The Economics of Proof-of-Stake Security

Mini-Lecture (30 min) – Cost of Corruption and Weak Subjectivity

Prof. Dr. Jörg Osterrieder

BSc Blockchain, Crypto Economy & NFTs

Spring 2026

Scope and Learning Objectives

Supplement to L09 (Proof of Stake). Three concepts L09 does not cover:

1. **Cost of corruption:** Why PoS attack capital is burnable (not just costly) – the capital-recoverability distinction [Analyze]
2. **Attacker's payoff inequality:** How the correlation penalty makes large-scale attacks economically irrational [Apply]
3. **Weak subjectivity:** The trust assumption PoW avoids and PoS requires for syncing nodes [Understand]

Prerequisite: L09 Proof of Stake (all frames)

This mini-lecture assumes: staking deposits, three slashing conditions, correlation penalty formula, Casper FFG finality, validator economics, stake centralization (33%/55%) – all covered in L09.

This deck focuses on the economic model behind PoS security, not the mechanism (L09's domain)

Cost-of-Attack vs Cost-of-Corruption: The Central Distinction

PoW: cost-of-attack

- Hardware (ASICs) plus energy
- After failure: hardware **resalable** at 60-70%
- Hash power is *rentable* (NiceHash)
- No permanent capital commitment required

Like a loan where you keep the collateral.

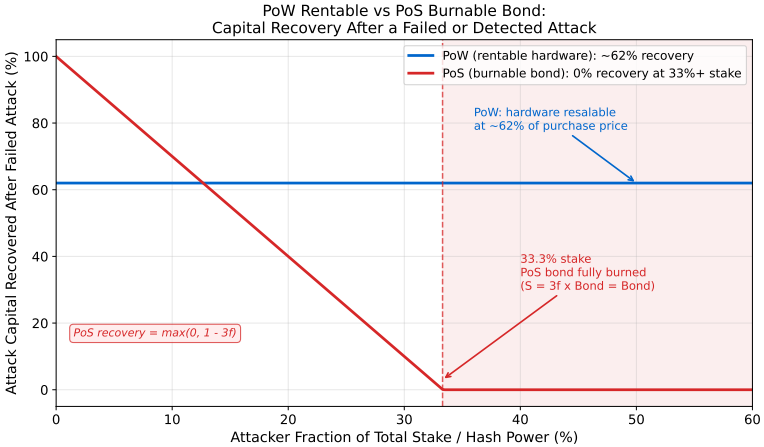
PoS: cost-of-corruption

- Bonded ETH, locked on-chain
- After detection: bond is **burned** by correlation penalty
- ETH is *non-rentable*: must own outright
- Loss on detection: up to 100% of bonded capital

Like posting collateral the protocol burns if you cheat.

"Cost of corruption" (Zamfir, 2015): the cost to cause the protocol to produce a given output (Source: Ethereum Foundation research blog, as of 2023)

PoW Rentable vs PoS Burnable Bond: Capital Recovery



PoS recovery = 0% at 33%+ stake: formula $S = 3f \times$ Bond, so at $f = 0.33$, full bond burned. PoW: 62% hardware resale (Source: Galaxy Research, as of 2023)

The Attacker's Payoff Inequality

Expected payoff:

$$\mathbb{E}[\text{payoff}] = p \cdot G - (1 - p) \cdot S - C_b$$

- p = probability of success before detection
- G = gain (double-spend value)
- S = slashed bond (burned on detection)
- C_b = opportunity cost of bonded ETH

For attack to be rational: $\mathbb{E}[\text{payoff}] > 0$

Recall from L09: correlation penalty

$$S = 3f \cdot \text{Bond}$$

where f = attacker's fraction of total stake.

- $f = 0.33$: S = full bond (100% loss)
- $f = 0.40$: $S >$ bond (capped at bond)
- $p \approx 1$ for equivocation: $G \ll S$ in practice

As stake grows, S grows faster than G: tightens non-linearly.

