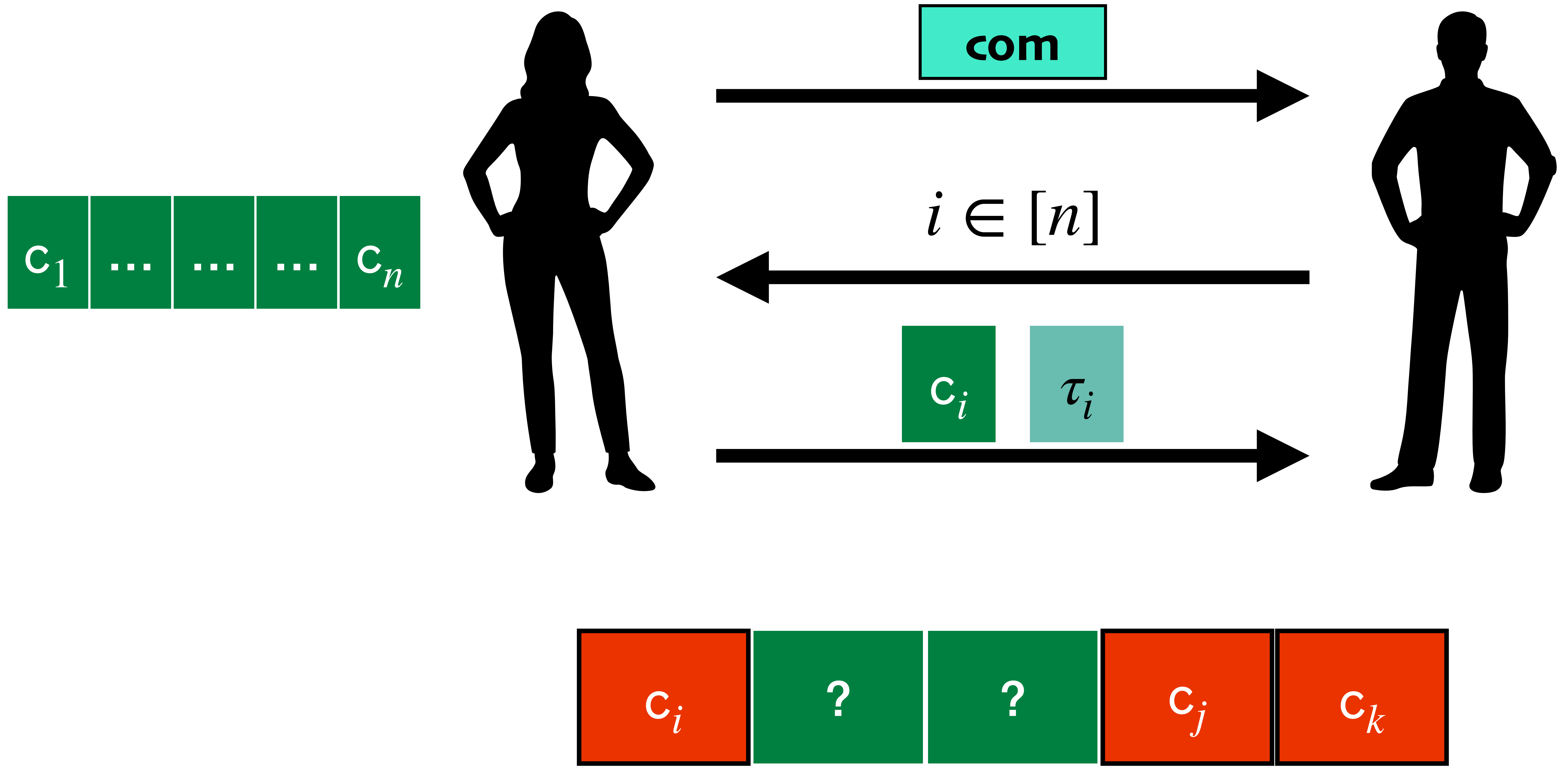
Data Availability Sampling



Always Consistent with a Codeword

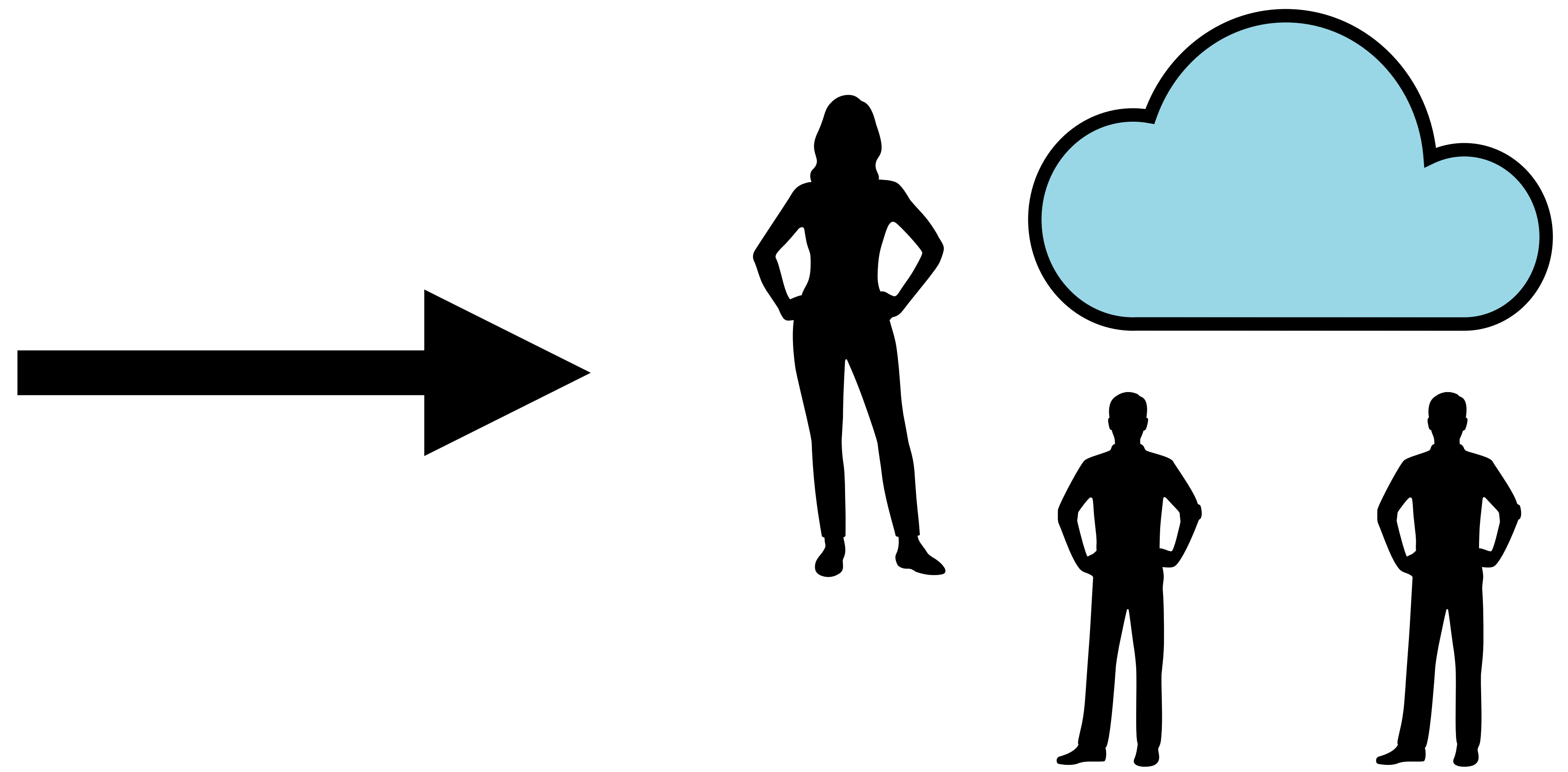
Data Availability Sampling

Erasure Code Commitment



Always Consistent with a Codeword

Data Availability Sampling



Sound

Consistent

Constructions

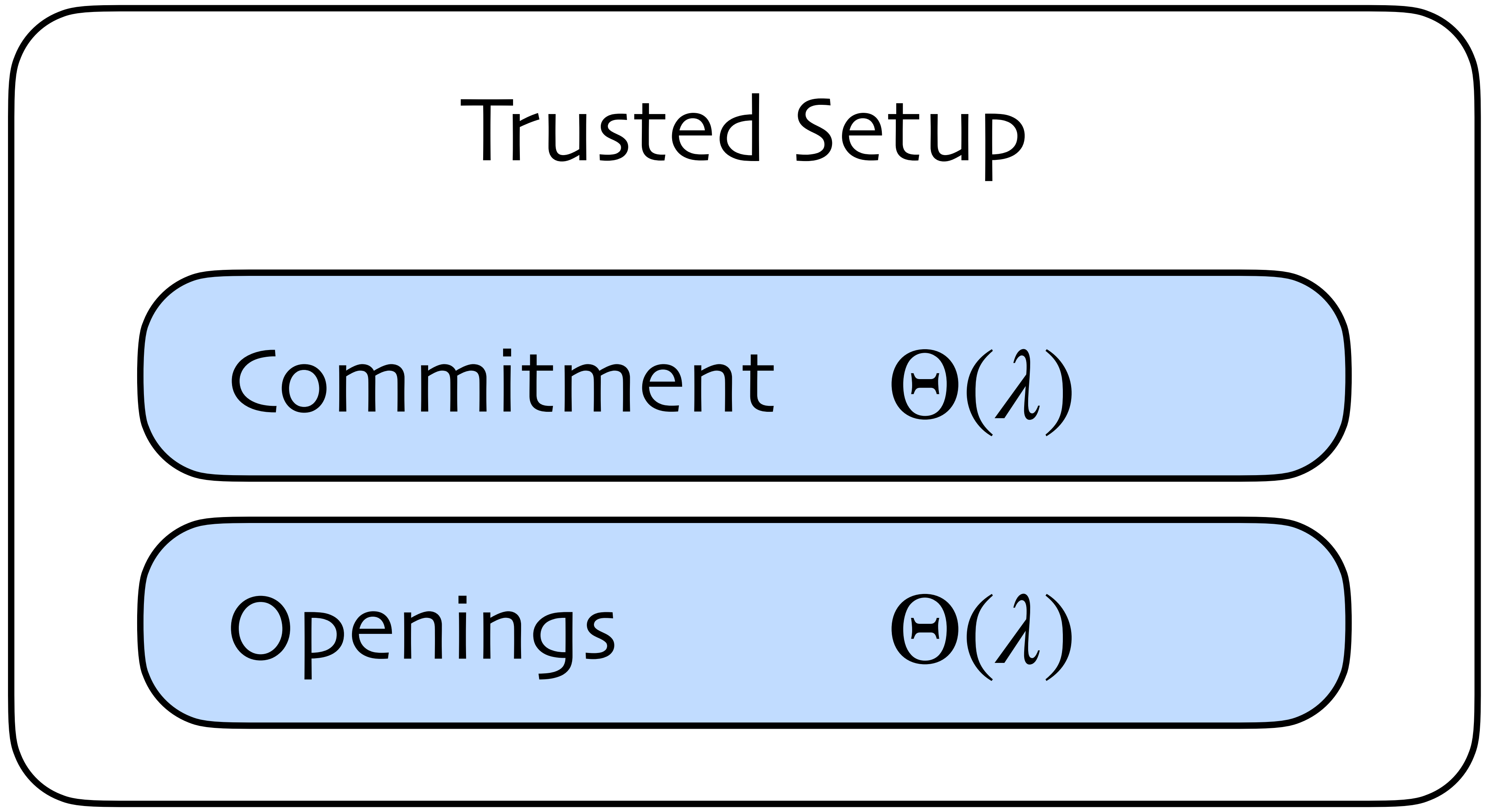
Erasure Code Commitments / DAS

D: size of data
λ : security parameter

Constructions

Erasure Code Commitments / DAS

D: size of data
λ : security parameter



Trusted Setup

Commitment $\Theta(\lambda)$

Openings $\Theta(\lambda)$

Constructions

Erasure Code Commitments / DAS

D: size of data
λ : security parameter

Trusted Setup

Commitment	$\Theta(\lambda)$
Openings	$\Theta(\lambda)$

Hash-Based

Commitment	$\Theta(\lambda\sqrt{D})$
Openings	$\Theta(\sqrt{D})$

Constructions

Erasure Code Commitments / DAS

D: size of data
λ : security parameter

Trusted Setup

Commitment $\Theta(\lambda)$

Openings $\Theta(\lambda)$

Hash-Based

Commitment $\Theta(\lambda\sqrt{D})$

Openings $\Theta(\sqrt{D})$

Hash-Based with Polylog Overhead?

