Building Software Systems

Lecture 6.1

Blockchain Fundamentals – Part 1

SAURABH SRIVASTAVA

ASSISTANT PROFESSOR

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING

IIT (ISM) DHANBAD

How hard is editing something in Records?

Assume that you have a printed document

- Can you change some text in this document so that the overall meaning of the text changes?
- Probably yes – we often do it with official documents as well
- While these changes may be easier till the authority of the change lies with one person, e.g., the author …
- … it becomes more complex, when the change needs to be approved by multiple people

Editing Text on Paper – Post-Ratification

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
IIT (ISM) Dhanbad

Date: 25.06.2022

Minutes of the DPAC meeting

An urgent DPAC meeting was held on 25.06.2022 at 11:00 AM in the Conference Room of CSE Department to discuss the procurement of **Twenty All-in-one Machines for CSE laboratories.**

The following members were present:

1. Prof. [redacted] Associate Professor & Head, CSE - Chairman
2. Prof. [redacted] Assistant Professor, CSE - Member
3. Prof. [redacted] Assistant Professor, CSE - Member
4. Prof. [redacted] Assistant Professor, CSE & Indenter - Member

The DPAC observed that all of the bids received against the indent were technically qualified, as observed during the meeting held on 25.06.2022.

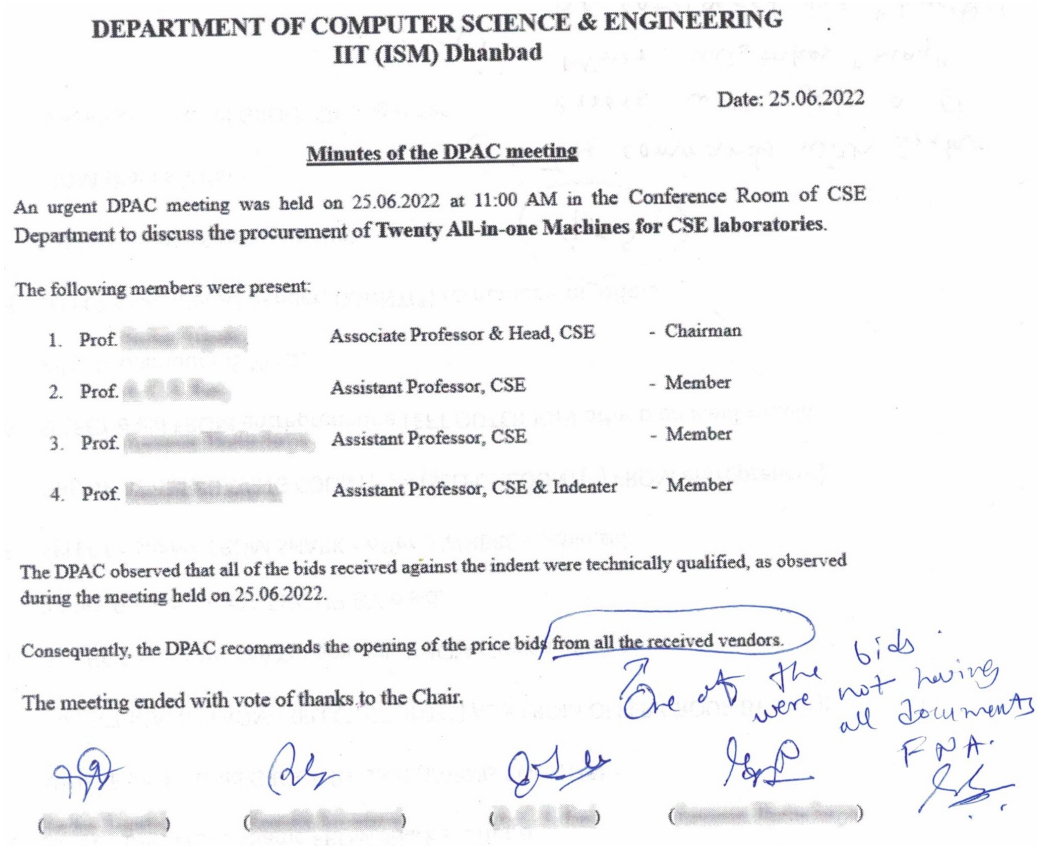
Consequently, the DPAC recommends the opening of the price bids from all the received vendors.

One of the bids were not having all documents FNA.

The meeting ended with vote of thanks to the Chair.

([redacted]) ([redacted]) ([redacted]) ([redacted])

Editing Text on Paper – Post-Ratification



DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
IIT (ISM) Dhanbad

Date: 25.06.2022

Minutes of the DPAC meeting

An urgent DPAC meeting was held on 25.06.2022 at 11:00 AM in the Conference Room of CSE Department to discuss the procurement of **Twenty All-in-one Machines for CSE laboratories.**

The following members were present:

1. Prof. [redacted] Associate Professor & Head, CSE - Chairman
2. Prof. [redacted] Assistant Professor, CSE - Member
3. Prof. [redacted] Assistant Professor, CSE - Member
4. Prof. [redacted] Assistant Professor, CSE & Indenter - Member

The DPAC observed that all of the bids received against the indent were technically qualified, as observed during the meeting held on 25.06.2022.

Consequently, the DPAC recommends the opening of the price bids from all the received vendors.

One of the bids were not having all documents FNA.

The meeting ended with vote of thanks to the Chair.

([redacted]) ([redacted]) ([redacted]) ([redacted])

In most organisations, editing documents post its ratification from all stakeholders is either not permitted, or requires an approval from all the stakeholders !!

How hard is editing something in Records?