Detection probability $p \approx 1$ for equivocation makes S the dominant term (Source: Ethereum Proof-of-Stake Design Rationale, Buterin et al., as of 2024)

Weak Subjectivity: The Problem PoW Does Not Have

PoW: no weak subjectivity

- Block production costs energy – cannot fake history
- New node counts cumulative proof-of-work from genesis
- Honest chain is objectively longest (most cumulative work)
- Zero trust in external parties required

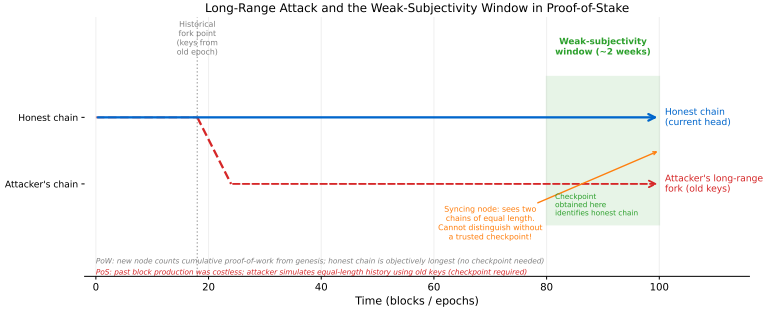
PoS: weak subjectivity problem

- Past block production was costless – attacker with old keys simulates alternate history
- Syncing node sees two chains of equal length and validity
- Cannot distinguish honest from attacker's fork by protocol rules alone
- Solution: obtain a recent checkpoint from a trusted source

Named concept (Buterin, 2014): "Weakly subjective" because a syncing node needs one external piece of social information beyond what the protocol proves from genesis.

Casper FFG prevents reorgs for nodes already online (L09 frames 22-24). Weak subjectivity is the residual assumption for syncing nodes (Source: Ethereum Foundation, as of 2023)

Long-Range Attack and the Weak-Subjectivity Window



Ethereum mainnet weak subjectivity period: ≈2 weeks (Source: Ethereum Foundation documentation, as of 2023). Nodes offline longer than this must obtain a checkpoint from a trusted source

Checkpoints and the Trust Hierarchy

What a checkpoint provides:

- A (block hash, state root) pair from a finalized block within ≈2 weeks (as of 2023)
- Anchor for chain selection: eliminates attacker's alternate history from before the checkpoint
- Sources: Ethereum Foundation portal, Consensys, Infura, peers with long uptime

Trust Hierarchy

1. **PoW:** Trust math and energy only. No external input needed.
2. **PoS without checkpoints:** Vulnerable to long-range attack.
3. **PoS with checkpoints:** One small, auditable, multi-source piece of social consensus.

Not a fatal flaw. A bounded, auditable trust assumption.

Client software (Lighthouse, Prysm) ships with an embedded checkpoint; trust is in the client release process (Source: ethstaker community documentation, as of 2023)

PoW vs PoS: Capital-Recoverability Contrast

Dimension	PoW (Bitcoin)	PoS (Ethereum)
Attack capital type	Hardware + energy OPEX	Bonded ETH, on-chain
Capital recoverability	≈60-70% (resale market)	0% on detection (burned)
Finality type	Probabilistic (Nakamoto)	Economic (Casper FFG)
Ongoing security cost	Energy (perpetual miner pay)	ETH issuance (inflation)
New entrant attack path	Rent hash power (hours)	Buy and bond ETH (weeks)

This table uses the capital-recoverability axis. L10 covers energy/TPS/finality-speed (a different axis). Both are needed for a complete economic picture.

Sources: Galaxy Research (PoW hardware resale, as of 2023); Ethereum Proof-of-Stake Design Rationale (PoS recoverability, as of 2024); Ethereum Foundation (issuance model)

Key Takeaways

1. **PoS attack capital is burnable, not rentable.** The correlation penalty converts the attacker's bond into a non-recoverable loss on detection. At 33%+ stake, the full bond is destroyed. PoW hardware recovers 62% on resale.