Data Availability Sampling

Data Availability Sampling from FRI

Data Availability Sampling

Data Availability Sampling from FRI

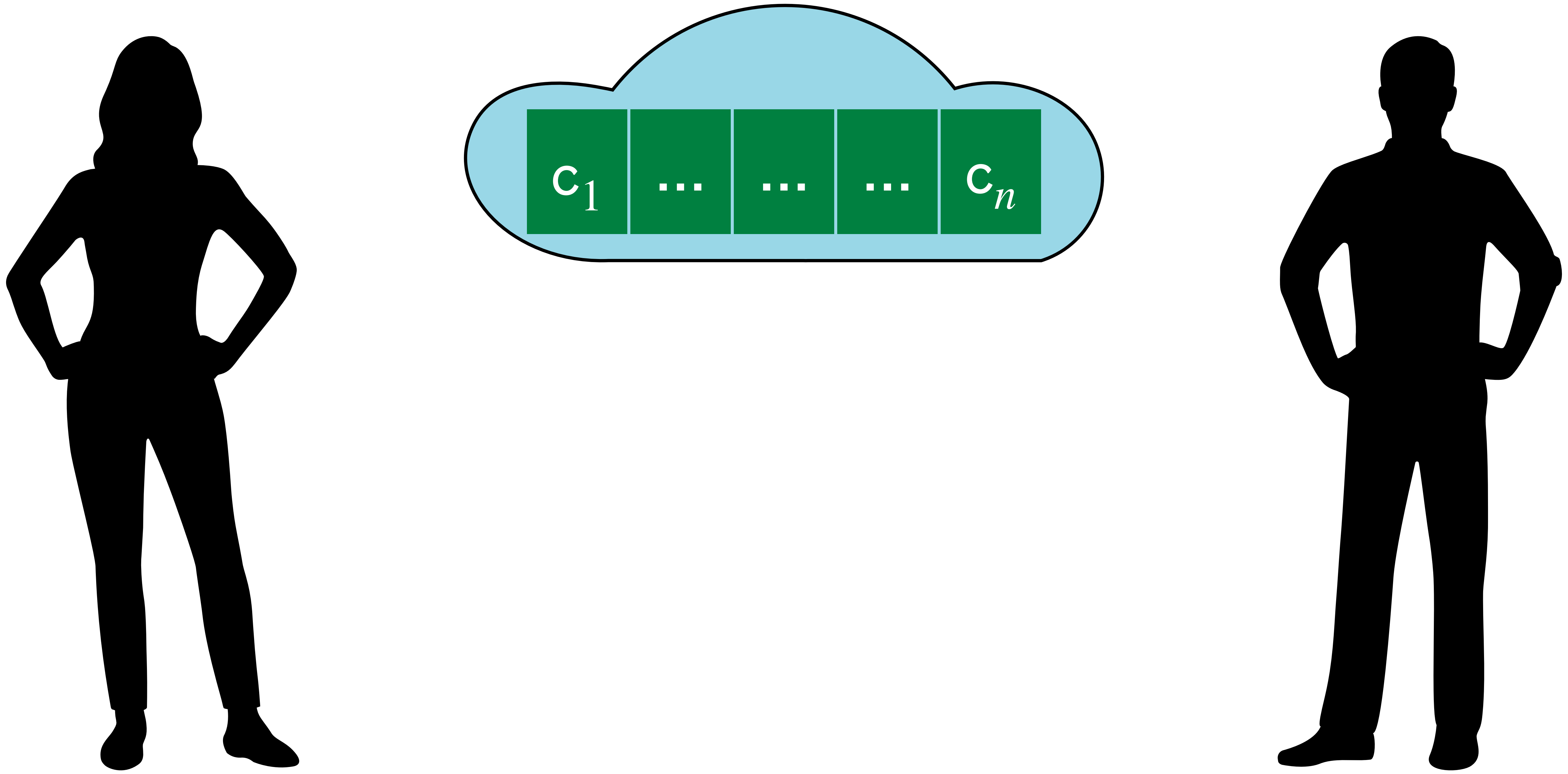
FRI = Fast Reed-Solomon IOPP

Interactive Oracle Proofs of Proximity

IOPPs

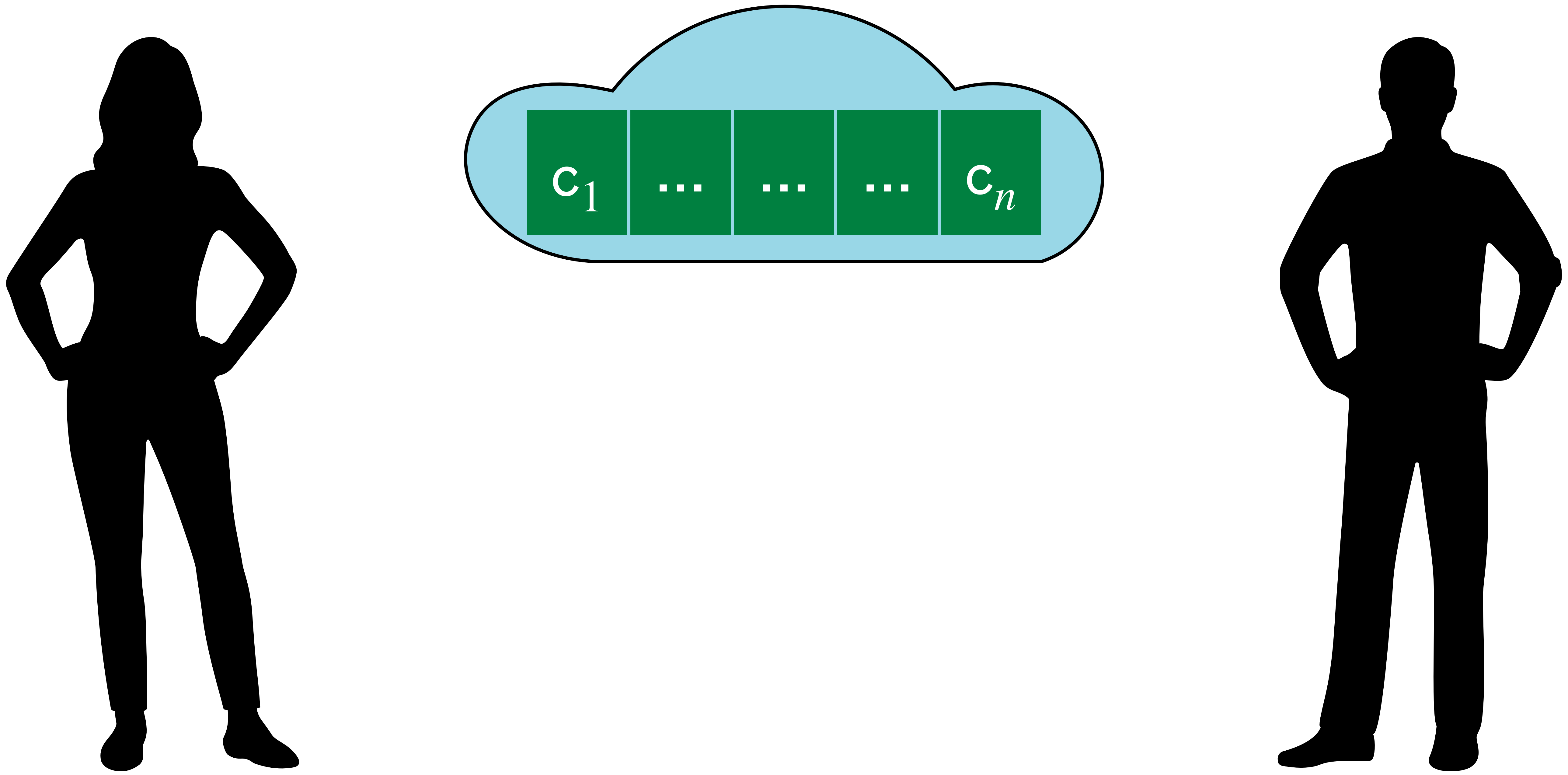
Interactive Oracle Proofs of Proximity

IOPPs



Interactive Oracle Proofs of Proximity

IOPPs

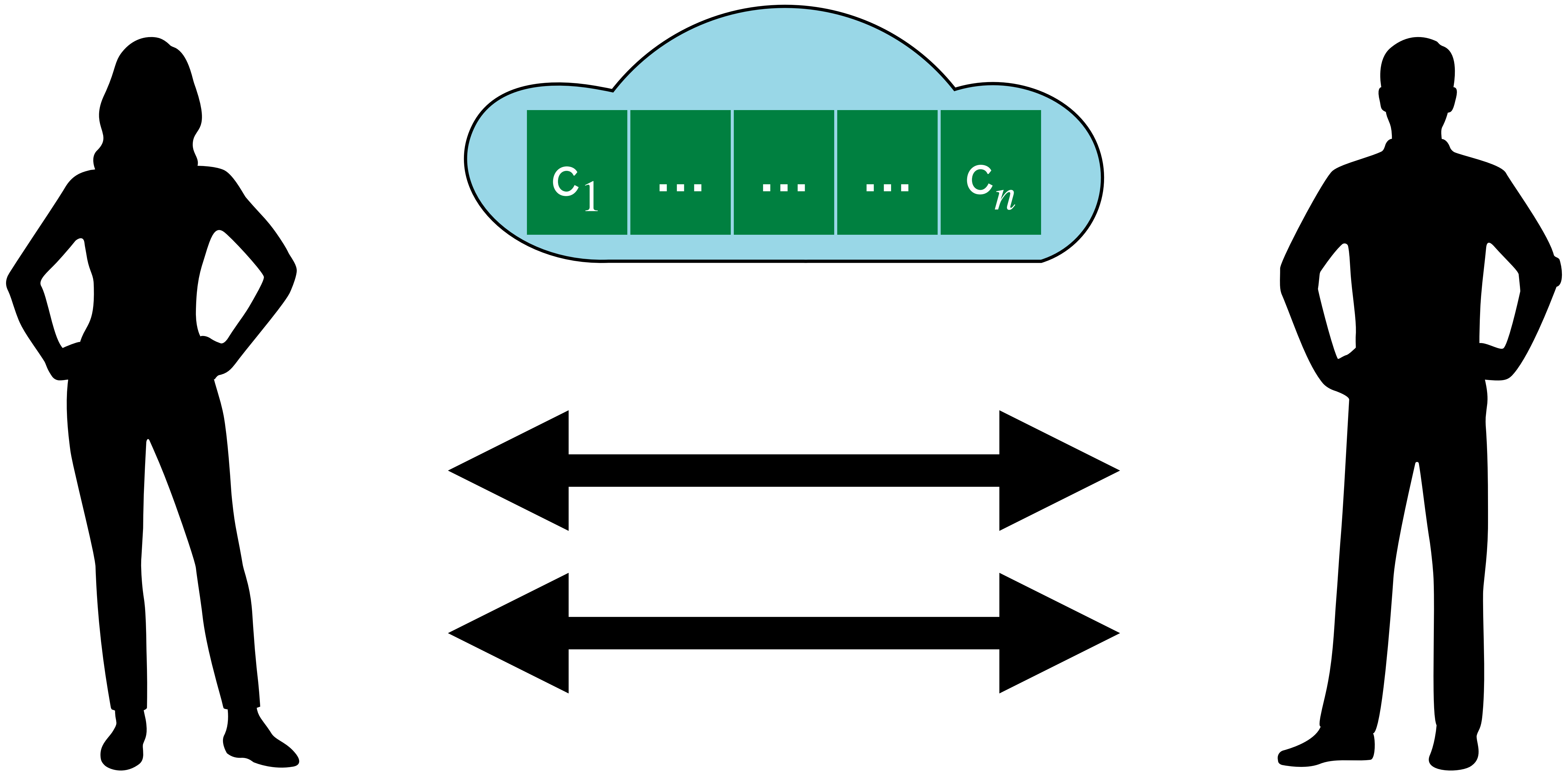


c_1	…	…	…	c_n

Close to the Code?

Interactive Oracle Proofs of Proximity

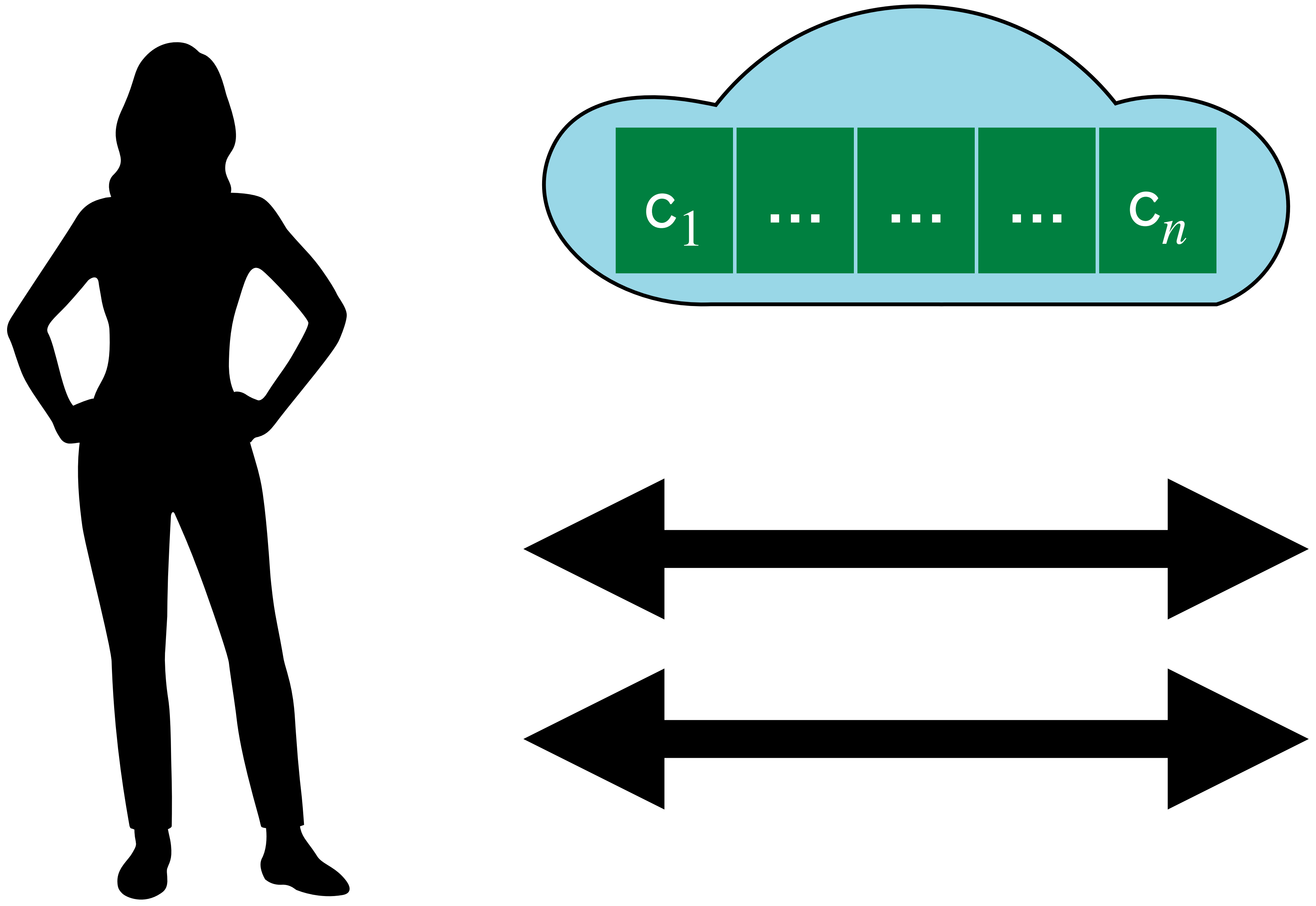
IOPPs



Close to the Code?