Assume that you have a printed document

- Can you change some text in this document so that the overall meaning of the text changes?
- Probably yes – we often do it with official documents as well
- While these changes may be easier till the authority of the change lies with one person, e.g., the author, …
- … it becomes more complex, when the change needs to be approved by multiple people
- Nevertheless, humans are used to paper for a very long time, and we often handle such issues in the *process*

The digital world of documents has its own pros and cons

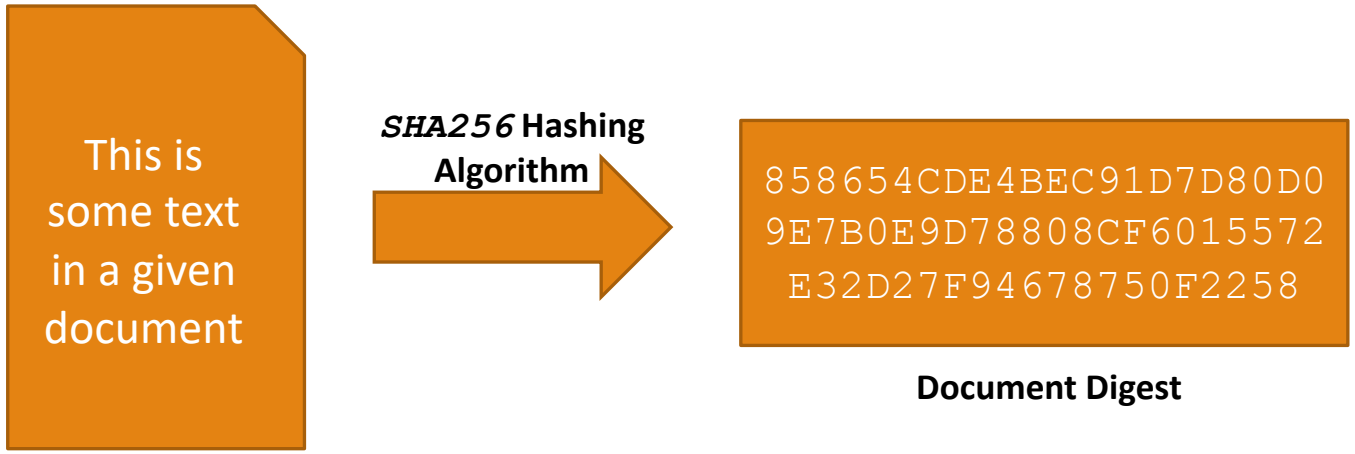
- While it easier to make changes to the documents – which may be as simple as editing a file in *MS Word*, …
- … it does have its own problems as well (in fact, even a *PDF* file might not be as immutable as it seems !!)
- For example, someone with access to the document, may modify it with a malicious or corrupt intent as well
- Thus, we need a "layer of protection" over digital documents, which prevents their ad-hoc modifications
- Keep in mind that we are not supposed to restrict the advantages of easy modifications in the digital world, …
- … we are only expected to come up with a "system of change" that matches an Organisation's needs