2. **Weak subjectivity is a bounded trust assumption, not a fatal flaw.** Syncing nodes need one external checkpoint (a recent finalized block hash). It is small, auditable, and available from multiple independent sources.

3. **The PoW/PoS comparison needs two axes.** L10 covers energy/TPS/finality speed. This supplement adds the capital-recoverability axis: recoverable vs. burnable attack capital, probabilistic vs. economic finality, hours vs. weeks to assemble attack capital.

Further reading: Buterin (2014) "Proof of Stake: How I Learned to Love Weak Subjectivity"; Ethereum Foundation Proof-of-Stake Design Rationale; Zamfir (2015) "Casper the Friendly Ghost"

Quiz Questions 1–5

Q1. What distinguishes "cost-of-attack" from "cost-of-corruption" in blockchain security?
A) Cost-of-attack measures hardware; cost-of-corruption measures energy B) Cost-of-attack is execution cost; cost-of-corruption is the capital permanently destroyed on detection
C) They are equivalent terms D) Cost-of-attack applies to PoS; cost-of-corruption applies to PoW

Quiz Questions 1–5

Q1. What distinguishes "cost-of-attack" from "cost-of-corruption" in blockchain security?
A) Cost-of-attack measures hardware; cost-of-corruption measures energy B) Cost-of-attack is execution cost; cost-of-corruption is the capital permanently destroyed on detection
C) They are equivalent terms D) Cost-of-attack applies to PoS; cost-of-corruption applies to PoW **Answer: B** – Zamfir (2015) introduced cost-of-corruption as the amount permanently destroyed when an attack is detected, distinct from the upfront execution cost.

Q2. An attacker holds 35% of staked ETH. Using S = 3f x Bond, what fraction of the bond is burned?
A) 35% B) 105% (capped at 100%) C) 0% (penalty only applies to individual validators) D) 65%

Quiz Questions 1–5

Q1. What distinguishes "cost-of-attack" from "cost-of-corruption" in blockchain security?
A) Cost-of-attack measures hardware; cost-of-corruption measures energy B) Cost-of-attack is execution cost; cost-of-corruption is the capital permanently destroyed on detection
C) They are equivalent terms D) Cost-of-attack applies to PoS; cost-of-corruption applies to PoW **Answer: B** – Zamfir (2015) introduced cost-of-corruption as the amount permanently destroyed when an attack is detected, distinct from the upfront execution cost.

Q2. An attacker holds 35% of staked ETH. Using S = 3f x Bond, what fraction of the bond is burned?
A) 35% B) 105% (capped at 100%) C) 0% (penalty only applies to individual validators) D) 65% **Answer: B** – $3 \times 0.35 = 1.05$: exceeds the bond, so the full bond is burned (capped at 100%).

Q3. Why is PoW attack capital described as "rentable" but PoS attack capital is "non-rentable"?
A) PoW miners pay rent; PoS validators own stake B) PoW pools rent hardware C) PoS validators earn rental income D) PoW hash power can be leased via NiceHash without owning hardware; PoS attack requires owning and bonding ETH outright

Quiz Questions 1–5

Q1. What distinguishes "cost-of-attack" from "cost-of-corruption" in blockchain security?
A) Cost-of-attack measures hardware; cost-of-corruption measures energy B) Cost-of-attack is execution cost; cost-of-corruption is the capital permanently destroyed on detection
C) They are equivalent terms D) Cost-of-attack applies to PoS; cost-of-corruption applies to PoW **Answer: B** – Zamfir (2015) introduced cost-of-corruption as the amount permanently destroyed when an attack is detected, distinct from the upfront execution cost.

