The Economics of Proof-of-Stake Security

Cost of Corruption and Weak Subjectivity: A Supplement to L09

Prof. Dr. Jörg Osterrieder

BSc Blockchain, Crypto Economy & NFTs

Spring 2026

What L09 Taught You (and What This Deck Adds)

Prerequisite: L09 already establishes

- Staking deposits and validator lifecycle (L09 frames 2-5)
- Nothing-at-stake problem and resolution (L09 frames 10-12)
- Three slashing conditions: double-vote, surround-vote, inactivity (L09 frames 19-21)
- Correlation penalty formula (L09 frame 21)
- Casper FFG finality: justified then finalized (L09 frames 22-24)
- Validator economics and APR curve (L09 frames 14-16)
- Stake centralization risks: 33%/55% thresholds (L09 frames 26-27)

What This Deck Adds

1. **Cost of corruption:** Why PoS attack capital is burnable, not just costly
2. **Weak subjectivity:** The trust assumption PoW avoids and PoS requires
3. **Security-economics contrast:** Capital recoverability as the defining axis

Recall from L09: the correlation penalty means an attacker loses stake non-linearly. This deck asks: what does that imply for the attacker's expected payoff?

This deck assumes L09 as prerequisite. Frames 5-7 build on the correlation penalty; Frame 8 builds on Casper FFG finality

Learning Objectives

By the end of this supplement, you will be able to:

1. Distinguish *cost-of-attack* from *cost-of-corruption* and explain why PoS attack capital is burnable while PoW attack capital is rentable and recoverable [Analyze]
2. Derive the attacker's expected payoff inequality and show why slashing is the dominant term at high stake fractions [Apply]
3. Define *weak subjectivity* and explain why a syncing node cannot distinguish the honest chain from a long-range attacker's history without a trusted checkpoint [Understand]
4. Compare PoW and PoS on the capital-recoverability axis (not the energy/TPS axis covered by L10) [Evaluate]
5. Assess which trust assumptions a PoS network requires that a PoW network does not [Evaluate]

Bloom's levels: Understand, Apply, Analyze, Evaluate

Prerequisites: L09 Proof of Stake (all frames) and L10 Consensus Comparison

The Economic Security Gap: From Mechanism to Model

What L09 explains (mechanism):

- Validators bond ETH, behave honestly, earn rewards
- Slashing destroys the bond if you cheat
- Attack cost $>$ potential gain implies honest behavior is Nash equilibrium

What L09 does not explain (model):

- *Why* is PoS different from PoW, where attack cost is also high?
- What happens to a miner's hardware after a failed PoW attack?
- Why does PoS require a social trust assumption that PoW avoids?

The Three Questions This Deck Answers

1. What is the **structure** of attack capital? (Rentable? Burnable?)
2. What is the **expected payoff** of an attack after slashing?
3. What new **trust assumptions** does PoS introduce that PoW avoids?

Mechanism describes the rules. Model predicts behavior under incentives.

Frames 5-7 answer Q1 and Q2 (capital structure and payoff); Frames 8-10 answer Q3 (trust assumptions); Frames 11-12 synthesize both

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

Cost-of-attack (PoW):

- Must acquire 51% of hash power
- Capital form: hardware (ASICs) plus energy
- After a failed attack: hardware is **resalable** (market price minus depreciation)
- Capital is *rentable*: hash power can be leased via NiceHash without permanent commitment
- Loss on failure: energy cost plus opportunity cost; hardware recovers 50-70%

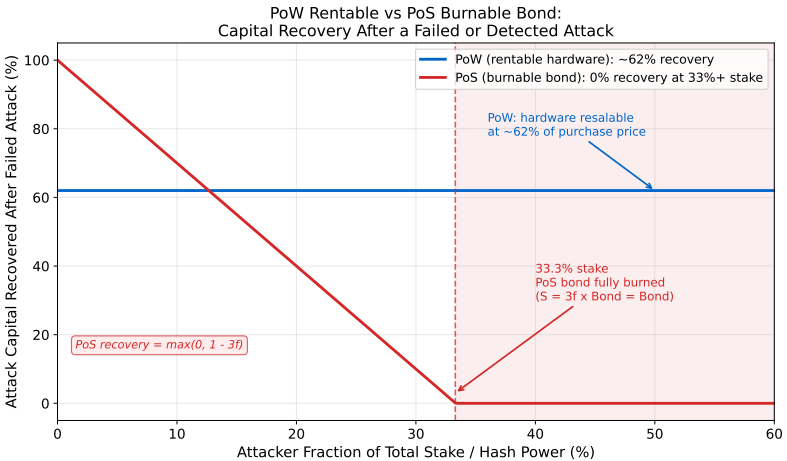
Cost-of-corruption (PoS):