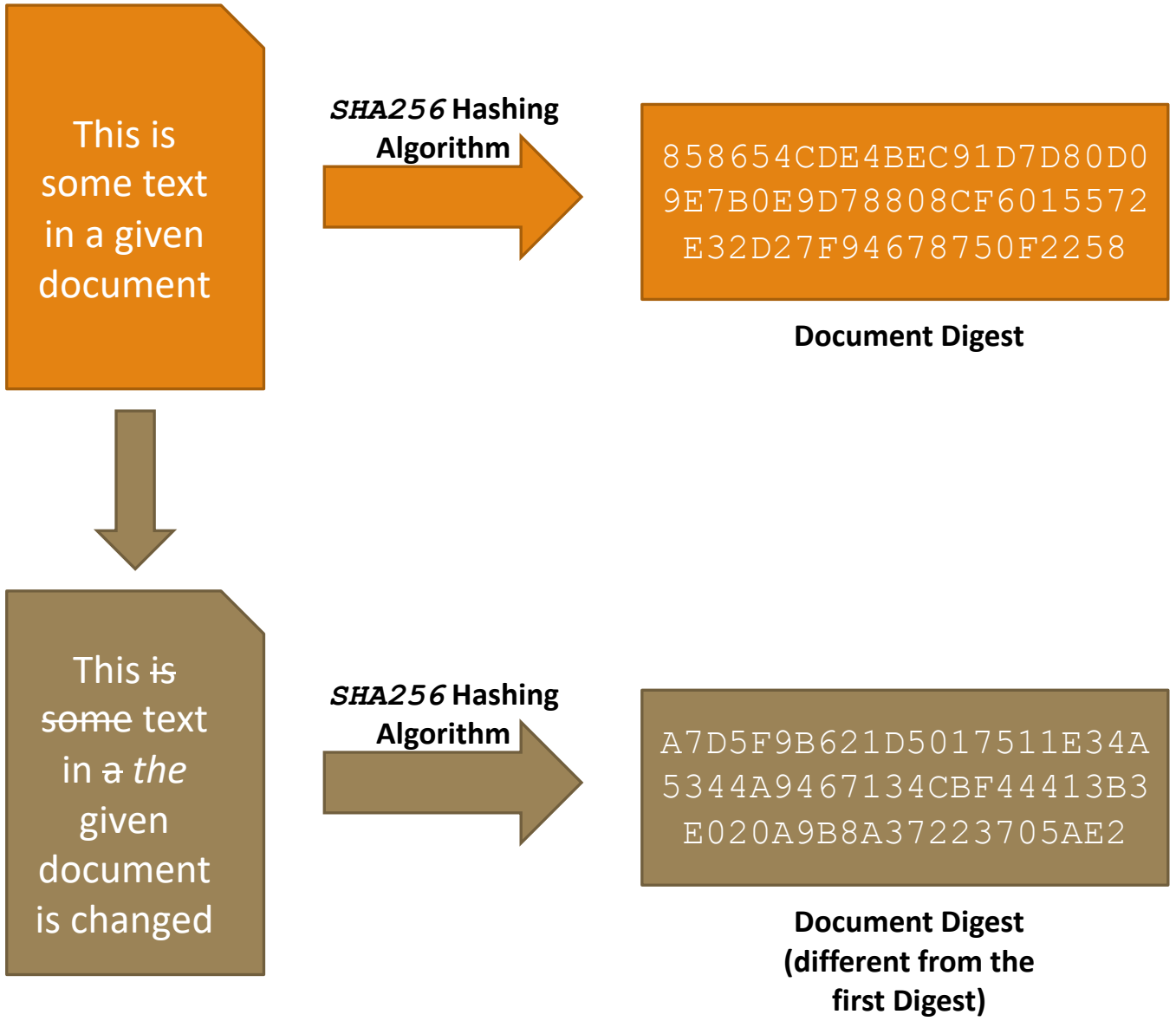
Technologies That Protect Change

Certain technologies from the Public Key Infrastructure (PKI) domain, which are useful for this protection

Hashes

- A hash is essentially a digest of a chunk of data
- The data may be in textual form, or in any other intelligible form – its hash can be computed
- A small change to the initial data, should ideally lead to a large change in the computed hash


This is some text in a given document
SHA256 Hashing Algorithm
858654CDE4BEC91D7D80D0
9E7B0E9D78808CF6015572
E32D27F94678750F2258
Document Digest



Any changes to the document will not yield the same Digest !!

Technologies That Protect Change

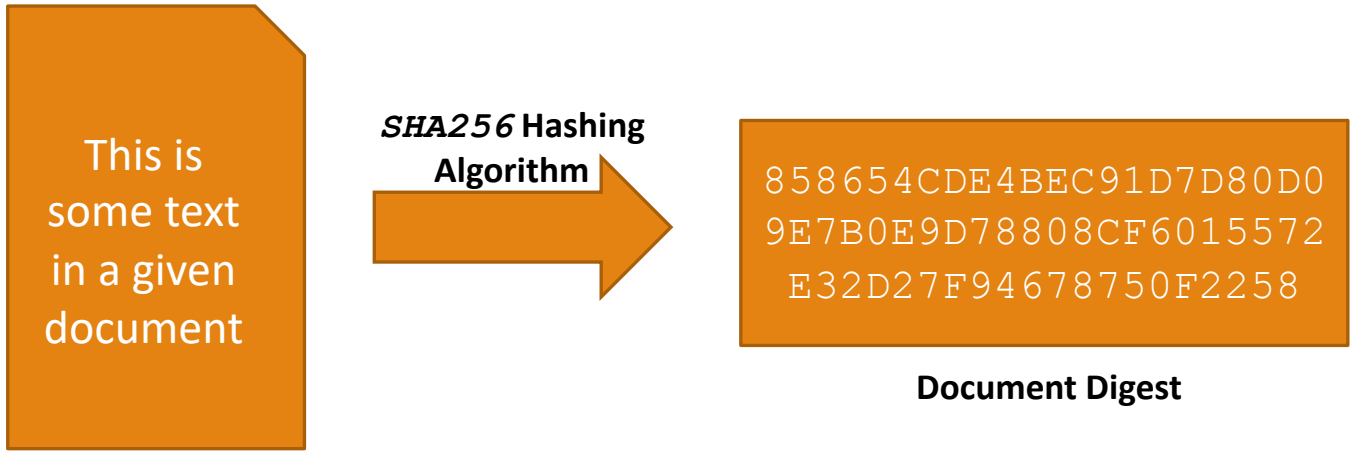
Certain technologies from the Public Key Infrastructure (PKI) domain, which are useful for this protection

Hashes

- A hash is essentially a digest of a chunk of data
- The data may be in textual form, or in any other intelligible form – its hash can be computed
- A small change to the initial data, should ideally lead to a large change in the computed hash

Public Key Encryption

- Encryption is a strategy for “hiding” data from “unauthorised” access; while encryption doesn’t protect changes …
- … Public Key Encryption (PKE), along with Hashing, provides a descent protection against changes


This is some text in a given document
SHA256 Hashing Algorithm
858654CDE4BEC91D7D80D0
9E7B0E9D78808CF6015572
E32D27F94678750F2258
Document Digest

The *Private Key* is known to only the signer of the document – so only the signer can produce this encrypted digest

Technologies That Protect Change

Certain technologies from the Public Key Infrastructure (PKI) domain, which are useful for this protection

Hashes

- A hash is essentially a digest of a chunk of data
- The data may be in textual form, or in any other intelligible form – its hash can be computed
- A small change to the initial data, should ideally lead to a large change in the computed hash