Q2. An attacker holds 35% of staked ETH. Using S = 3f x Bond, what fraction of the bond is burned?
A) 35% B) 105% (capped at 100%) C) 0% (penalty only applies to individual validators) D) 65% **Answer: B** – $3 \times 0.35 = 1.05$: exceeds the bond, so the full bond is burned (capped at 100%).

Q3. Why is PoW attack capital described as "rentable" but PoS attack capital is "non-rentable"?
A) PoW miners pay rent; PoS validators own stake B) PoW pools rent hardware C) PoS validators earn rental income D) PoW hash power can be leased via NiceHash without owning hardware; PoS attack requires owning and bonding ETH outright **Answer: D** – NiceHash and similar platforms allow temporary hash power rental. No equivalent market exists for PoS attack bonds.

Q4. What is the approximate PoW hardware capital recovery after a failed attack?
A) Approximately 60-70% (hardware resale market) B) 0% C) 100% D) Exactly 33%

Quiz Questions 1–5

Q1. What distinguishes "cost-of-attack" from "cost-of-corruption" in blockchain security?
A) Cost-of-attack measures hardware; cost-of-corruption measures energy B) Cost-of-attack is execution cost; cost-of-corruption is the capital permanently destroyed on detection
C) They are equivalent terms D) Cost-of-attack applies to PoS; cost-of-corruption applies to PoW **Answer: B** – Zamfir (2015) introduced cost-of-corruption as the amount permanently destroyed when an attack is detected, distinct from the upfront execution cost.

Q2. An attacker holds 35% of staked ETH. Using S = 3f x Bond, what fraction of the bond is burned?
A) 35% B) 105% (capped at 100%) C) 0% (penalty only applies to individual validators) D) 65% **Answer: B** – $3 \times 0.35 = 1.05$: exceeds the bond, so the full bond is burned (capped at 100%).

Q3. Why is PoW attack capital described as "rentable" but PoS attack capital is "non-rentable"?
A) PoW miners pay rent; PoS validators own stake B) PoW pools rent hardware C) PoS validators earn rental income D) PoW hash power can be leased via NiceHash without owning hardware; PoS attack requires owning and bonding ETH outright **Answer: D** – NiceHash and similar platforms allow temporary hash power rental. No equivalent market exists for PoS attack bonds.

Q4. What is the approximate PoW hardware capital recovery after a failed attack?
A) Approximately 60-70% (hardware resale market) B) 0% C) 100% D) Exactly 33% **Answer: A** – ASICs retain market value (depreciation aside) and can be resold.

Q5. In the payoff formula, what does S represent?
A) Success probability B) Gain from a successful attack C) Slashed bond burned on detection D) Validator staking reward

Quiz Questions 1–5

Q1. What distinguishes "cost-of-attack" from "cost-of-corruption" in blockchain security?
A) Cost-of-attack measures hardware; cost-of-corruption measures energy B) Cost-of-attack is execution cost; cost-of-corruption is the capital permanently destroyed on detection
C) They are equivalent terms D) Cost-of-attack applies to PoS; cost-of-corruption applies to PoW **Answer: B** – Zamfir (2015) introduced cost-of-corruption as the amount permanently destroyed when an attack is detected, distinct from the upfront execution cost.

Q2. An attacker holds 35% of staked ETH. Using S = 3f x Bond, what fraction of the bond is burned?
A) 35% B) 105% (capped at 100%) C) 0% (penalty only applies to individual validators) D) 65% **Answer: B** – $3 \times 0.35 = 1.05$: exceeds the bond, so the full bond is burned (capped at 100%).

Q3. Why is PoW attack capital described as "rentable" but PoS attack capital is "non-rentable"?
A) PoW miners pay rent; PoS validators own stake B) PoW pools rent hardware C) PoS validators earn rental income D) PoW hash power can be leased via NiceHash without owning hardware; PoS attack requires owning and bonding ETH outright **Answer: D** – NiceHash and similar platforms allow temporary hash power rental. No equivalent market exists for PoS attack bonds.

