Lecture #5: Byzantine Faults and Digital Signature Schemes

> COMS 4995-001: The Science of Blockchains URL: https://timroughgarden.org/s25/

> > Tim Roughgarden

#### Goals for Lecture #5

- 1. The challenges of Byzantine faults.
  - faulty validators that can behave in arbitrary (worst-case) ways
- 2. Digital signature schemes.
  - key tool for limiting the space of Byzantine validator strategies
- 3. Limits on what is achievable.
  - Byzantine faults make the SMR problem harder in partial synchrony
- 4. Key ideas behind Tendermint.
  - Full protocol description and analysis on Monday.

### State Machine Replication (SMR)

SMR: version of consensus appropriate for a blockchain protocol.

- "state machine" = for us, current state of virtual machine
- "replication" = all validators perform same state transitions
- "clients" submit transactions ("txs") to validators
- each validator maintains an append-only list of finalized txs (a.k.a. "log" or "history")

Goal: a protocol that satisfies consistency and liveness.



3

#### A Road Map to Practical SMR Protocols



Lecture #3: Protocol B solves SMR with crash faults in synchrony.

Lecture #4: if strict majority of validators are non-faulty, Protocol C (≈ Paxos/Raft) solves SMR with crash faults in partial synchrony.

Next challenge: Byzantine faults.

• faulty validators can act arbitrarily

Next challenge: Byzantine faults.

- faulty validators can act arbitrarily
  - original motivation (1980s): hard-to-model software errors
  - blockchain protocols: might literally get attacked by hostile actor
    - e.g., hacks into validators previously controlled by good actors
  - alternative names for non-faulty validators: "honest," "correct"

Next challenge: Byzantine faults.

- faulty validators can act arbitrarily
  - original motivation (1980s): hard-to-model software errors
  - blockchain protocols: might literally get attacked by hostile actor
    - e.g., hacks into validators previously controlled by good actors
  - alternative names for non-faulty validators: "honest," "correct"
  - question: what are Byzantine validators capable of?

Next challenge: Byzantine faults.

- faulty validators can act arbitrarily
  - original motivation (1980s): hard-to-model software errors
  - blockchain protocols: might literally get attacked by hostile actor
    - e.g., hacks into validators previously controlled by good actors
  - alternative names for non-faulty validators: "honest," "correct"
  - question: what are Byzantine validators capable of?

Next challenge: Byzantine faults (i.e., faulty validators can act arbitrarily).

Next challenge: Byzantine faults (i.e., faulty validators can act arbitrarily).



Next challenge: Byzantine faults (i.e., faulty validators can act arbitrarily).



Next challenge: Byzantine faults (i.e., faulty validators can act arbitrarily).



Next challenge: Byzantine faults (i.e., faulty validators can act arbitrarily).



Question: is Protocol C still live and consistent with Byzantine faults?

 key property for consistency: read quorum must intersect write quorums from all previous views → if leader makes a proposal, must be up-to-date



Issue #1: Byzantine leader could ignore read quorum requirement and make an (out-of-date) proposal anyway.

- maybe didn't receive chains from > n/2 validators, or maybe it did and chose to ignore them
- out-of-date proposal (if adopted) → consistency violation

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

**Issue #2:** Byzantine leader could propose different chains to different validators [a.k.a. "equivocation"].

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

**Issue #2:** Byzantine leader could propose different chains to different validators [a.k.a. "equivocation"].

but can't only one proposal garner the necessary >n/2 acks?

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

**Issue #2:** Byzantine leader could propose different chains to different validators [a.k.a. "equivocation"].

- but can't only one proposal garner the necessary >n/2 acks?
- no: Byzantine validators can ack multiple proposals
  - non-faulty validators might simultaneous finalize inconsistent chains

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

Issue #2: Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

Issue #2: Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)

**Issue #3:** Byzantine validators could lie about messages received from other validators.

- e.g., frame a non-faulty validator for its own misbehavior
- will tackle this issue with cryptography (next)

### **Digital Signature Schemes in Blockchains**

• one of the two most ubiquitous cryptographic primitives used in blockchain protocols (along with cryptographic hash functions)

### **Digital Signature Schemes in Blockchains**

• one of the two most ubiquitous cryptographic primitives used in blockchain protocols (along with cryptographic hash functions)

Application #1: allows a user of a blockchain to authorize a transaction (e.g., making a payment).

• fundamental to the vision of shared computer in the sky

## Digital Signature Schemes in Blockchains

• one of the two most ubiquitous cryptographic primitives used in blockchain protocols (along with cryptographic hash functions)

Application #1: allows a user of a blockchain to authorize a transaction (e.g., making a payment).