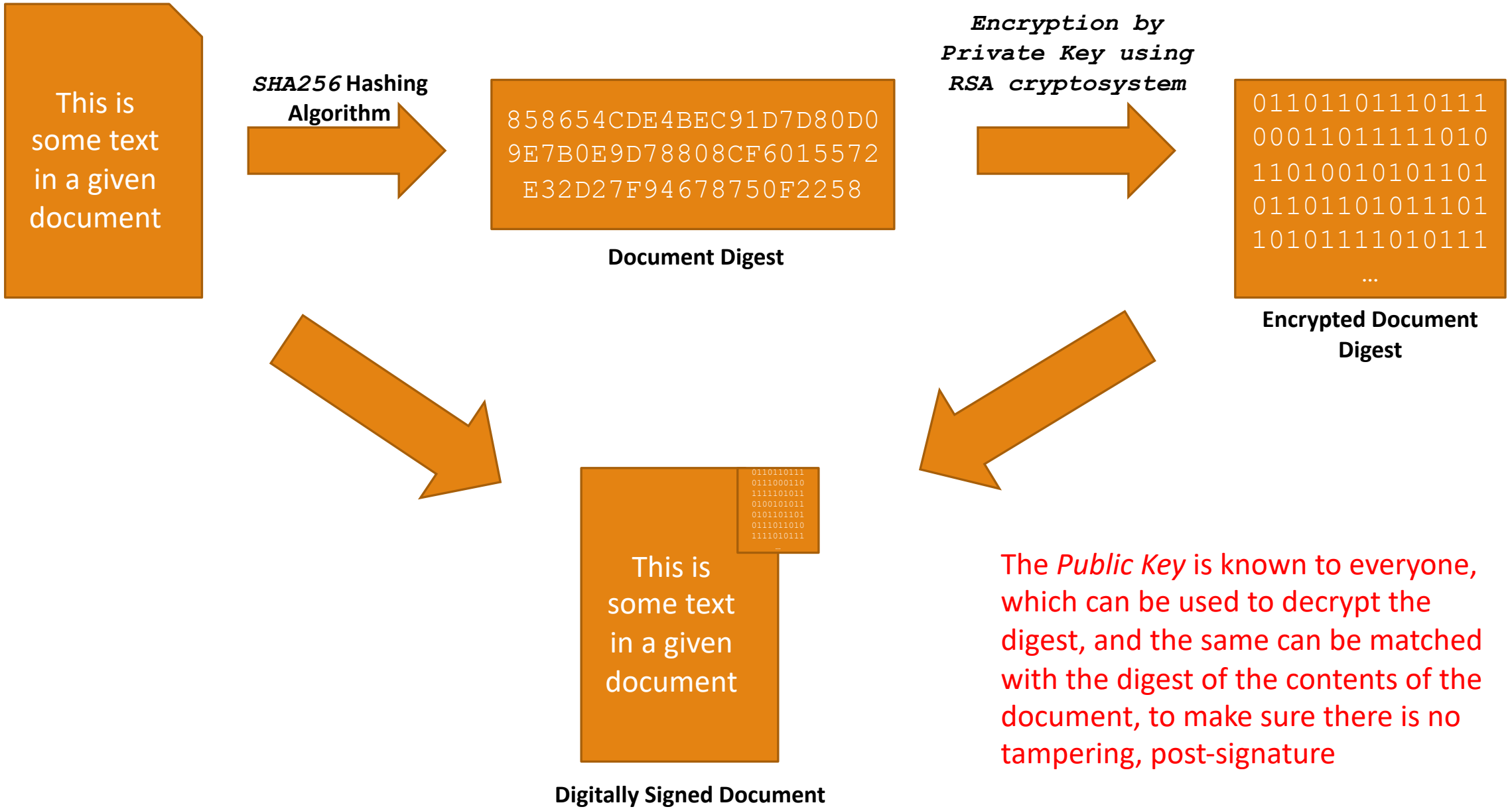
Public Key Encryption

- Encryption is a strategy for “hiding” data from “unauthorised” access; while encryption doesn’t protect changes …
- … Public Key Encryption (PKE), along with Hashing, provides a descent protection against changes

Digital Signatures

- Digital Signatures are a way by which a file’s content is “digitally ratified” by an individual or organisation
- It is performed by first finding the hash of the document, proceeded by encrypting the same through PKE
- This encrypted digest is attached to the document, which may be verified by others to check for modifications


This is some text in a given document
SHA256 Hashing Algorithm
858654CDE4BEC91D7D80D0
9E7B0E9D78808CF6015572
E32D27F94678750F2258
Document Digest
Encryption by Private Key using RSA cryptosystem
01101101110111
00011011111010
11010010101101
01101101011101
10101111010111
...
Encrypted Document Digest



The *Public Key* is known to everyone, which can be used to decrypt the digest, and the same can be matched with the digest of the contents of the document, to make sure there is no tampering, post-signature

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is based on the previous block's data. As a result, if the data is changed in one block, it's unique identifier changes, which can be seen in every subsequent block (providing tamper evidence).

Source: https://www.nist.gov/blockchain

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is based on the previous block's data. As a result, if the data is changed in one block, it's unique identifier changes, which can be seen in every subsequent block (providing tamper evidence).

Source: https://www.nist.gov/blockchain

Let us go through this definition piece-by-piece !!

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are connected to identifier that is based on the previous block's data. As a result, if the data is changed in one block, it's unique identifier changes, which can be seen in every subsequent block (providing tamper evidence).

A ledger in this context is a digital record of transactional data; it is a term that is often used by people in Accounting to refer to a book that keeps track of debits and credits of money in an account (or for an organisation as a whole) !!

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are connected to identifier that is based on the previous block's data. As a result, if the data identifier changes, which can be seen in every subsequent block (providing tamper evidence).

A ledger in this context is a digital record of transactional data; it is a term that is often used by people in Accounting to refer to a book that keeps track of debits and credits of money in an account (or for an organisation as a whole) !!

As per the definition, a blockchain is a "collaborative, temper-resistant ledger"

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) connected to identifier that is based on the previous block's data. As a result, if the da identifier changes which can be seen in every subsequent blo

A ledger in this context is a digital record of transactional data; it is a term that is often used by people in Accounting to refer to a book that keeps track of debits and credits of money in an account (or for an organisation as a whole) !!

As per the definition, a blockchain is a "collaborative, temper-resistant ledger"

It is collaborative, because it is maintained jointly by multiple parties, and it is temper-resistant, because it adopts a process for changing already existing data; without following the proper protocol, even if changes are attempted, they will be rejected !!

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is [illegible] result, if the [illegible] identifier changes, which can be seen in every subsequent block (providing tamper evidence).