Q4. What is the approximate PoW hardware capital recovery after a failed attack?
A) Approximately 60-70% (hardware resale market) B) 0% C) 100% D) Exactly 33% **Answer: A** – ASICs retain market value (depreciation aside) and can be resold.

Q5. In the payoff formula, what does S represent?
A) Success probability B) Gain from a successful attack C) Slashed bond burned on detection D) Validator staking reward **Answer: C** – S is the slashed bond destroyed by the correlation penalty upon detection.

Quiz Questions 6–10

Q6. What is “weak subjectivity” in Proof-of-Stake?
A) Validators choose transactions subjectively B) PoS networks favor validators with moral preferences
C) Validators must subjectively confirm blocks D) The need for a syncing node to obtain a trusted recent checkpoint because it cannot distinguish the honest chain from a long-range fork by protocol rules alone

Quiz Questions 6–10

Q6. What is "weak subjectivity" in Proof-of-Stake?
A) Validators choose transactions subjectively B) PoS networks favor validators with moral preferences
C) Validators must subjectively confirm blocks D) The need for a syncing node to obtain a trusted recent checkpoint because it cannot distinguish the honest chain from a long-range fork by protocol rules alone **Answer: D** – Named by Buterin (2014): a syncing node needs one piece of external social information.

Q7. Why does PoW avoid the weak subjectivity problem?
A) PoW uses checkpoints from mining pools B) PoW block production requires real energy; an attacker cannot simulate an equal-length alternate history without burning the same electricity as the honest chain C) PoW validators are randomly selected D) PoW chains are shorter

Quiz Questions 6–10

Q6. What is "weak subjectivity" in Proof-of-Stake?
A) Validators choose transactions subjectively B) PoS networks favor validators with moral preferences
C) Validators must subjectively confirm blocks D) The need for a syncing node to obtain a trusted recent checkpoint because it cannot distinguish the honest chain from a long-range fork by protocol rules alone **Answer: D** – Named by Buterin (2014): a syncing node needs one piece of external social information.

Q7. Why does PoW avoid the weak subjectivity problem?
A) PoW uses checkpoints from mining pools B) PoW block production requires real energy; an attacker cannot simulate an equal-length alternate history without burning the same electricity as the honest chain C) PoW validators are randomly selected D) PoW chains are shorter **Answer: B** – Energy expenditure makes history unforgeable in PoW. In PoS, past block production was costless.

Q8. What is the approximate Ethereum weak subjectivity period (as of 2023)?
A) 10 minutes B) 6 hours C) Approximately 2 weeks D) 1 year

Quiz Questions 6–10

Q6. What is "weak subjectivity" in Proof-of-Stake?
A) Validators choose transactions subjectively B) PoS networks favor validators with moral preferences
C) Validators must subjectively confirm blocks D) The need for a syncing node to obtain a trusted recent checkpoint because it cannot distinguish the honest chain from a long-range fork by protocol rules alone **Answer: D** – Named by Buterin (2014): a syncing node needs one piece of external social information.

Q7. Why does PoW avoid the weak subjectivity problem?
A) PoW uses checkpoints from mining pools B) PoW block production requires real energy; an attacker cannot simulate an equal-length alternate history without burning the same electricity as the honest chain C) PoW validators are randomly selected D) PoW chains are shorter **Answer: B** – Energy expenditure makes history unforgeable in PoW. In PoS, past block production was costless.

Q8. What is the approximate Ethereum weak subjectivity period (as of 2023)?
A) 10 minutes B) 6 hours C) Approximately 2 weeks D) 1 year **Answer: C** – The Ethereum Foundation documentation specifies approximately 2 weeks as the weak subjectivity period (Source: Ethereum Foundation, as of 2023).