Interactive Oracle Proofs of Proximity

IOPPs

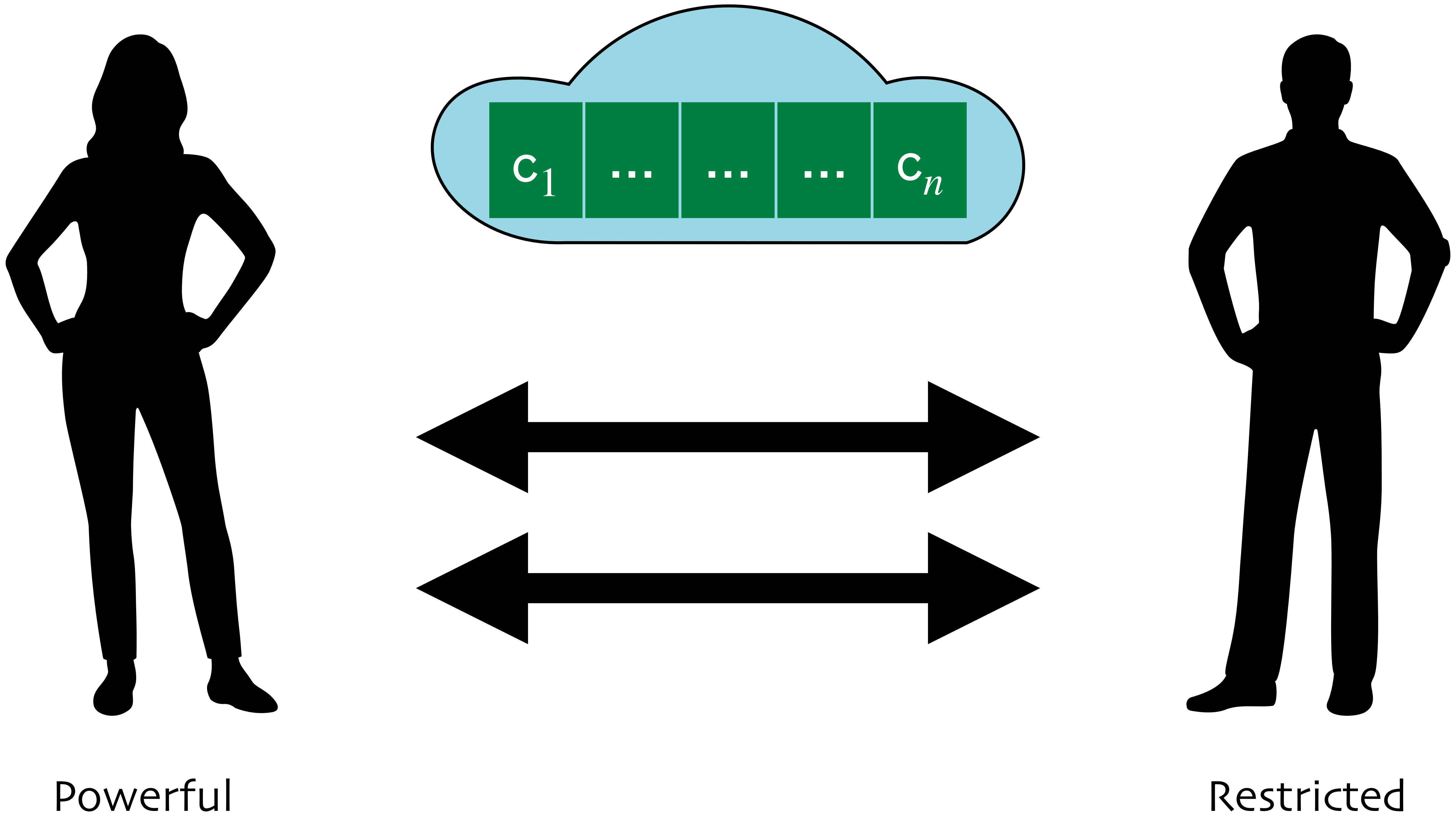


Close to the Code?

c* in Code	c_1^*	…	…	…	…	…	…	c_n^*
c close	c_1	…	🟥 …	…	…	…	…	c_n

Interactive Oracle Proofs of Proximity

IOPPs



Close to the Code?

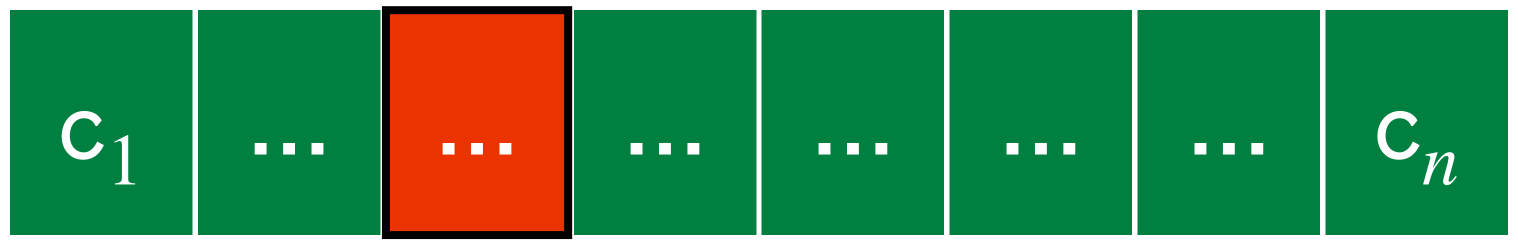
c* in Code	c_1^*	…	…	…	…	…	…	c_n^*
c close	c_1	…	🟥 …	…	…	…	…	c_n