Several transactions are grouped together on this ledger, to form a unit of data *addition* (or in certain cases, following the adopted protocol, the *modification*) on the blockchain

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is [illegible] result, if the [illegible] identifier changes, which can be seen in every subsequent blo[illegible]

Several transactions are grouped together on this ledger, to form a unit of data *addition* (or in certain cases, following the adopted protocol, the *modification*) on the blockchain

The number of transactions on each block may differ widely, and is a function of how often new data arrives in the system, and how critical, is the delay in the addition of the new data to the blockchain; As an example, on a typical block, there are around 2000 transactions in case of *bitcoin* (https://bitcoin.org/)

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is based on the previous block's data. As a result, if the da identifier char subsequent blo

This identifier has some properties; one of these properties is that if the content of the block (which is essentially the concatenation of transactional data for all the transactions in the block, plus, something more) is changed, this identifier changes too !!

NIST Definition of Blockchain

A blockchain is a collaborative, tamper-resistant ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is based on the previous block's data. As a result, if the data in a block changes, the unique identifier changes and all subsequent blocks are affected.

This identifier has some properties; one of these properties is that if the content of the block (which is essentially the concatenation of transactional data for all the transactions in the block, plus, something more) is changed, this identifier changes too !!

The concept of a **hash** is central to the computation of this identifier; in fact, it is actually a hash of the block's data, with some additional constraints !!

NIST Definition of Blockchain

As we said earlier, the data over which the identifier is calculated has the transactional data, plus "something" more …

A blockchain is a type of distributed ledger that maintains transactional records. The transactional records (data) are grouped into blocks. A block is connected to the previous one by including a unique identifier that is based on the previous block's data. As a result, if the data is changed in one block, it's unique identifier changes, which can be seen in every subsequent block (providing tamper evidence).

NIST Definition of Blockchain

As we said earlier, the data over which the identifier is calculated has the transactional data, plus "something" more …

… this something more, is actually the identifier from the *previous* block; this means that the identifier of block i is dependent on the identifier of block $i-1$, creating a chain between the blocks

A blockchain is a tamper evident ledger that maintains transactional records. The transactional records (data) in each block are connected to the previous one by including a unique identifier that is based on the previous block's data. As a result, if the data is changed in one block, it's unique identifier changes, which can be seen in every subsequent block (providing tamper evidence).

NIST Definition of Blockchain

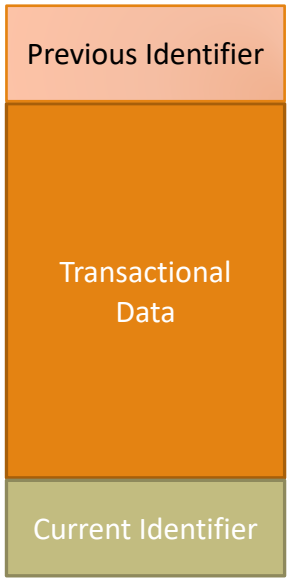
A blockchain is a type of distributed ledger that maintains transactional records. The transactional records (data) in each block are connected to the previous one by, including a unique identifier that is based on the previous block's data. *As a result, if the data is changed in one block, it's unique identifier changes, which can be seen in every subsequent block (providing tamper evidence).*

As we said earlier, the data over which the identifier is calculated has the transactional data, plus "something" more …

… this something more, is actually the identifier from the *previous* block; this means that the identifier of block i is dependent on the identifier of block $i-1$, creating a chain between the blocks

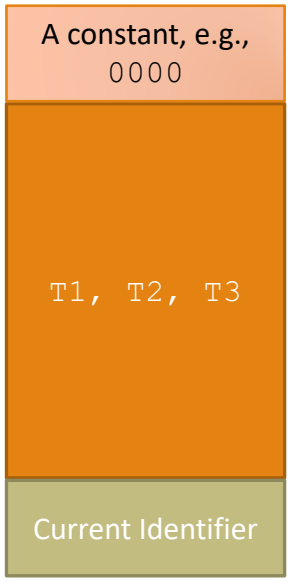
Any change in the block j, would thus require making a corresponding change in block $j+1$, which in turn, would require a change in block $j+2$, and so on !!

Pictorial Representation of a Blockchain



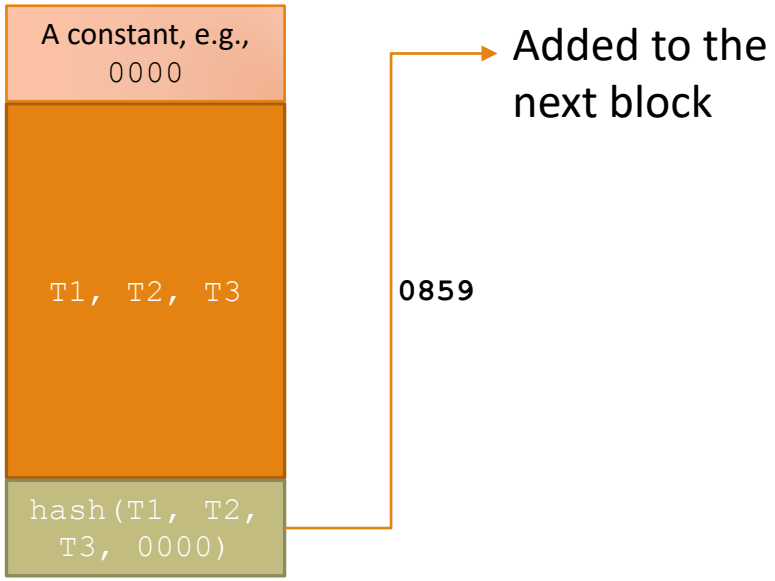
This is a simplified view of a typical block in a blockchain