Q9. An attacker uses keys from a validator set that withdrew 3 months ago. What can they attempt against a syncing PoS node?
A) 51% attack with current hash power B) Long-range attack: simulate an alternate chain history from 3 months ago that appears equally valid to a naive syncing node C) Nothing-at-stake attack D) Slashing attack

Quiz Questions 6–10

Q6. What is "weak subjectivity" in Proof-of-Stake?
A) Validators choose transactions subjectively B) PoS networks favor validators with moral preferences
C) Validators must subjectively confirm blocks D) The need for a syncing node to obtain a trusted recent checkpoint because it cannot distinguish the honest chain from a long-range fork by protocol rules alone **Answer: D** – Named by Buterin (2014): a syncing node needs one piece of external social information.

Q7. Why does PoW avoid the weak subjectivity problem?
A) PoW uses checkpoints from mining pools B) PoW block production requires real energy; an attacker cannot simulate an equal-length alternate history without burning the same electricity as the honest chain C) PoW validators are randomly selected D) PoW chains are shorter **Answer: B** – Energy expenditure makes history unforgeable in PoW. In PoS, past block production was costless.

Q8. What is the approximate Ethereum weak subjectivity period (as of 2023)?
A) 10 minutes B) 6 hours C) Approximately 2 weeks D) 1 year **Answer: C** – The Ethereum Foundation documentation specifies approximately 2 weeks as the weak subjectivity period (Source: Ethereum Foundation, as of 2023).

Q9. An attacker uses keys from a validator set that withdrew 3 months ago. What can they attempt against a syncing PoS node?
A) 51% attack with current hash power B) Long-range attack: simulate an alternate chain history from 3 months ago that appears equally valid to a naive syncing node C) Nothing-at-stake attack D) Slashing attack **Answer: B** – Old keys have no current stake but can simulate costless PoS history. Checkpoints defeat this.

Q10. Which dimension does this supplement add to L10's PoW/PoS comparison?
A) Transaction throughput (TPS) B) Energy consumption C) Block time D) Capital recoverability after a failed or detected attack (recoverable vs. burnable)

Quiz Questions 6–10

Q6. What is "weak subjectivity" in Proof-of-Stake?
A) Validators choose transactions subjectively B) PoS networks favor validators with moral preferences
C) Validators must subjectively confirm blocks D) The need for a syncing node to obtain a trusted recent checkpoint because it cannot distinguish the honest chain from a long-range fork by protocol rules alone **Answer: D** – Named by Buterin (2014): a syncing node needs one piece of external social information.

Q7. Why does PoW avoid the weak subjectivity problem?
A) PoW uses checkpoints from mining pools B) PoW block production requires real energy; an attacker cannot simulate an equal-length alternate history without burning the same electricity as the honest chain C) PoW validators are randomly selected D) PoW chains are shorter **Answer: B** – Energy expenditure makes history unforgeable in PoW. In PoS, past block production was costless.

Q8. What is the approximate Ethereum weak subjectivity period (as of 2023)?
A) 10 minutes B) 6 hours C) Approximately 2 weeks D) 1 year **Answer: C** – The Ethereum Foundation documentation specifies approximately 2 weeks as the weak subjectivity period (Source: Ethereum Foundation, as of 2023).

Q9. An attacker uses keys from a validator set that withdrew 3 months ago. What can they attempt against a syncing PoS node?
A) 51% attack with current hash power B) Long-range attack: simulate an alternate chain history from 3 months ago that appears equally valid to a naive syncing node C) Nothing-at-stake attack D) Slashing attack **Answer: B** – Old keys have no current stake but can simulate costless PoS history. Checkpoints defeat this.

Q10. Which dimension does this supplement add to L10's PoW/PoS comparison?
A) Transaction throughput (TPS) B) Energy consumption C) Block time D) Capital recoverability after a failed or detected attack (recoverable vs. burnable) **Answer: D** – L10 covers energy/TPS/finality speed. This deck adds the capital-recoverability / finality-type / attack-assembly axis.