Interactive Oracle Proofs of Proximity

IOPPs vs Erasure Code Commitments

Interactive Oracle Proofs of Proximity

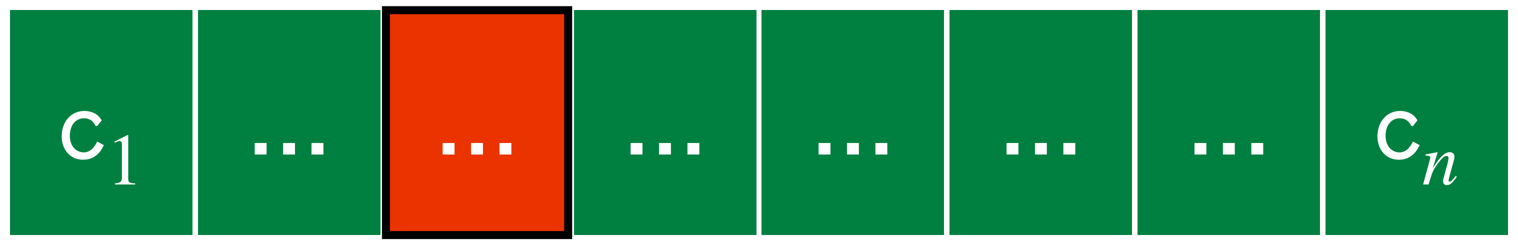
IOPPs vs Erasure Code Commitments



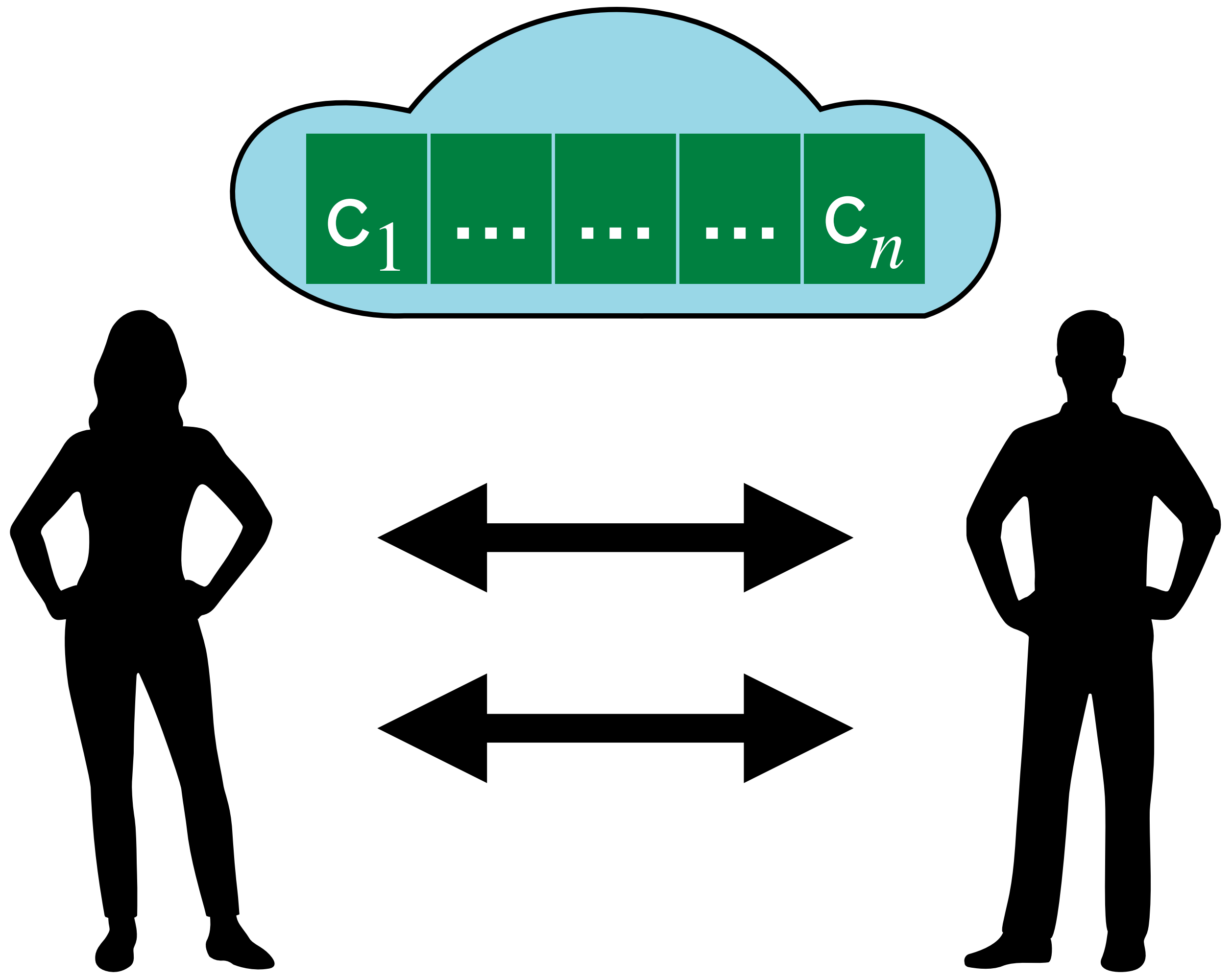
c_1	…	…	…	…	…	…	c_n

Interactive Oracle Proofs of Proximity

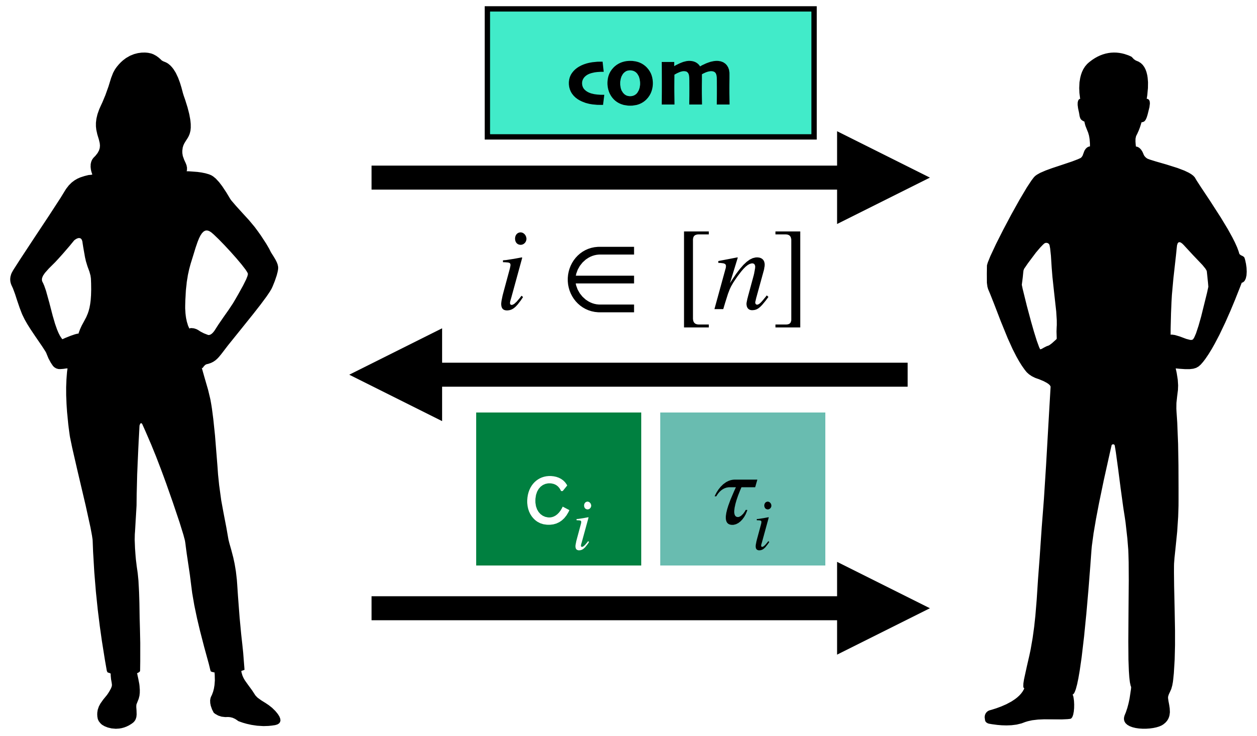
IOPPs vs Erasure Code Commitments



IOPP

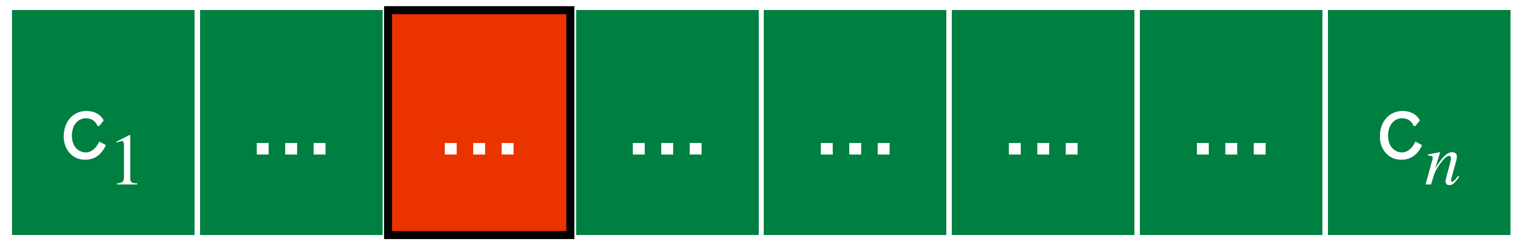


Erasure Code Commitments

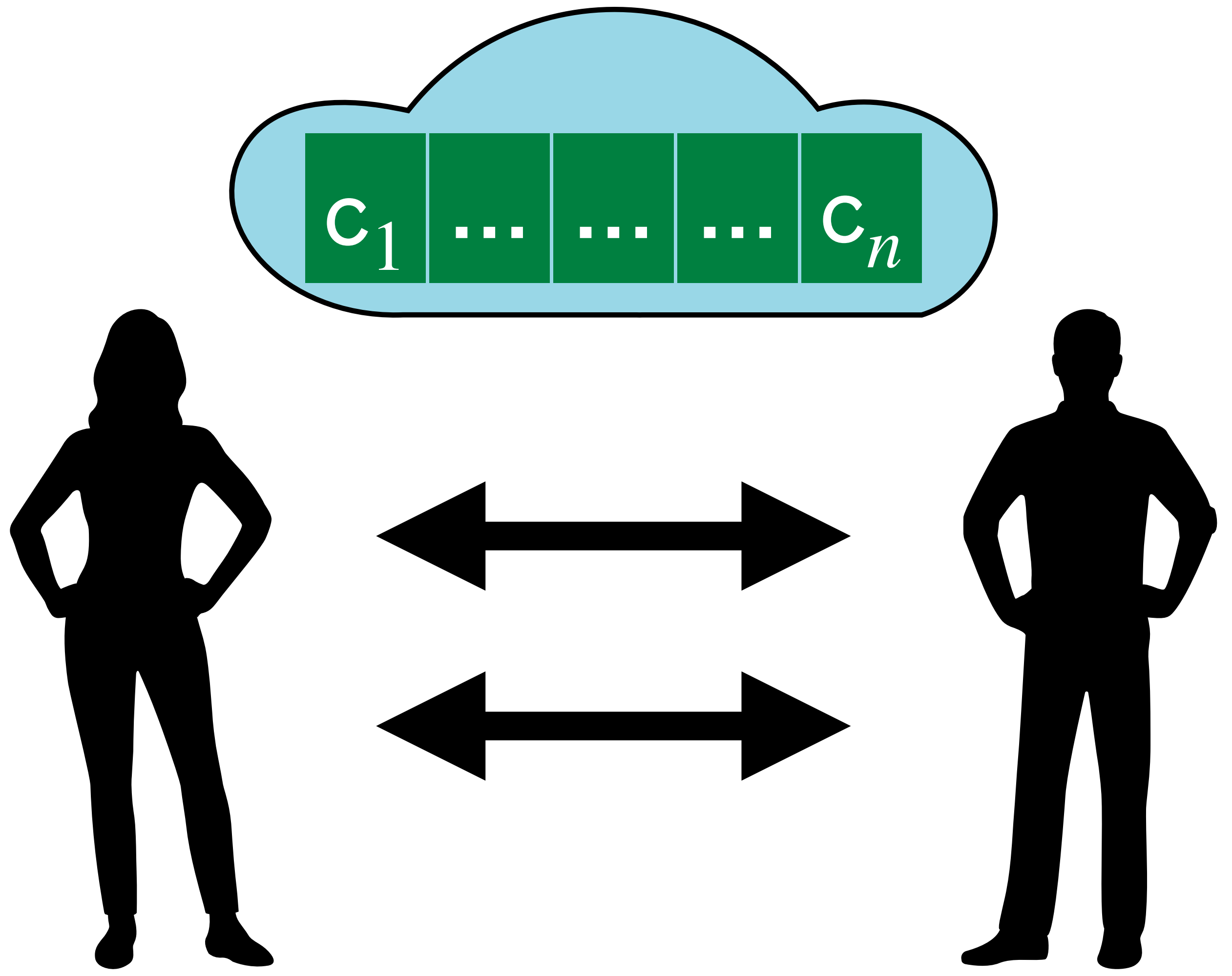


Interactive Oracle Proofs of Proximity

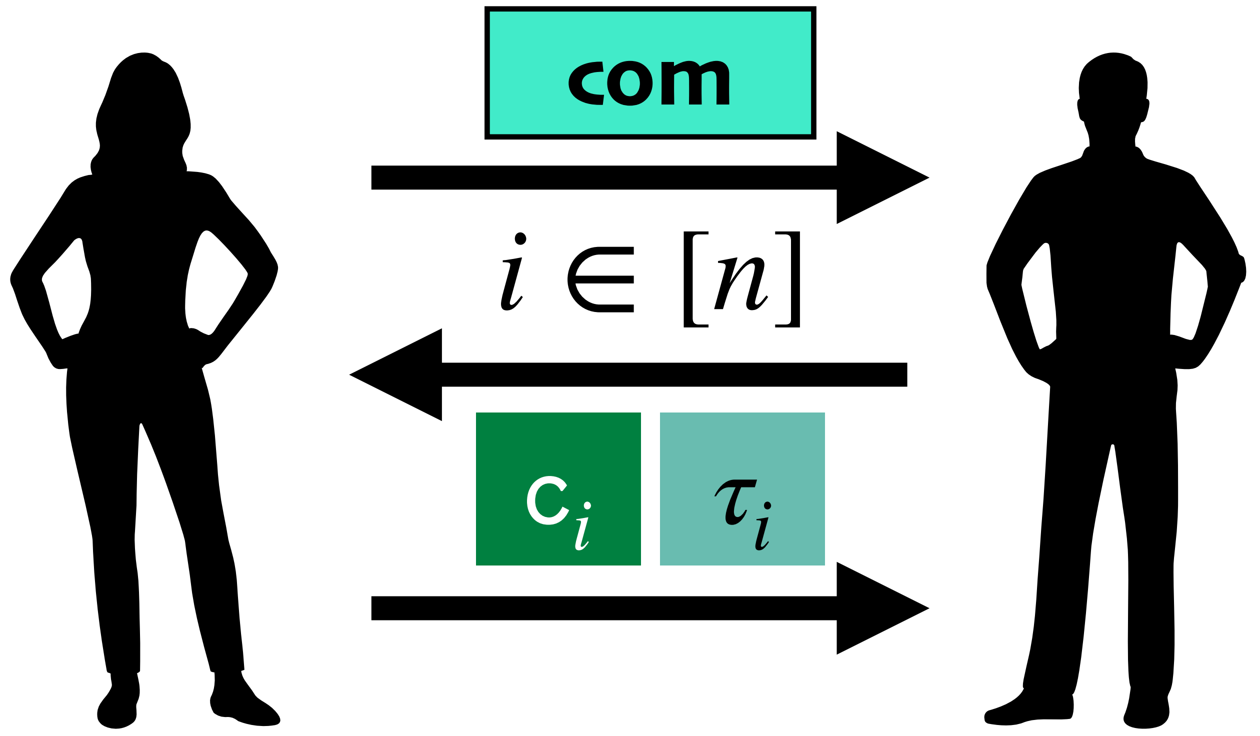
IOPPs vs Erasure Code Commitments



IOPP



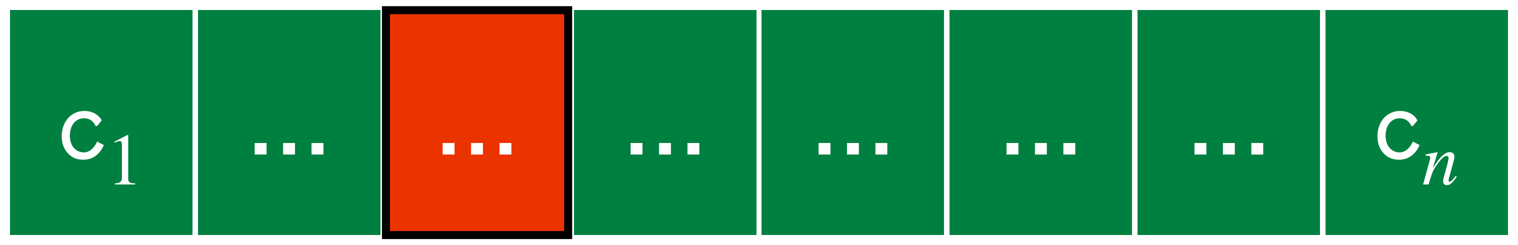
Erasure Code Commitments



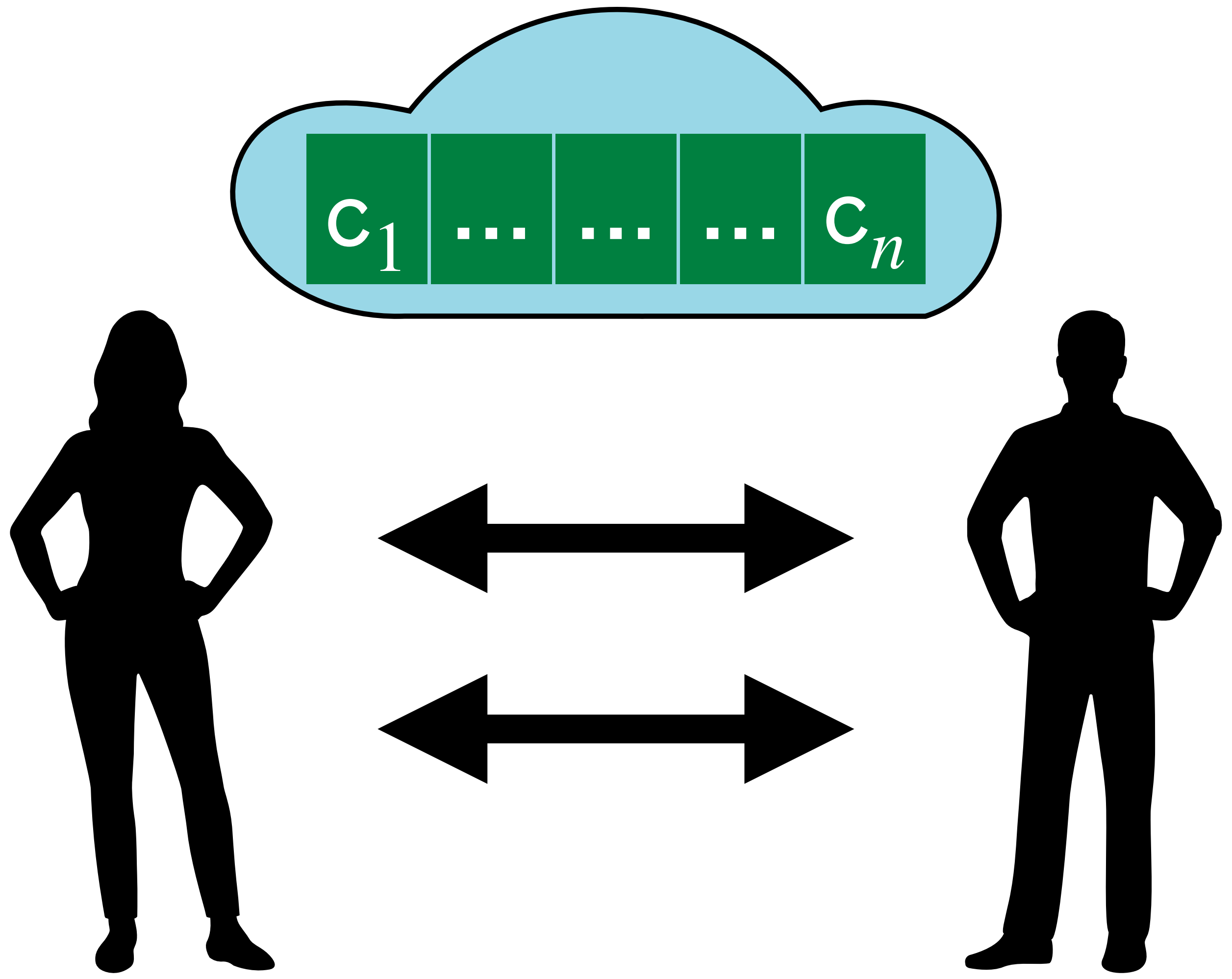
Few Red Positions

Interactive Oracle Proofs of Proximity

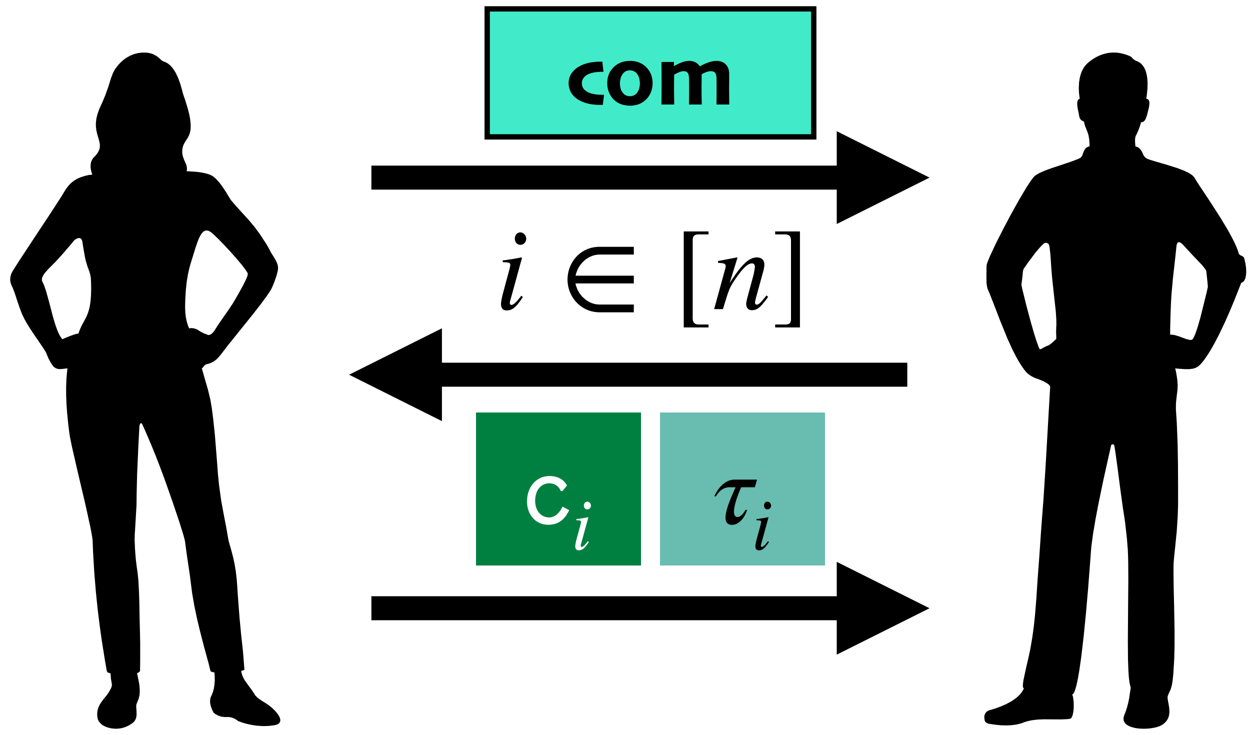
IOPPs vs Erasure Code Commitments



IOPP



Erasure Code Commitments



Few Red Positions

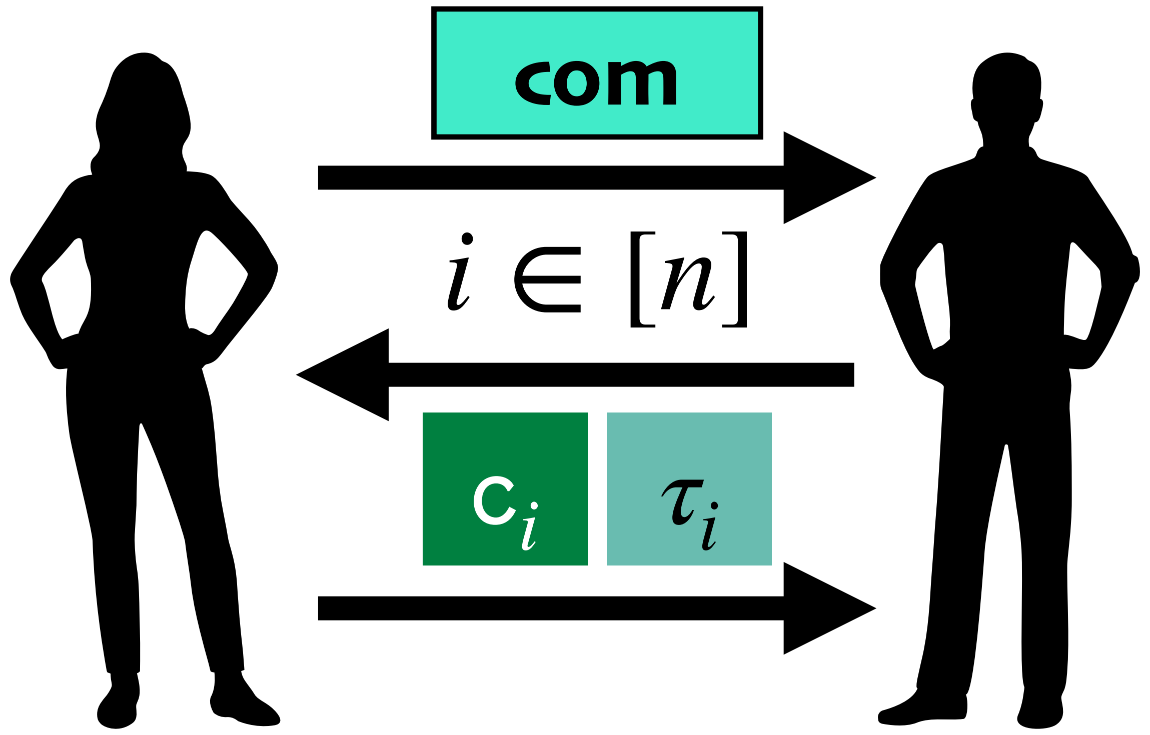
Can’t Open Red Positions

Our Compiler

IOPPs to Erasure Code Commitments

Our Compiler

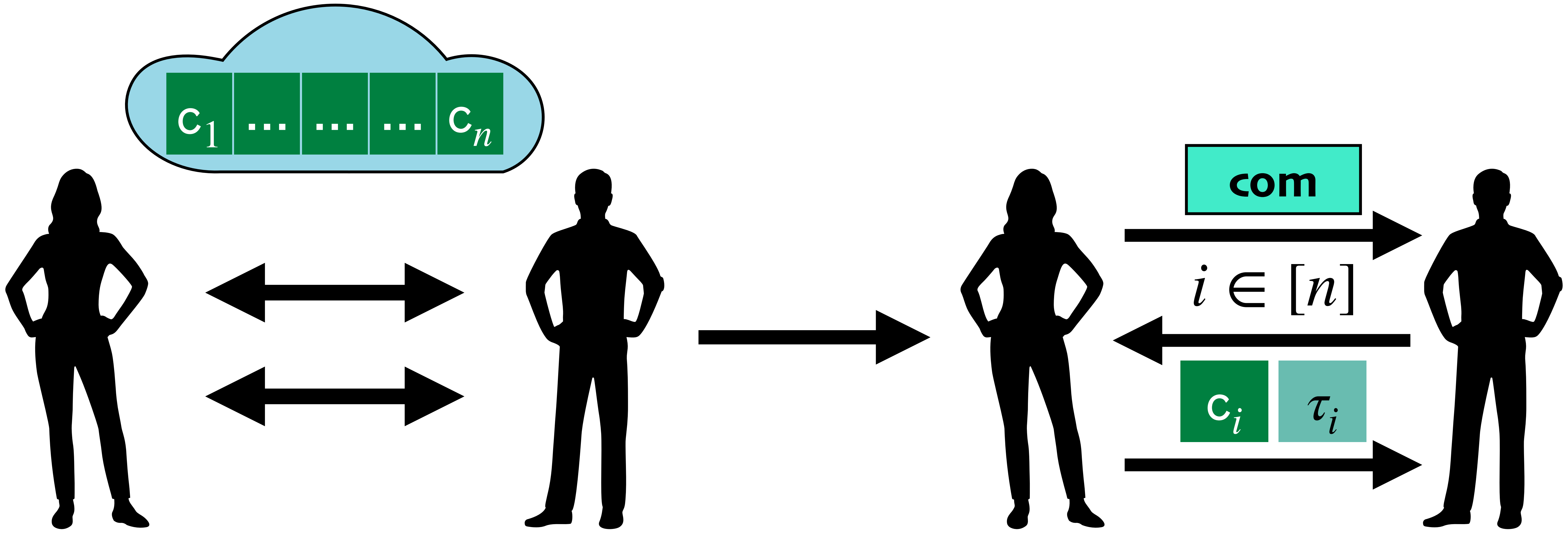
IOPPs to Erasure Code Commitments



Erasure Code Commitment

Our Compiler

IOPPs to Erasure Code Commitments