Pictorial Representation of a Blockchain



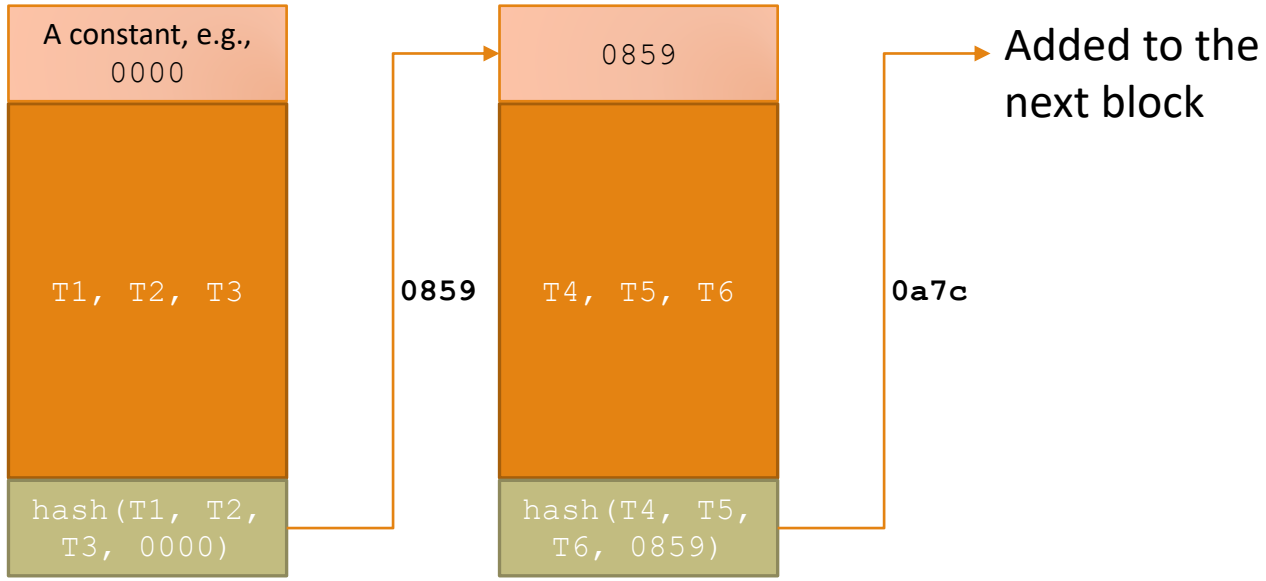
The transactions are added to a block,
as and when they enter the system

Pictorial Representation of a Blockchain



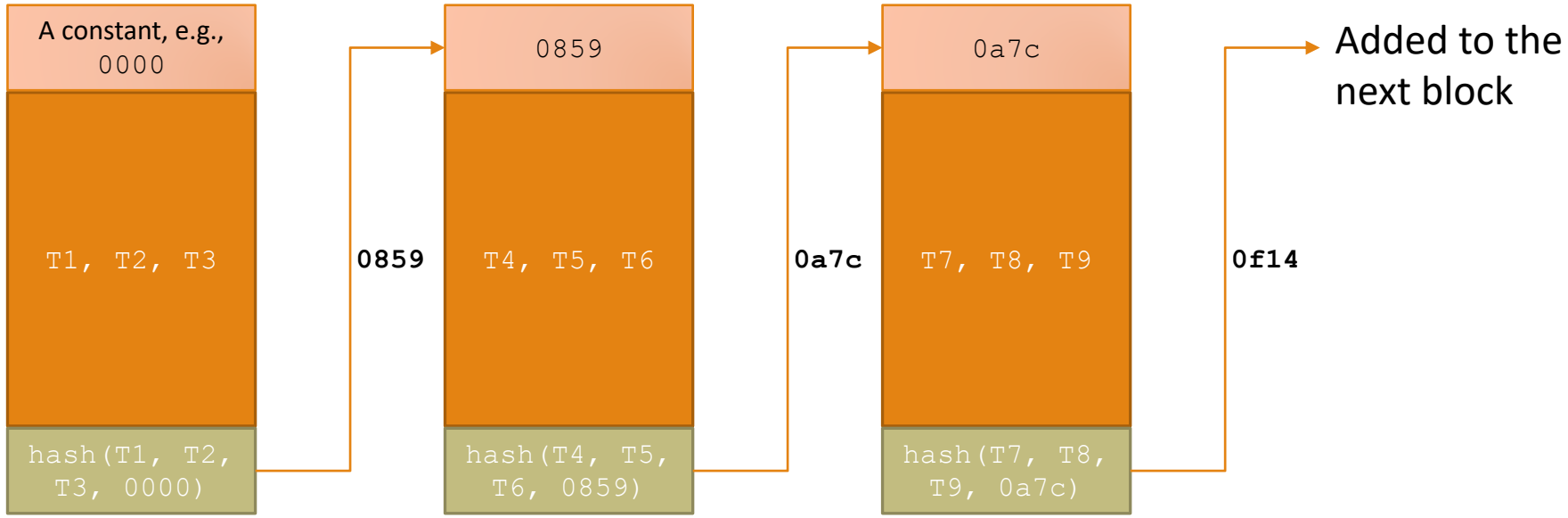
When the block is “full” (or, it has been sufficient amount of time since the last block was added to the blockchain), the identifier is calculated (which is a hash, with some other properties)