- Must bond 33%+ of staked ETH to threaten finality
- Capital form: ETH bond, locked on-chain
- After a detected attack: bond is **burned** by the correlation penalty
- Capital is *non-rentable*: must own the ETH outright (no liquid attack lease market)
- Loss on detection: up to 100% of bonded capital

Key insight: PoW attack is like a loan where you keep the collateral. PoS attack is like posting collateral that the protocol burns if you cheat.

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

PoW Rentable vs PoS Burnable Bond: Attacker Capital Recovery



PoW recovery floor $\approx$ 60-70% (hardware resale market); PoS recovery = 0% at $\geq$33% stake due to non-linear correlation penalty (Source: Galaxy Research hardware analysis) (as of 2023)

The Attacker's Payoff Inequality: Slashing as the Dominant Term

Expected payoff of a PoS attack:

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

- p = probability attack succeeds before detection
- G = gain if successful (double-spend or censorship value)
- S = slashed bond (burned on detection)
- C_b = borrow/opportunity cost of bonded ETH

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

Why slashing dominates:

Recall from L09: correlation penalty

$$= \text{Bond} \times 3 \times \frac{\text{slashed}}{\text{total}}$$

If attacker holds fraction f of total stake, all colluders slashed together:

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

- At $f = 0.33$: $S = \text{Bond}$ (full burn)
- At $f = 0.40$: $S > \text{Bond}$ (capped at bond)

As the attacker accumulates more stake, S grows faster than G: the inequality tightens non-linearly.

Detection probability $p \approx 1$ for equivocation offences, making $G \cdot p \ll S \cdot (1 - p)$ in practice (Source: Ethereum Proof-of-Stake Design Rationale, Buterin et al.) (as of 2024)

Weak Subjectivity: The Problem PoW Does Not Have

PoW has no weak subjectivity problem:

- Block production requires energy; you cannot fake a long history without burning real electricity
- A new node verifies by counting cumulative proof-of-work from genesis
- The honest chain is objectively the one with most cumulative work
- No trust in external parties required for chain selection

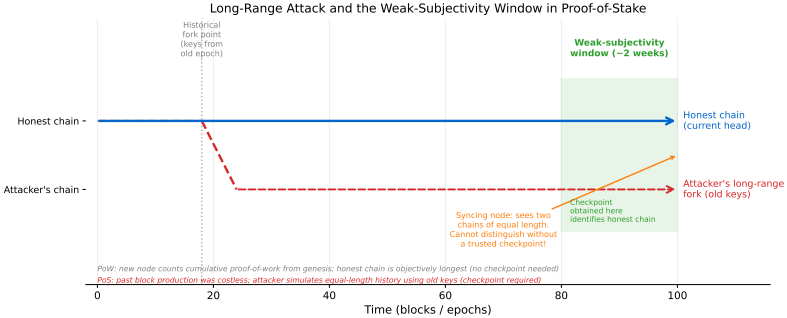
PoS has a weak subjectivity problem:

- Past block production was costless: an attacker with old keys can simulate an alternate history back to any checkpoint
- A syncing node sees two chains of equal length and equal apparent validity
- It cannot distinguish the honest chain from the attacker's long-range fork by protocol rules alone
- Solution: obtain a recent *weak subjectivity checkpoint* from a trusted source

Named concept (Buterin, 2014): "Weakly subjective" because a new client needs one piece of external social information beyond what the protocol can prove from genesis.

Casper FFG prevents long-range reorgs for nodes already online (L09 frames 22-24). Weak subjectivity is the residual trust 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 period must obtain a checkpoint from a trusted source before syncing

Weak Subjectivity Checkpoints: What Trust Do They Require?