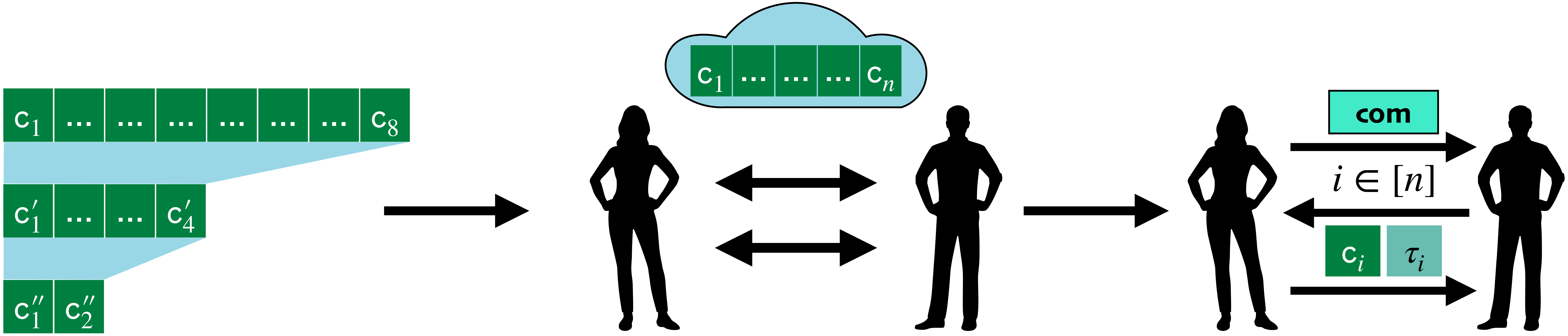


IOPP with Opening-Consistency

Erasure Code Commitment

Our Compiler

IOPPs to Erasure Code Commitments



FRI

IOPP with Opening-Consistency

Erasure Code Commitment

Efficiency of FRIDA

D: size of data
λ : security parameter

Efficiency of FRIDA

D: size of data
λ : security parameter

	Trusted Setup	Hash-Based
Commitment	$\Theta(\lambda)$	$\Theta(\lambda\sqrt{D})$
Openings	$\Theta(\lambda)$	$\Theta(\sqrt{D})$

Efficiency of FRIDA

D: size of data
λ : security parameter

Trusted Setup

Commitment $\Theta(\lambda)$

Openings $\Theta(\lambda)$

Hash-Based

Commitment $\Theta(\lambda\sqrt{D})$

Openings $\Theta(\sqrt{D})$

FRIDA - Commitment from FRI

Commitment $\Theta(\lambda^2 \log^2 D)$

Openings $\Theta(\lambda \log^2 D)$

Summary and Future Work

Summary and Future Work

DAS and Erasure Code Commitments

Summary and Future Work

DAS and Erasure Code Commitments

DAS from FRI

Summary and Future Work

DAS and Erasure Code Commitments

DAS from FRI

No Trusted Setup

Summary and Future Work

DAS and Erasure Code Commitments

DAS from FRI

No Trusted Setup

Polylog Overhead

Summary and Future Work

DAS and Erasure Code Commitments

DAS from FRI

No Trusted Setup

Polylog Overhead