Pictorial Representation of a Blockchain



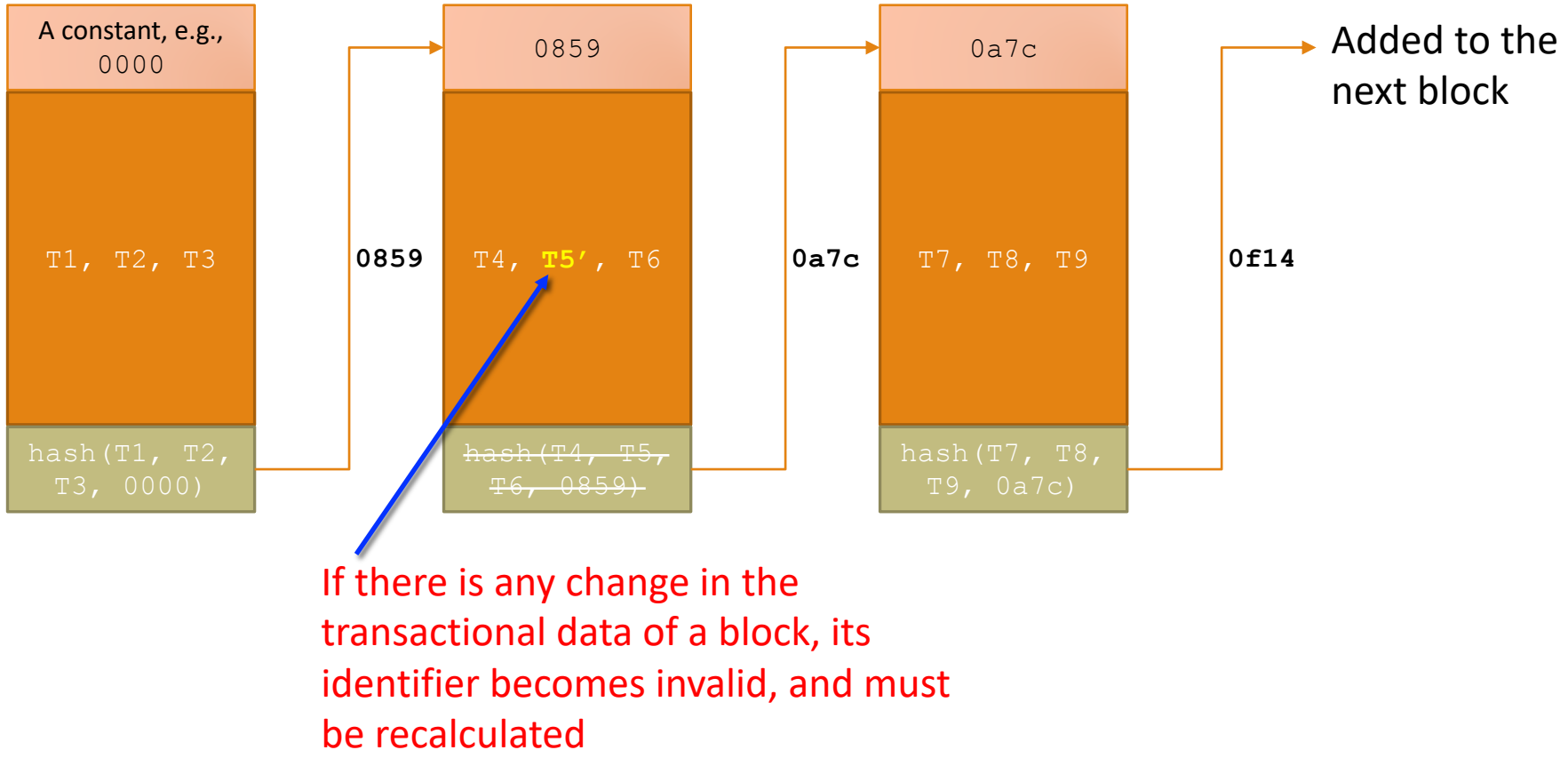
The chain of blocks continue getting added at the back of this chain …

Pictorial Representation of a Blockchain

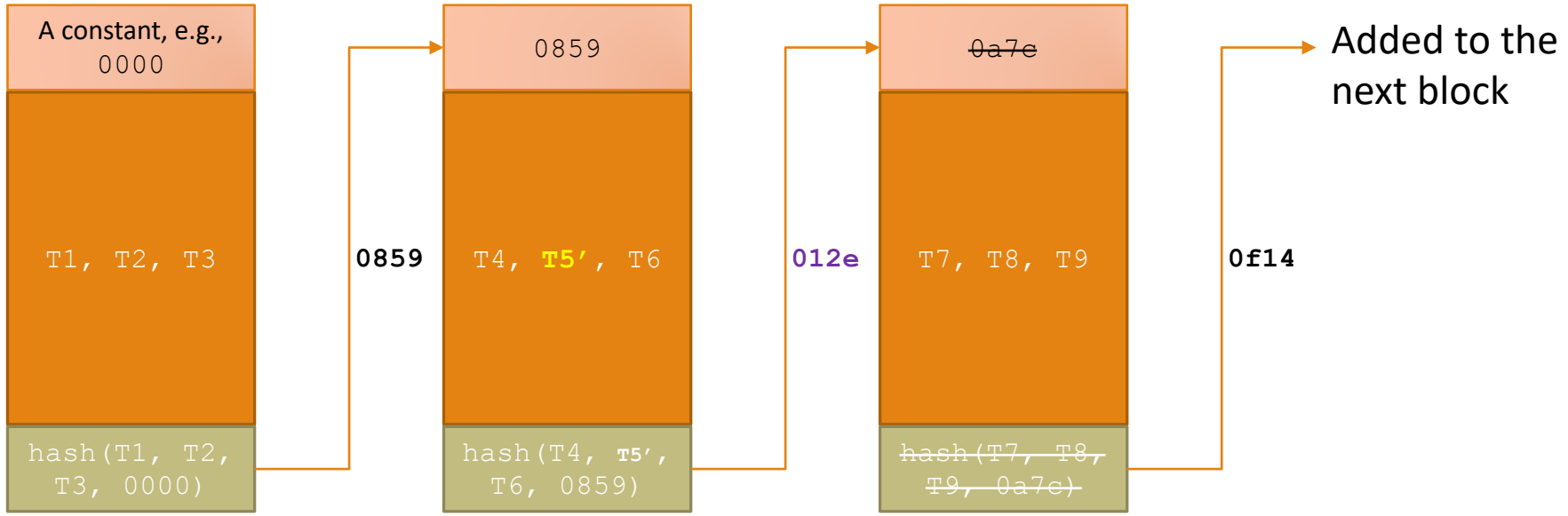


… with each block requiring the identifier from the previous block for its own identifier's collection !!