• fundamental to the vision of shared computer in the sky

Application #2: under the hood, allows validators of a blockchain protocol to sign their messages.

- used in most blockchain protocols for this purpose
  - with Bitcoin a notable exception

- 1. Key generation algorithm: maps seed  $r \rightarrow (pk,sk)$  pair.
  - in some cases, may generate r itself (e.g., ssh-keygen)

- 1. Key generation algorithm: maps seed  $r \rightarrow (pk,sk)$  pair.
  - in some cases, may generate r itself (e.g., ssh-keygen)
- 2. Signing algorithm: maps message + sk  $\rightarrow$  signature.
  - signature depends on both sk and the message being signed

- 1. Key generation algorithm: maps seed  $r \rightarrow (pk,sk)$  pair.
  - in some cases, may generate r itself (e.g., ssh-keygen)
- 2. Signing algorithm: maps message + sk  $\rightarrow$  signature.
  - signature depends on both sk and the message being signed
- 3. Verification algorithm: maps  $msg + sig + pk \rightarrow "yes"/"no"$ .
  - anyone who knows pk can verify correctness of an alleged signature

Digital signature scheme: defined by 3 (efficient) algorithms:

- 1. Key generation algorithm: maps seed  $r \rightarrow (pk,sk)$  pair.
  - in some cases, may generate r itself (e.g., ssh-keygen)
- 2. Signing algorithm: maps message + sk  $\rightarrow$  signature.
  - signature depends on both sk and the message being signed
- 3. Verification algorithm: maps  $msg + sig + pk \rightarrow "yes"/"no"$ .
  - anyone who knows pk can verify correctness of an alleged signature

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

note: not literally true (e.g., could reverse engineer sk by brute force)

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

note: not literally true (e.g., could reverse engineer sk by brute force)

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

note: not literally true (e.g., could reverse engineer sk by brute force)

For a formal security guarantee: need to assume...

• attacker is computationally bounded (polynomial-time)

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

note: not literally true (e.g., could reverse engineer sk by brute force)

- attacker is computationally bounded (polynomial-time)
- secret key length is sufficiently long (so brute force infeasible)

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

note: not literally true (e.g., could reverse engineer sk by brute force)

- attacker is computationally bounded (polynomial-time)
- secret key length is sufficiently long (so brute force infeasible)
- no way to forge signatures much faster than brute-forcing sk
  ideally, related to "standard" hardness assumption (like discrete log)

Ideal signature scheme: can't produce valid signatures (that you haven't already seen) unless you know the private key sk.

note: not literally true (e.g., could reverse engineer sk by brute force)

- attacker is computationally bounded (polynomial-time)
- secret key length is sufficiently long (so brute force infeasible)
- no way to forge signatures much faster than brute-forcing sk
  ideally, related to "standard" hardness assumption (like discrete log)
- non-zero (but negligible) chance an attacker gets lucky
# Defining Security for a DSS

For a formal security guarantee: need to assume...

- attacker is computationally bounded (polynomial-time)
- secret key length is sufficiently long (so brute force infeasible)
- no way to forge signatures much faster than brute-forcing sk
  - ideally, related to "standard" hardness assumption (like discrete log)
- non-zero (but negligible) chance an attacker gets lucky

(Semi-)formal DSS security statement: under suitable complexity assumptions, no randomized poly-time (in key length) algorithm with access to a bunch of signed messages can produce a valid signature for an unseen message with non-negligible probability.

# What Signatures Can and Can't Do

Issues:

- Byzantine leader could ignore read quorum requirement and make a proposal anyway.
  (→ consistency violation)
- Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)
- Byzantine validators could lie about messages received from other validators.

# What Signatures Can and Can't Do

**Issues**:

- Byzantine leader could ignore read quorum requirement and make a proposal anyway.
  (→ consistency violation)
- Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)
- Byzantine validators could lie about messages received from other validators.

Good news: signatures  $\rightarrow$  don't need to worry about issue #3 (Byzantine validators can't lie about messages sent by others).

# What Signatures Can and Can't Do

Issues:

- Byzantine leader could ignore read quorum requirement and make a proposal anyway.
  (→ consistency violation)
- Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)
- Byzantine validators could lie about messages received from other validators.

Good news: signatures  $\rightarrow$  don't need to worry about issue #3 (Byzantine validators can't lie about messages sent by others).

Bad news: even with signatures, SMR strictly harder with Byzantine faults than with crash faults.