Compiler from IOPP

Summary and Future Work

DAS and Erasure Code Commitments

DAS from FRI

No Trusted Setup

Polylog Overhead

Compiler from IOPP

Opening-Consistency

Summary and Future Work

DAS and Erasure Code Commitments

DAS from FRI

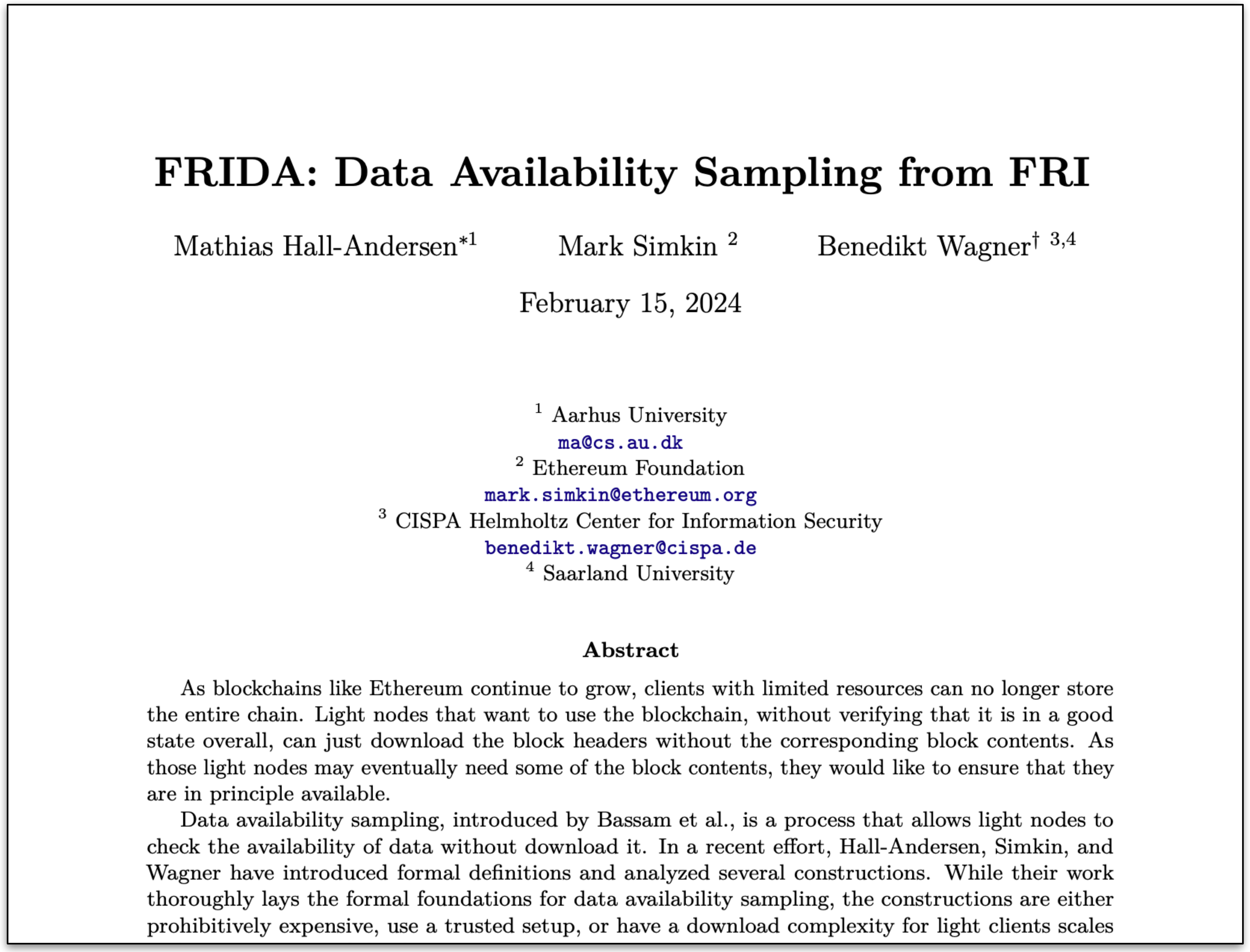
- No Trusted Setup
- Polylog Overhead
- Compiler from IOPP
- Opening-Consistency

Better IOPPs with Opening-Consistency

Mathias Hall-Andersen

Mark Simkin

Benedikt Wagner



FRIDA: Data Availability Sampling from FRI

Mathias Hall-Andersen*[1] Mark Simkin [2] Benedikt Wagner† [3,4]

February 15, 2024

[1] Aarhus University
ma@cs.au.dk
[2] Ethereum Foundation
mark.simkin@ethereum.org
[3] CISPA Helmholtz Center for Information Security
benedikt.wagner@cispa.de
[4] Saarland University

Abstract

As blockchains like Ethereum continue to grow, clients with limited resources can no longer store the entire chain. Light nodes that want to use the blockchain, without verifying that it is in a good state overall, can just download the block headers without the corresponding block contents. As those light nodes may eventually need some of the block contents, they would like to ensure that they are in principle available.