Pictorial Representation of a Blockchain

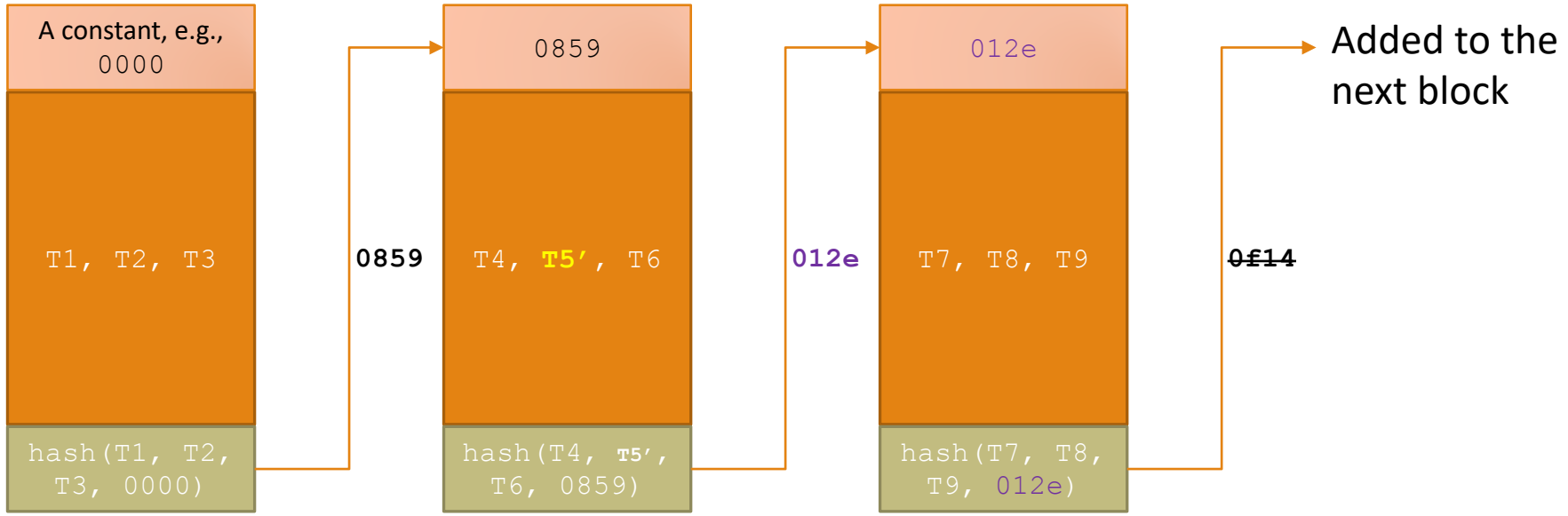


Pictorial Representation of a Blockchain



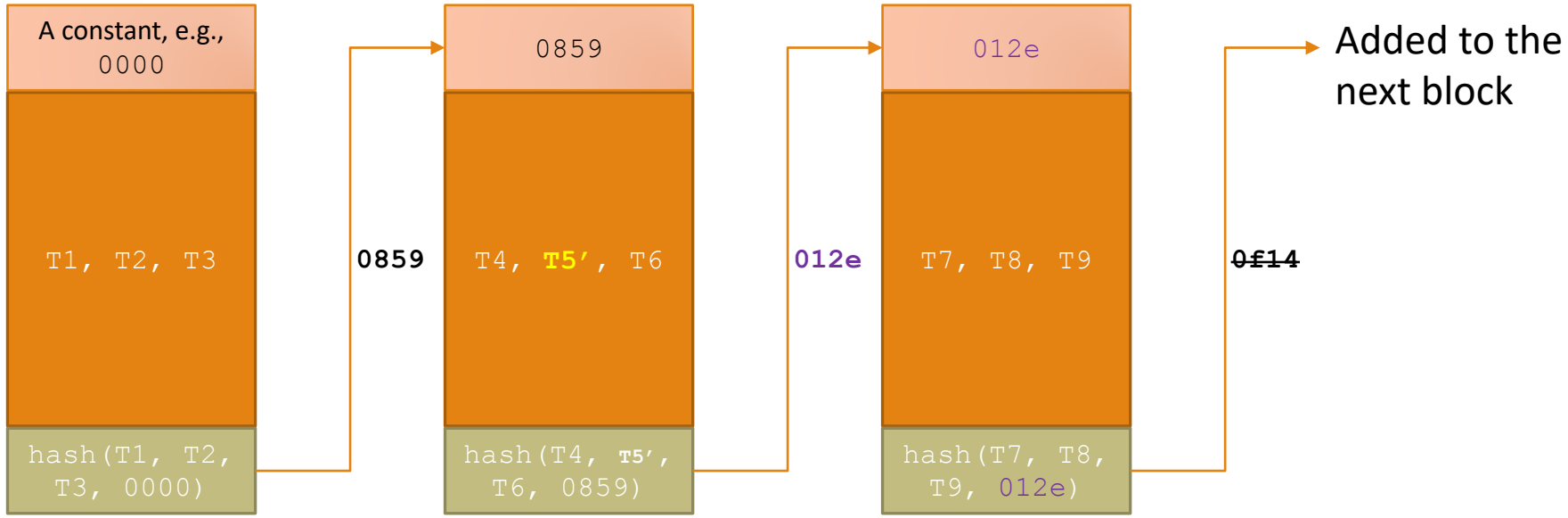
This recomputed identifier must now be updated in the next block, which will result in the need for re-computing the identifier for the next block

Pictorial Representation of a Blockchain



This chain continues till the most recently added block is also changed !!

Pictorial Representation of a Blockchain



This chain continues till the most recently added block is also changed !!

If a blockchain has sufficiently high number of blocks, this chain reaction makes a change difficult !!
(for perspective, a new block gets added to the *bitcoin* blockchain every 10 minutes or so on average)