- shared global clock (timesteps=0,1,2,...)
- known upper bound  $\Delta$  on message delays in normal conditions

- shared global clock (timesteps=0,1,2,...)
- known upper bound  $\Delta$  on message delays in normal conditions
- unknown transition time GST ("global stabilization time") from asynchrony to synchrony (i.e., end of attack/outage)

- protocol must work no matter what GST is

- shared global clock (timesteps=0,1,2,...)
- known upper bound  $\Delta$  on message delays in normal conditions
- unknown transition time GST ("global stabilization time") from asynchrony to synchrony (i.e., end of attack/outage)
  - protocol must work no matter what GST is

Recall goals:

• consistency, always (even pre-GST/"under attack")

- shared global clock (timesteps=0,1,2,...)
- known upper bound  $\Delta$  on message delays in normal conditions
- unknown transition time GST ("global stabilization time") from asynchrony to synchrony (i.e., end of attack/outage)
  - protocol must work no matter what GST is

#### Recall goals:

- consistency, always (even pre-GST/"under attack")
- liveness soon after GST (once "normal conditions" resume)
   FLP → need to give up one of consistency, liveness before GST

## Recap: Partial Synchrony + Crash Faults

Fact: crash faults + partial synchrony  $\rightarrow$  security threshold < 50%.

Suppose: validators in A don't hear from any validators in B for a long time.

• should they finalize any new txs?

#### Catch-22:



- if validators in A wait → possible liveness violation
  - if post-GST and all validators in B have crashed (will wait forever)
- if validators in A proceed → possible consistency violation
  - if pre-GST and all messages A ⇔ B have been delayed

Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

Intuition:

Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

- liveness → protocol must eventually finalize new transactions even if have heard from only n-f validators.
  - other f might well be Byzantine, could otherwise stall protocol forever

#### Post-GST Crashes or Pre-GST Delays?



Scenario #1

Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

- liveness → protocol must eventually finalize new transactions even if have heard from only n-f validators.
  - other f might well be Byzantine, could otherwise stall protocol forever

Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

- liveness → protocol must eventually finalize new transactions even if have heard from only n-f validators.
  - other f might well be Byzantine, could otherwise stall protocol forever
- ambiguity between crashes and long msg delays → might well be that f of the n-f contributing validators are Byzantine

#### Post-GST Crashes or Pre-GST Delays?





Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

- liveness → protocol must eventually finalize new transactions even if have heard from only n-f validators.
- ambiguity between crashes and long msg delays → might well be that f of the n-f contributing validators are Byzantine

Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

- liveness → protocol must eventually finalize new transactions even if have heard from only n-f validators.
- ambiguity between crashes and long msg delays → might well be that f of the n-f contributing validators are Byzantine
- 3. to avoid getting tricked, need strict majority of these n-f validators to be honest: (n-f)-f > f

Fact: Byzantine faults + partial synch  $\rightarrow$  security threshold < 33%.

i.e., no hope unless > two-thirds of validators are non-faulty

- liveness → protocol must eventually finalize new transactions even if have heard from only n-f validators.
- ambiguity between crashes and long msg delays → might well be that f of the n-f contributing validators are Byzantine
- 3. to avoid getting tricked, need strict majority of these n-f validators to be honest: (n-f)-f > f, i.e., f < n/3</p>

Starting point: Protocol C (≈ Paxos/Raft).

Starting point: Protocol C (≈ Paxos/Raft).

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

Starting point: Protocol C (≈ Paxos/Raft).

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

Issue #2: Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)

Starting point: Protocol C (≈ Paxos/Raft).

Issue #1: Byzantine leader could ignore read quorum requirement and make a proposal anyway. (→ consistency violation)

Issue #2: Byzantine leader could propose different chains to different validators, all supported by acks from Byzantine validators. (→ consistency violation)

**Issue #3:** Byzantine validators could lie about messages received from other validators.

## Key Ideas in Tendermint

Idea #1: every validator signs every message it sends.

- assume all validators know each others public keys (+ IDs + IP addrs)
- called a "public key infrastructure (PKI)" assumption

## Key Ideas in Tendermint

Idea #1: every validator signs every message it sends.

- assume all validators know each others public keys (+ IDs + IP addrs)
- called a "public key infrastructure (PKI)" assumption

**Recall:** in Protocol C, crucial that every write quorum (size > n/2) intersects every subsequent read quorum (size > n/2).

 reason: once a leader is in a position to make a proposal, it must be up-to-date on all txs already finalized by some non-faulty validator

## Key Ideas in Tendermint

Idea #1: every validator signs every message it sends.

- assume all validators know each others public keys (+ IDs + IP addrs)
- called a "public key infrastructure (PKI)" assumption

**Recall:** in Protocol C, crucial that every write quorum (size > n/2) intersects every subsequent read quorum (size > n/2).

 reason: once a leader is in a position to make a proposal, it must be up-to-date on all txs already finalized by some non-faulty validator

**Issue:** a Byzantine validator in a read quorum could deliberately submit an out-of-date chain (ignoring its past write quorums).