Data availability sampling, introduced by Bassam et al., is a process that allows light nodes to check the availability of data without download it. In a recent effort, Hall-Andersen, Simkin, and Wagner have introduced formal definitions and analyzed several constructions. While their work thoroughly lays the formal foundations for data availability sampling, the constructions are either prohibitively expensive, use a trusted setup, or have a download complexity for light clients scales

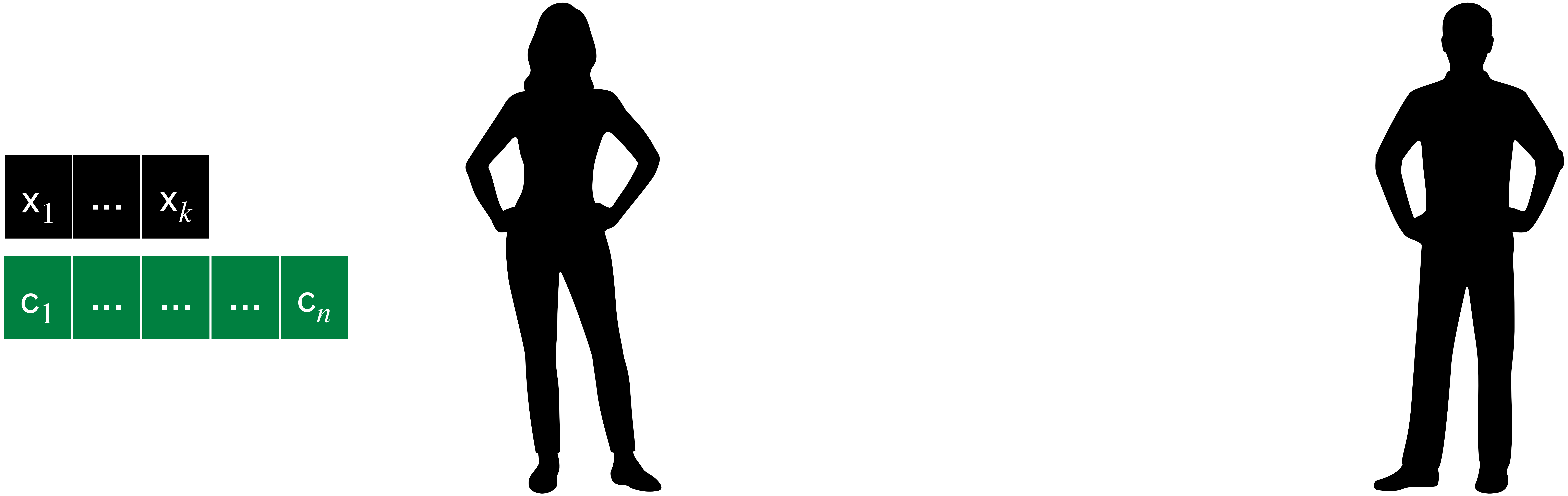
eprint: 2024/248

Construction Framework

Erasure Code Commitments

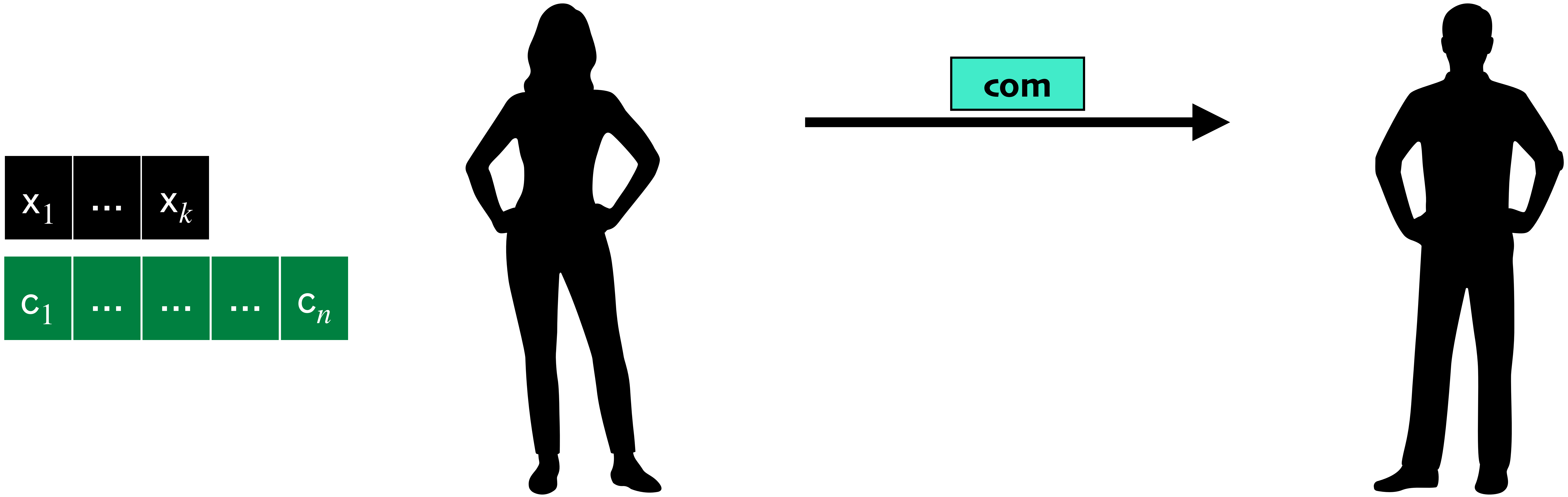
Construction Framework

Erasure Code Commitments



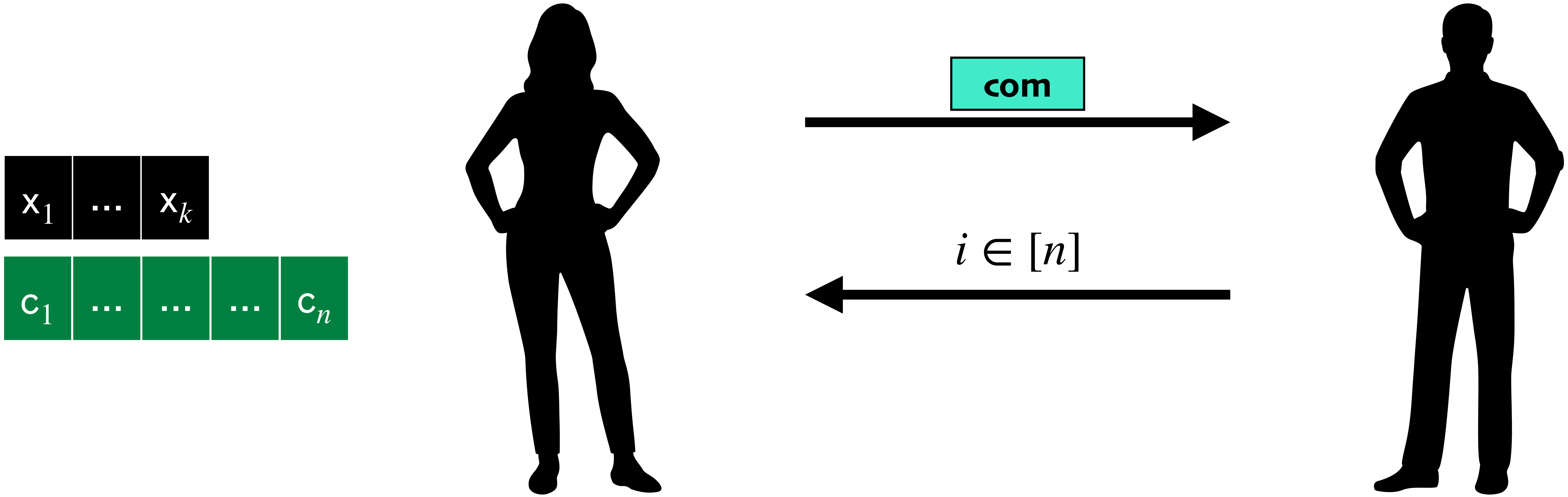
Construction Framework

Erasure Code Commitments



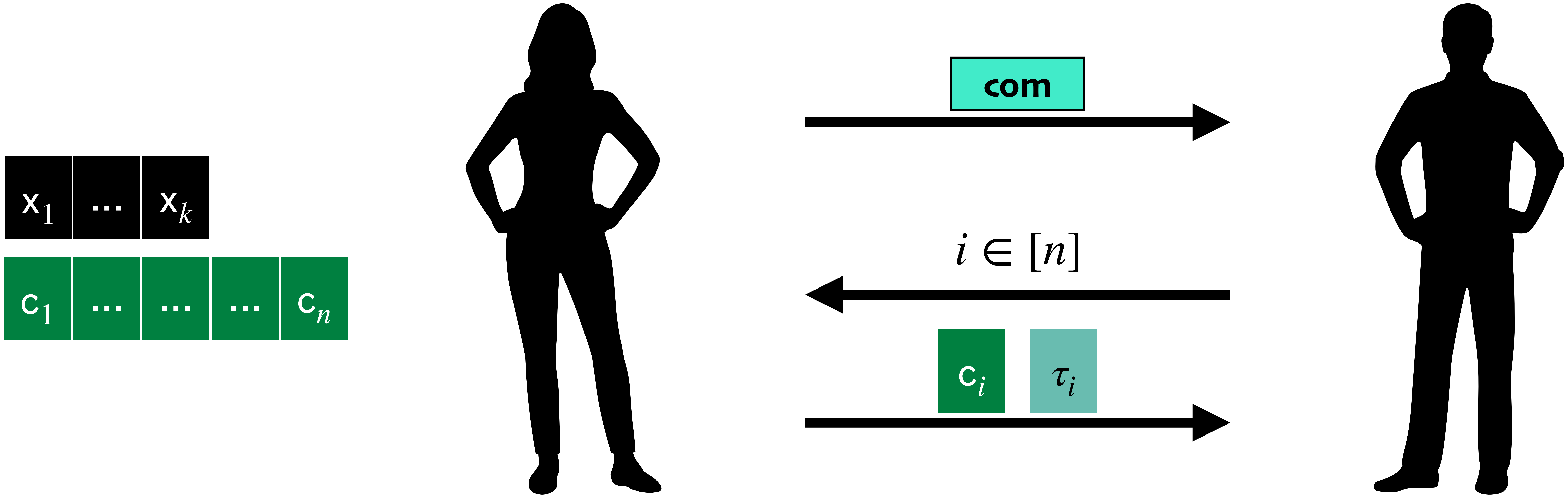
Construction Framework

Erasure Code Commitments



Construction Framework

Erasure Code Commitments



Construction Framework

Erasure Code Commitments

Construction Framework

Erasure Code Commitments

Position-Binding

Code-Binding

Construction Framework

Erasure Code Commitments

Position-Binding

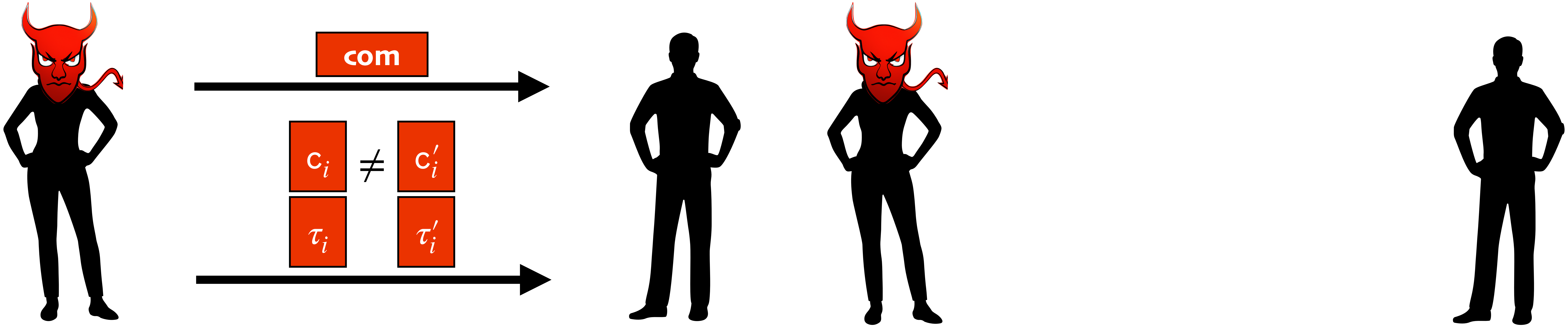
Code-Binding

Construction Framework

Erasure Code Commitments

Position-Binding

Code-Binding

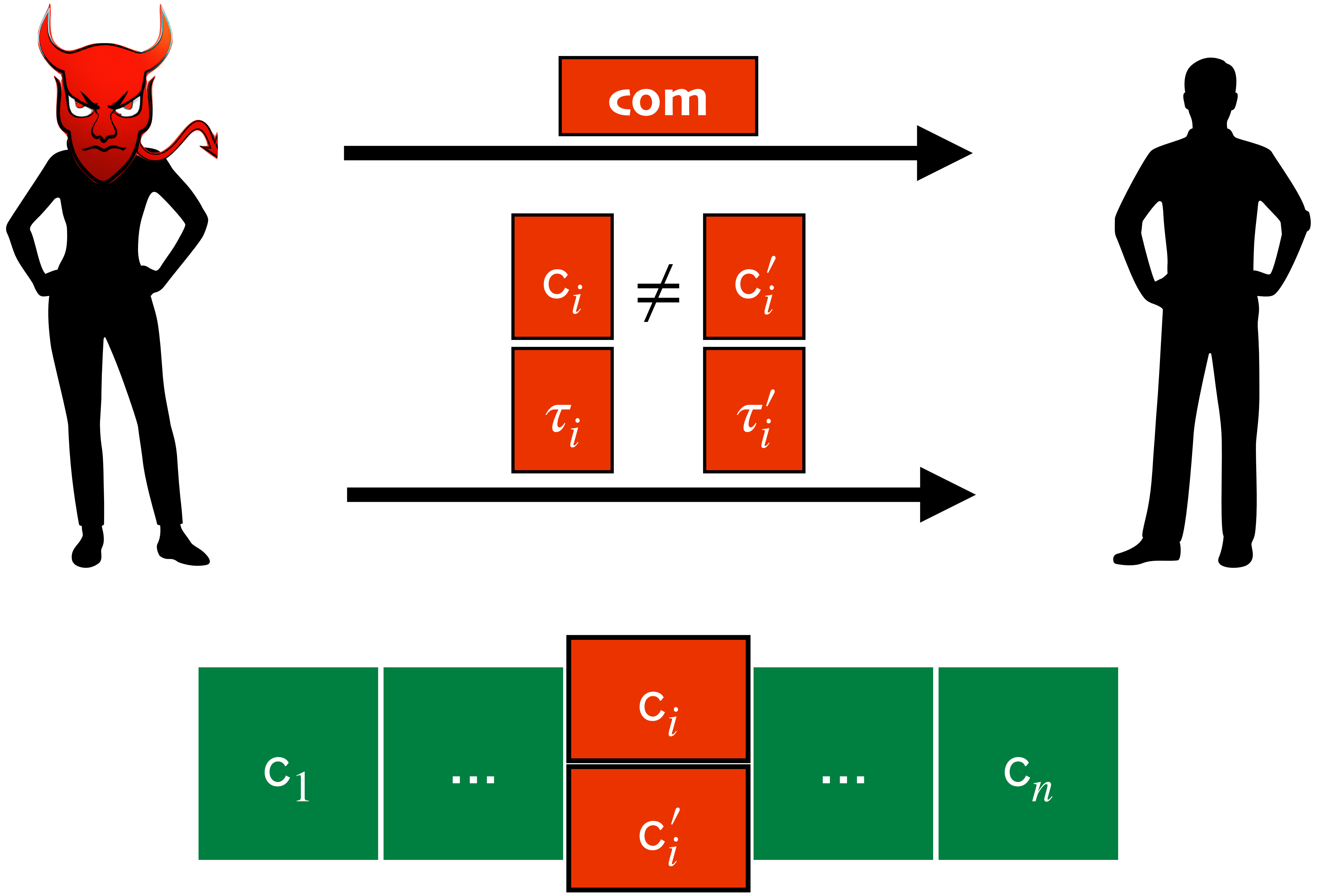


Construction Framework

Erasure Code Commitments

Position-Binding

Code-Binding

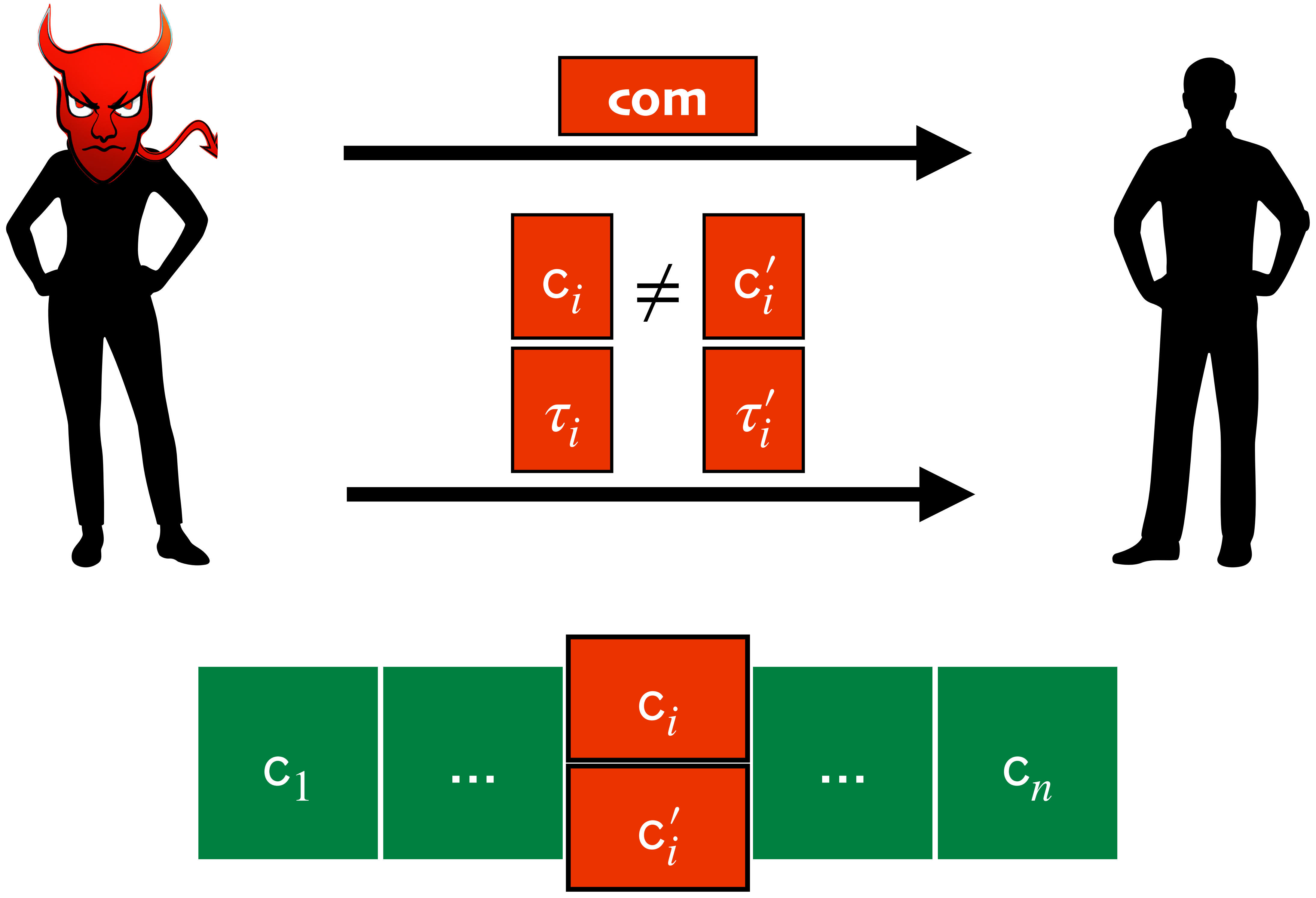


Construction Framework

Erasure Code Commitments

Position-Binding

Code-Binding



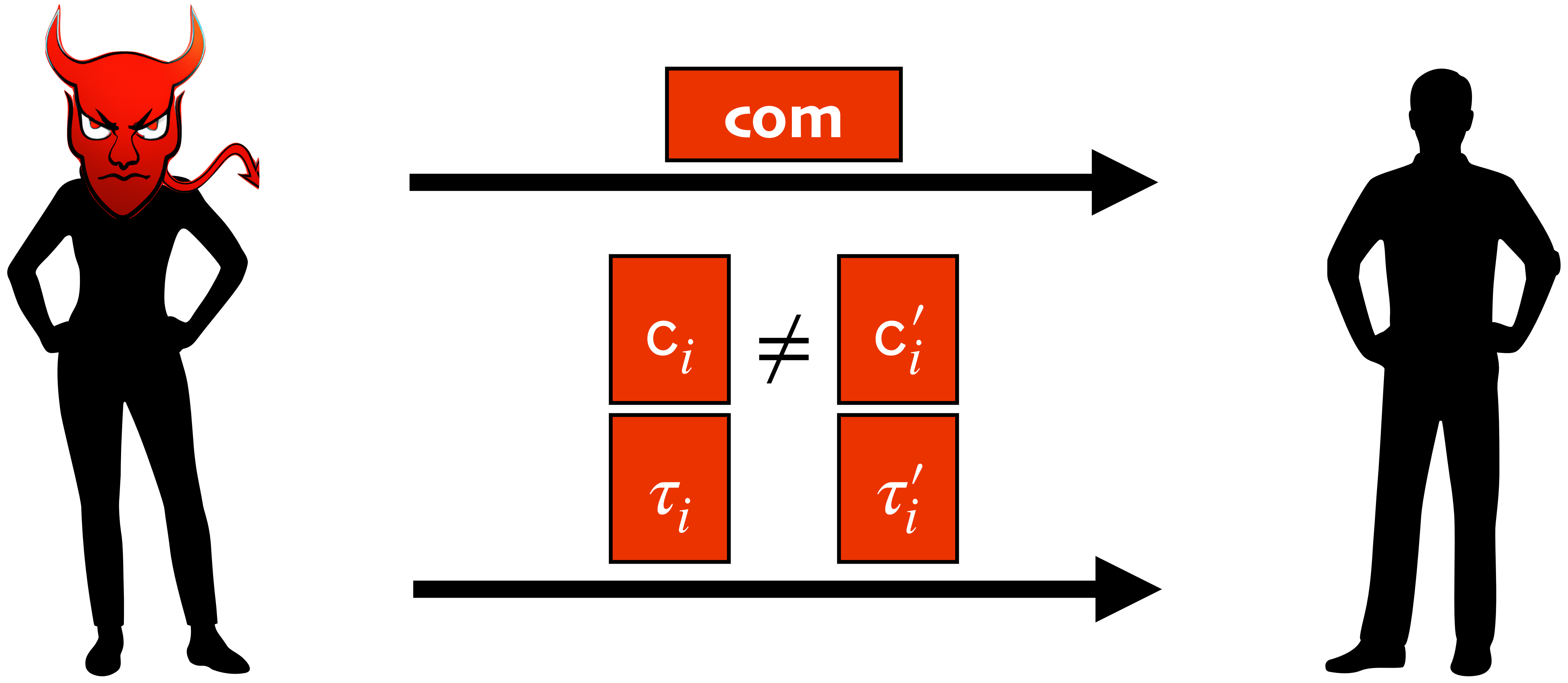
No two distinct openings
for same position