What a checkpoint provides:

- A (block hash, state root) pair from a finalized block within the last ≈2 weeks (as of 2023)
- Allows a syncing node to start chain selection from a known honest anchor
- Eliminates the attacker's ability to present an alternate history from before the checkpoint

Checkpoint sources (Ethereum):

- Ethereum Foundation checkpoint portal
- Consensys, Infura, known validator operators
- Peers with verified long uptime

Trust Hierarchy

1. **PoW (Bitcoin):** Trust only math and energy. No external input needed to sync.
2. **PoS without checkpoints:** Vulnerable to long-range attack from any historical key holder.
3. **PoS with checkpoints (Ethereum):** Trust is weakly subjective: requires one piece of social consensus (the checkpoint), but it is small, auditable, and multi-source.

The checkpoint requirement is not a fatal flaw. It is a bounded, auditable trust assumption.

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

PoW vs PoS Security-Economics: The Capital-Recoverability Axis

Proof-of-Work:

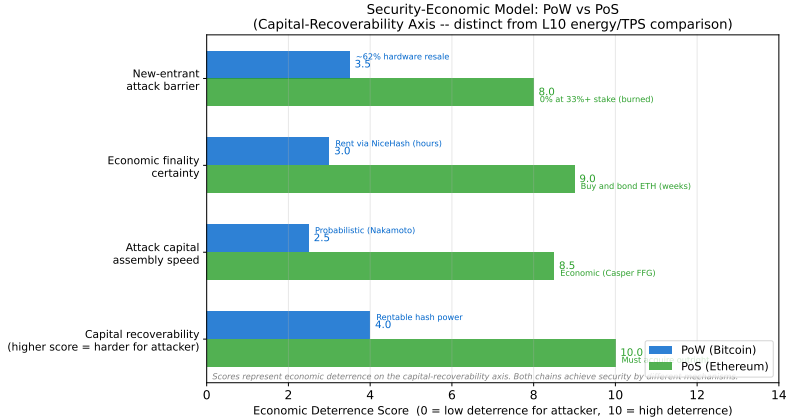
- Attack capital: hardware (ASICs) plus energy OPEX
- Recoverability: ≈60-70% (hardware resale; Source: Galaxy Research, as of 2023)
- Finality type: probabilistic (Nakamoto), reversal risk never reaches zero
- Ongoing security cost: energy, miners paid perpetually
- New entrant attack path: rent hash power in hours (NiceHash)

Proof-of-Stake:

- Attack capital: bonded ETH, on-chain, non-rentable
- Recoverability: 0% on detection, correlation penalty burns bond (Source: Ethereum Proof-of-Stake Design Rationale, as of 2024)
- Finality type: economic (Casper FFG), reversal costs >33% of staked ETH
- Ongoing security cost: ETH issuance (inflation to pay validators)
- New entrant attack path: buy and bond ETH (slow, visible, weeks)

This frame uses the capital-recoverability axis. L10 covers energy/TPS/finality-speed separately. Both axes are needed for a complete economic picture

Security-Economic Model: PoW vs PoS Comparison



Capital recoverability and finality type are the axes where PoS differs fundamentally from PoW (Source: Ethereum Foundation; Rated Network; as of 2024)

Key Takeaways: The Economic Case for PoS Security

Three core takeaways:

1. **PoS attack capital is burnable, not rentable.** The correlation penalty converts a validator's bond into a non-recoverable loss on detection, making large-scale attacks economically irrational at 33%+ stake.
2. **Weak subjectivity is a bounded trust assumption.** Syncing nodes need one external checkpoint (a recent finalized block hash) to defeat long-range attacks. This is a known, auditable, multi-source requirement.
3. **The security-economics comparison requires two axes:** energy/TPS (L10's domain) and capital recoverability plus finality type (this deck's domain).

Open Questions

1. If ETH issuance is the ongoing cost of PoS security, what happens when burns outpace issuance? (Post-EIP-1559 dynamics, as of 2024)
2. Can a liquid staking protocol (Lido, Rocket Pool) create a rental market for PoS attack capital, narrowing the recoverability gap?
3. How does the weak subjectivity period change if the validator set rotates faster?

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"