**Recall:** in Protocol C, crucial that every write quorum (size > n/2) intersects every subsequent read quorum (size > n/2).

 reason: once a leader is in a position to make a proposal, it must be up-to-date on all txs already finalized by some non-faulty validator

**Issue:** a Byzantine validator in a read quorum could deliberately submit an out-of-date chain (ignoring its past write quorums).

**Recall:** in Protocol C, crucial that every write quorum (size > n/2) intersects every subsequent read quorum (size > n/2).

 reason: once a leader is in a position to make a proposal, it must be up-to-date on all txs already finalized by some non-faulty validator

**Issue:** a Byzantine validator in a read quorum could deliberately submit an out-of-date chain (ignoring its past write quorums).

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2:

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary]

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary] (ii) increase all quorum sizes to > 2n/3 validators

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary] (ii) increase all quorum sizes to > 2n/3 validators

- note: given (i), (ii) does not immediately threaten liveness

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary] (ii) increase all quorum sizes to > 2n/3 validators

- note: given (i), (ii) does not immediately threaten liveness
- updated quorum intersection property:

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary] (ii) increase all quorum sizes to > 2n/3 validators

- note: given (i), (ii) does not immediately threaten liveness
- updated quorum intersection property: if S, T are quorums →
  |S|, |T| > 2n/3

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary] (ii) increase all quorum sizes to > 2n/3 validators

- note: given (i), (ii) does not immediately threaten liveness

updated quorum intersection property: if S, T are quorums →
 |S|, |T| > 2n/3 → S, T overlap in > n – n/3 – n/3 = n/3 validators

Fix: ensure that every read quorum, write quorum overlap in at least one *non-faulty* validator.

- can count on this non-faulty validator to keep leader up-to-date

Idea #2: (i) assume < n/3 validators are Byzantine. [necessary]</li>(ii) increase all quorum sizes to > 2n/3 validators

- note: given (i), (ii) does not immediately threaten liveness
- updated quorum intersection property: if S, T are quorums →
  |S|, |T| > 2n/3 → S, T overlap in > n n/3 n/3 = n/3 validators
  → S, T overlap in at least one non-faulty validator

Idea #2: (i) assume < n/3 validators are Byzantine.

(ii) increase all quorum sizes to > 2n/3 validators

- consequence: any two quorums have non-faulty validator in common
Idea #2: (i) assume < n/3 validators are Byzantine. (ii) increase all quorum sizes to > 2n/3 validators

- consequence: any two quorums have non-faulty validator in common

Bonus: can't have write quorums for two different chains in the same view (despite equivocating leader, Byzantine acks).

• will ensure that simultaneous updates must be consistent

Idea #2: (i) assume < n/3 validators are Byzantine. (ii) increase all quorum sizes to > 2n/3 validators

- consequence: any two quorums have non-faulty validator in common

Bonus: can't have write quorums for two different chains in the same view (despite equivocating leader, Byzantine acks).

- will ensure that simultaneous updates must be consistent
- reason: non-faulty validators will ack only one proposal per view
  - two write quorums → have a non-faulty validator in common → validator only acked one proposal → both WQs support same proposal

Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

 quorum certificate (QC): > 2n/3 validators attesting that a proposal by leader is up-to-date (as far as they can tell)

Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

- quorum certificate (QC): > 2n/3 validators attesting that a proposal by leader is up-to-date (as far as they can tell)
  - QCs included in blockchain as metadata



Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

- quorum certificate (QC): > 2n/3 validators attesting that a proposal by leader is up-to-date (as far as they can tell)
  - QCs included in blockchain as metadata
  - adds extra round to each view



Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

- quorum certificate (QC): > 2n/3 validators attesting that a proposal by leader is up-to-date (as far as they can tell)
  - QCs included in blockchain as metadata
  - adds extra round to each view



note: QCs don't even make sense without idea #1 (signatures)

Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

- quorum certificate (QC): > 2n/3 validators attesting that a proposal by leader is up-to-date (as far as they can tell)
  - QCs included in blockchain as metadata
  - adds extra round to each view



note: QCs don't even make sense without idea #1 (signatures)

Worry: Byzantine validators will manipulate QC formation.

Idea #3: can't trust leader to assemble a read quorum  $\rightarrow$  each validator assembles one itself before acking a proposal.

- quorum certificate (QC): > 2n/3 validators attesting that a proposal by leader is up-to-date (as far as they can tell)
  - QCs included in blockchain as metadata
  - adds extra round to each view



note: QCs don't even make sense without idea #1 (signatures)

Worry: Byzantine validators will manipulate QC formation.

 good news: idea #2 → impossible to have QCs for two different proposals in the same view (effectively, equivocation not possible)