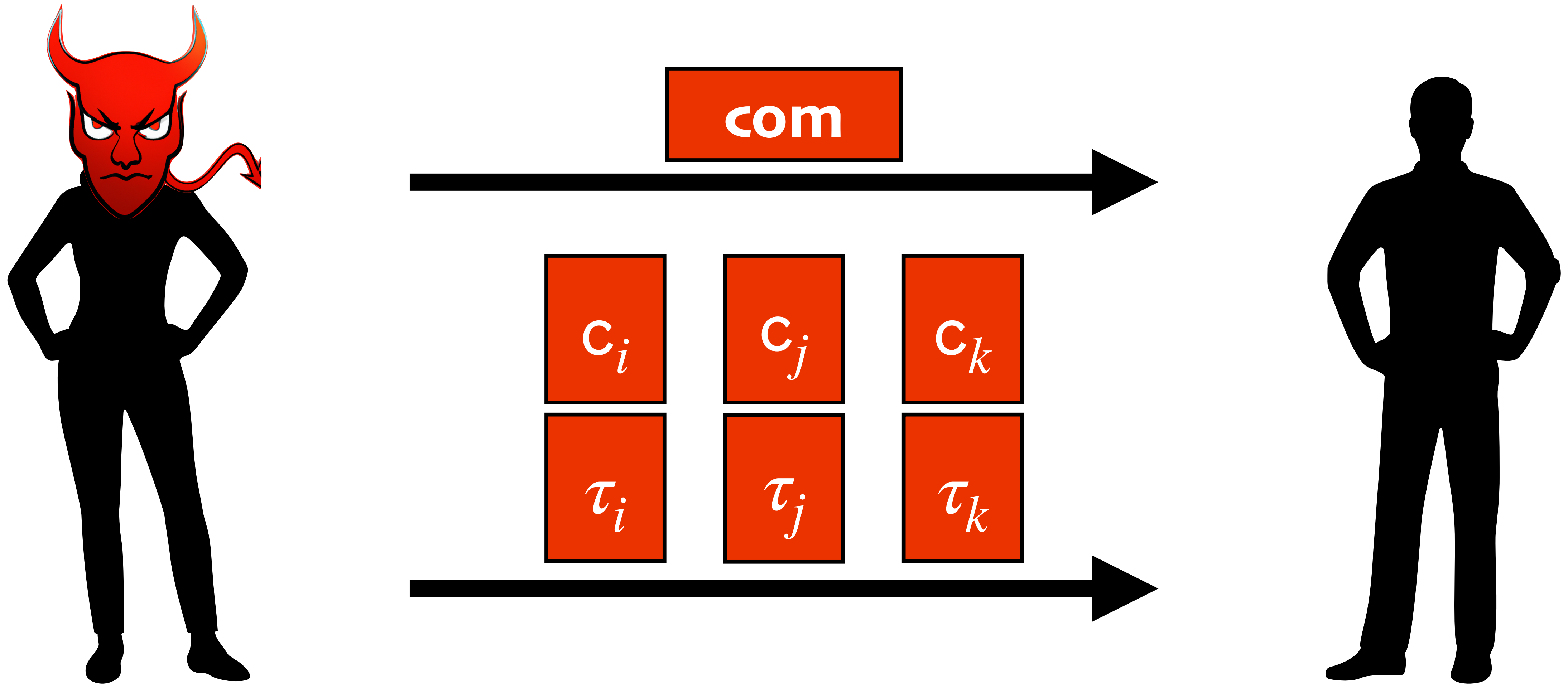
Construction Framework

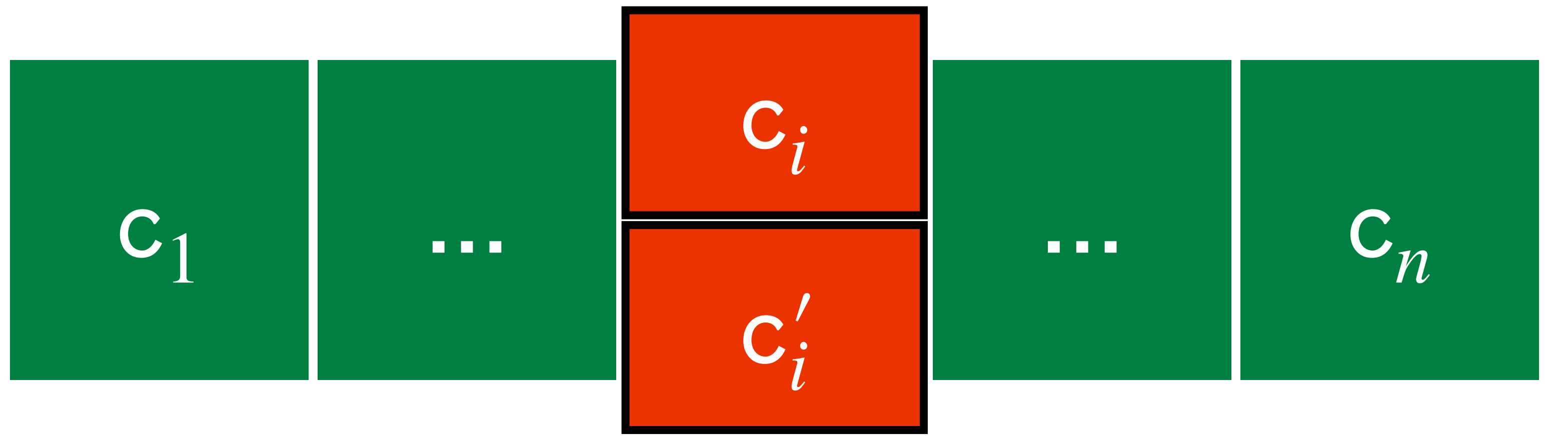
Erasure Code Commitments

Position-Binding

Code-Binding





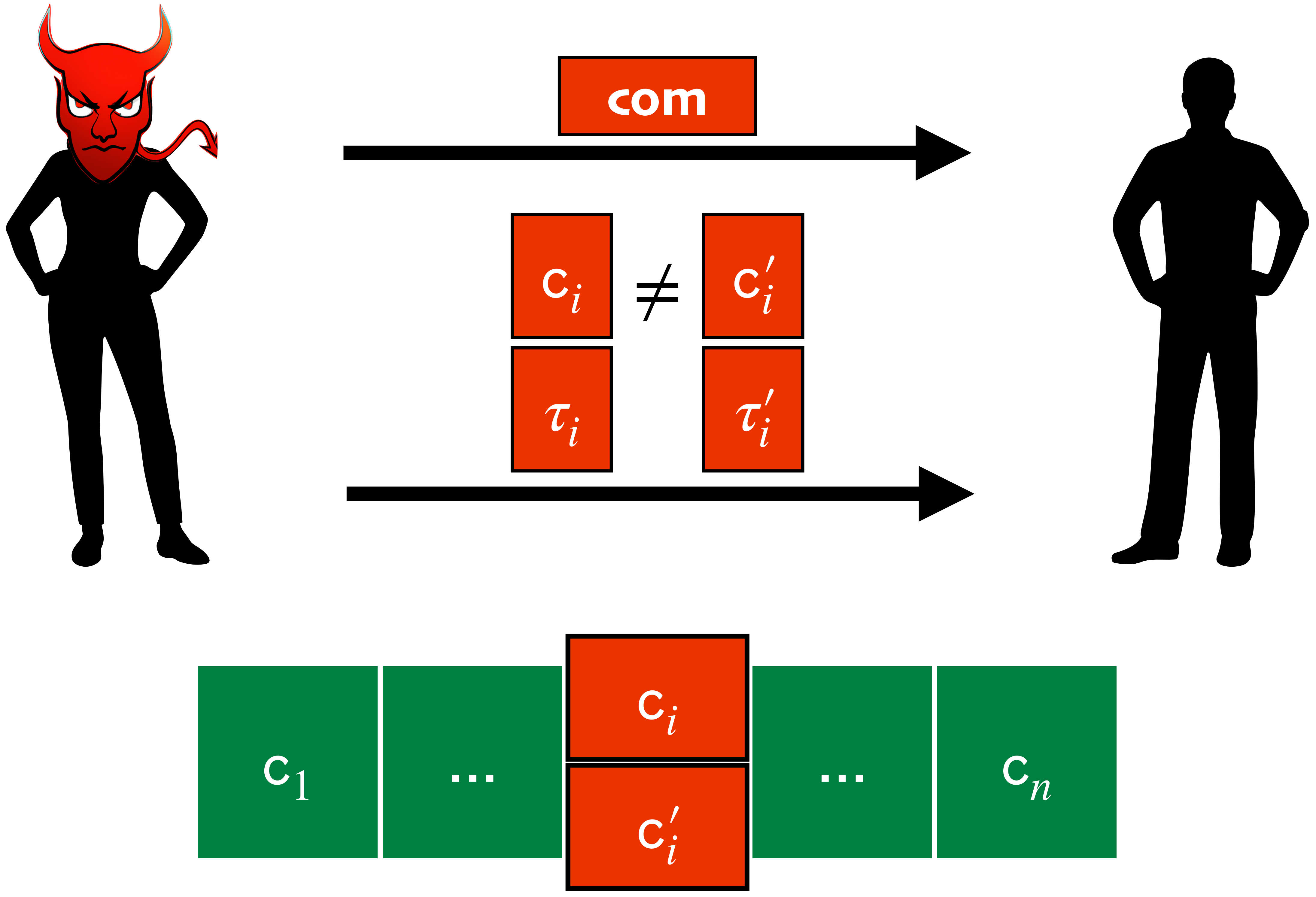


No two distinct openings
for same position

Construction Framework

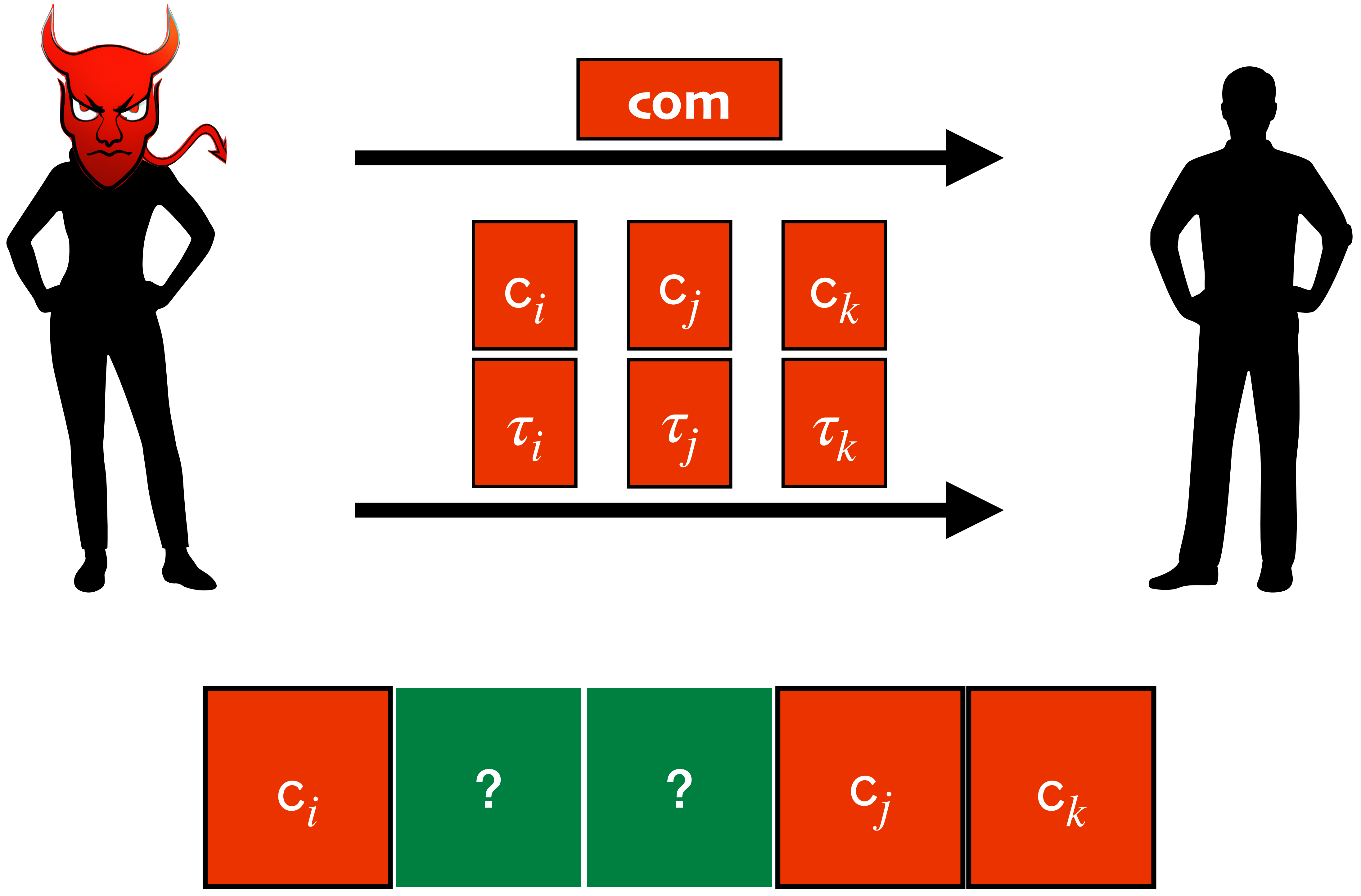
Erasure Code Commitments

Position-Binding



No two distinct openings
for same position

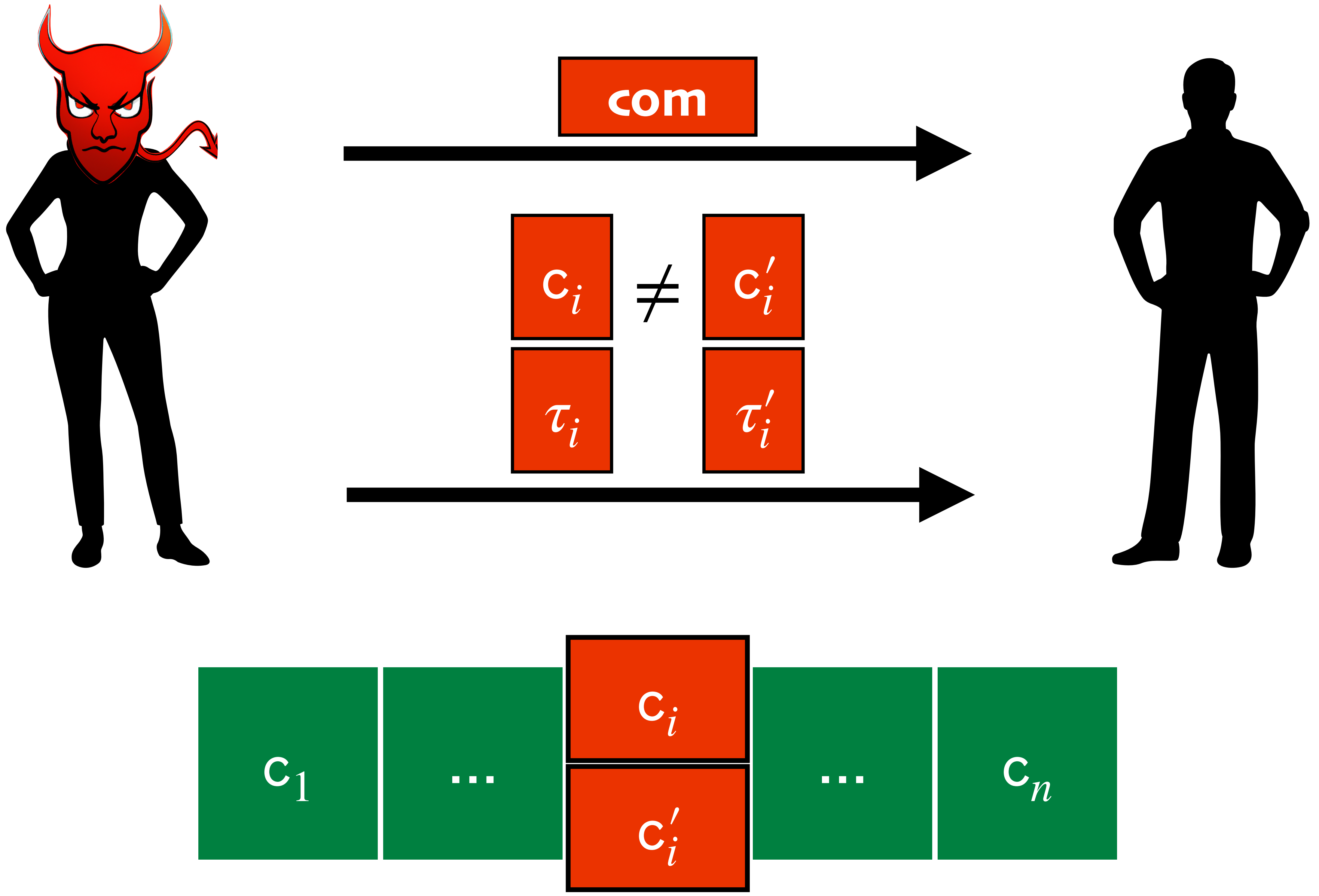
Code-Binding



Construction Framework

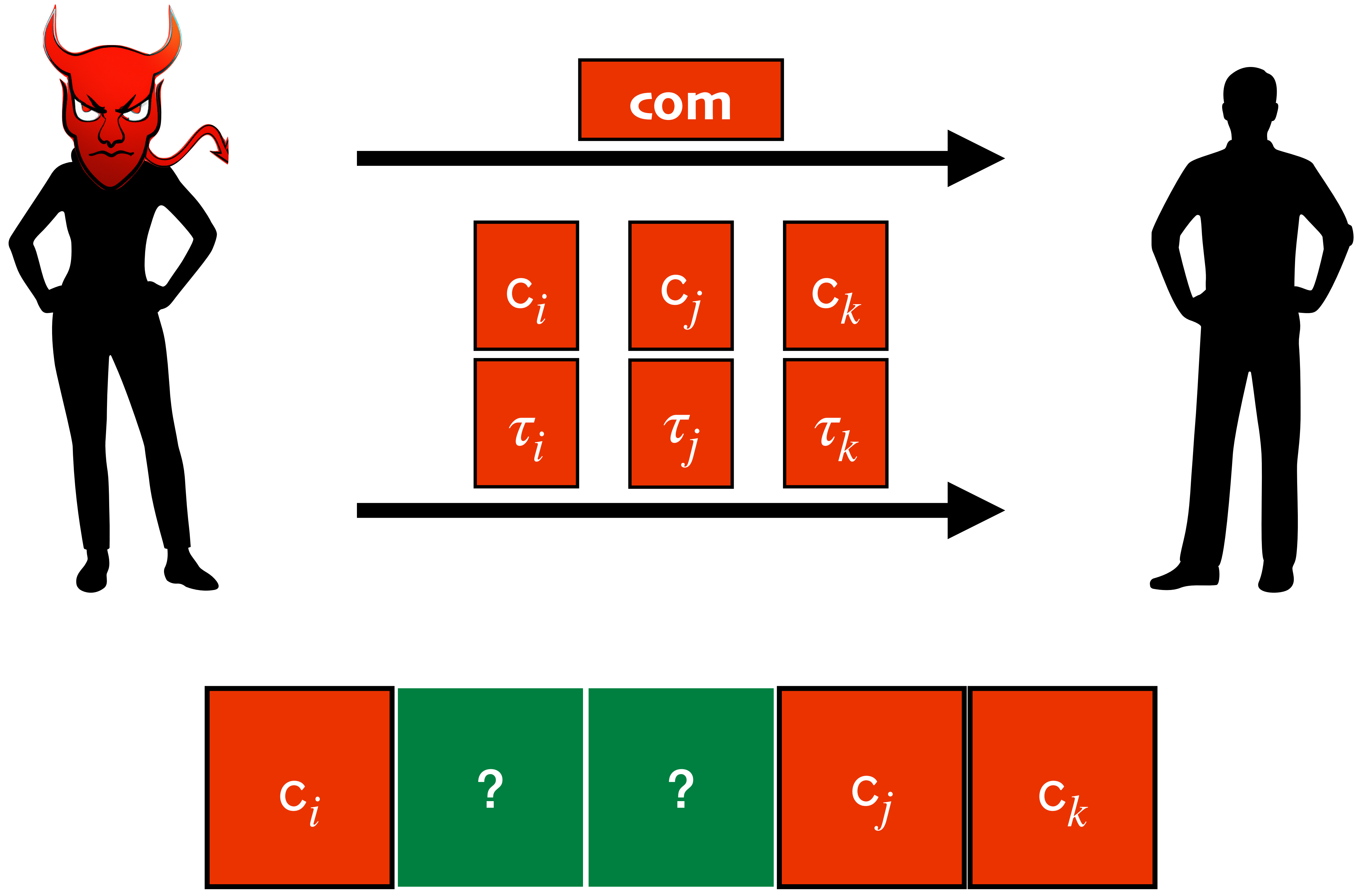
Erasure Code Commitments

Position-Binding



No two distinct openings
for same position

Code-Binding



Always consistent with
at least one codeword