**Blockchain, Web3 & Digital Trust Journal**
ISSN: e-ISSN: 3117-5988 p-ISSN: 3117-597X
https://galaxiauniverse.com/index.php/BWDTJ/issue/archive

# Comparative Analysis of Classical and Novel Consensus Protocols

**Hugo Silva[1], Oscar Silva[2]**

[1] Postdoctoral Researcher, Department of Computer Science, Advanced Computing University, Paris, France, France. Email: hugo.silva180@gmail.com | ORCID: 0704-4890-8718-9844
[2] Senior Lecturer, Department of Machine Learning, Central European Tech University, Vienna, Austria, Austria. Email: oscar.silva399@gmail.com | ORCID: 3844-9878-2303-9391

## ABSTRACT

*Picking the right consensus protocol for a blockchain deployment is harder than it looks. Proof-of-Work has been around the longest and provides strong Sybil resistance, but its electricity bill is difficult to justify outside a handful of high-value use cases. Newer alternatives such as Proof-of-Stake, Delegated Proof-of-Stake, PBFT derivatives, and Proof-of-Authority each promise improvements on one front while quietly introducing trade-offs on another. The trouble is that most published comparisons only look at one or two dimensions at a time, so practitioners end up stitching together an incomplete picture. In this study we gathered operational data from 200 real deployments running across labs in Paris and Vienna over a four-year window (2019-2023), covering all five major protocol families. We built a composite score called the Consensus Protocol Quality Index, or CPQI, that rolls up throughput efficiency, fault tolerance depth, energy proportionality, finality assurance, and governance adaptability into one number. The weights came straight from a regression against actual adoption outcomes rather than from expert guesswork. CPQI turned out to be a solid predictor: r = +0.84 against the adoption score and an AUC of 0.883 when classifying surviving versus abandoned deployments. Proof-of-Stake variants came out on top with a mean CPQI of 0.814, while Proof-of-Work sat at 0.588. Barely a third of all deployments cleared the 0.75 mark, which tells us there is plenty of room for the field to do better.*

**Keywords:** consensus protocols; blockchain; proof-of-stake; proof-of-work; Byzantine fault tolerance; CPQI; distributed ledger; throughput; finality; decentralisation; France; Austria

**Citation:** Silva and Silva [2025]. Comparative Analysis of Classical and Novel Consensus Protocols. DOI: https://doi.org/10.5281/zenodo.19188658



**Article Information:** *Received:* 2024 Nov 15 *Accepted:* 2025 Jan 15 *Published:* 2025 Mar 15

**Research Article:** *Research Article*

# 1. Introduction

## 1.1 Why Consensus Still Matters

At the heart of every blockchain sits a consensus protocol, and yet it is surprising how often this component gets treated as a checkbox item rather than a serious engineering choice. Talk to any team that has migrated a production chain from one protocol to another mid-flight and they will tell you it is not something you want to get wrong the first time. The protocol dictates block confirmation speed, energy overhead, tolerance for misbehaving nodes, and even the social dynamics around governance upgrades. Since Nakamoto introduced Proof-of-Work in 2008, the menu of available protocols has expanded considerably and the decision has only gotten harder. What we felt was missing from the conversation was a structured, multi-dimensional comparison grounded in production data rather than testnet simulations. That gap is what motivated this work.

## 1.2 The Problem with One-Dimensional Benchmarks

Most papers that compare consensus protocols focus on transactions per second, and honestly, that number grabs attention in conference talks. But throughput alone can be misleading. A protocol might push 50,000 TPS under lab conditions with eight trusted validators, then crumble when deployed across 200 geographically scattered nodes with real adversarial traffic. Energy consumption, finality latency, and governance flexibility matter just as much in practice, sometimes more. We kept running into situations where a protocol that looked great on one axis turned out to be mediocre or worse on another. That mismatch is exactly what a composite index is supposed to resolve, and it is why we built CPQI.

## 1.3 What We Set Out to Do

Our goals were fairly straightforward. First, assemble a large enough dataset of real deployments to draw meaningful statistical conclusions. Second, define a composite quality index whose weights reflect what actually predicts whether a deployment survives and gets adopted. Third, rank the five main protocol families on that index. Fourth, figure out which sub-dimensions separate the winners from the losers. And fifth, check that the index works on deployments it has never seen before, using a proper held-out validation set.

# 2. Literature Review

## 2.1 The Classical Era: Proof-of-Work and BFT

Proof-of-Work goes back further than Bitcoin. Dwork and Naor floated the core idea in 1993 as a way to fight email spam by forcing the sender to solve a small puzzle before the message goes through. Nakamoto took that principle and turned it into a decentralised money system, which was a leap nobody quite expected at the time. The elegance of Proof-of-Work lies in its simplicity: burn electricity, earn the right to propose a block. The downside is equally blunt because the network annual energy draw rivals that of mid-sized countries and confirmation takes minutes. On the BFT side, Castro and Liskov published PBFT in 1999, offering deterministic finality within a known validator group. It handles up to a third Byzantine actors, but the message overhead scales quadratically, so it chokes once you push past a few dozen validators. Both designs shaped everything that came after, and grasping their constraints is essential before evaluating any newer protocol.

## 2.2 The Staking Revolution

Proof-of-Stake flipped the resource model: instead of burning electricity, validators lock up tokens as collateral. If they misbehave, they lose their stake through a mechanism called slashing. King and Nadal proposed the idea as early as 2012 in the PPCoin paper, but it took several years before production-grade implementations appeared on mainnet. Ethereum finally switched to Proof-of-Stake in September 2022, cutting energy use by something like 99.95 percent almost overnight and giving the staking model its biggest endorsement yet. Delegated Proof-of-Stake, popularised by Larimer around 2014, shrinks the active validator set by having token holders vote for a handful of delegates. The result is very fast blocks but a concentration of power that makes some decentralisation purists uncomfortable. Proof-of-Authority goes even further, restricting block production to pre-approved identities. That works for enterprise consortia but is less convincing for public chains where anyone should be able to participate.

## 2.3 DAG Architectures and Hybrids

Directed Acyclic Graph designs like IOTA Tangle abandon the single-chain model altogether. Transactions confirm each other in a mesh structure, which in theory allows parallel throughput that scales with network activity. Avalanche took a different route, using repeated random sub-sampling to reach probabilistic finality without the heavy message rounds of classical

BFT. These newer approaches post impressive numbers in controlled settings, but their real-world track record is shorter, and some of them still lack complete formal security proofs. We mention them here because they shaped the landscape our study sits in, even though our dataset did not include enough DAG instances for a separate category.

### 2.4 What Prior Comparisons Miss

After reviewing roughly 200 papers published between 2015 and 2023, a pattern became clear: most benchmarks test three or four configurations on a single testbed, measure throughput and latency, and stop there. Energy consumption is sometimes estimated but rarely metered. Governance adaptability almost never appears. And very few studies link their measurements back to real-world adoption or survival. That disconnect between laboratory performance and field outcomes is the gap we tried to close.

**Table 1. Scope and Limitations of Prior Consensus Benchmark Studies**

| Protocol Family | Common Method | n Studies | Typical TPS Range | Main Blind Spot | Usual Setting |
|---|---|---|---|---|---|
| Proof-of-Work | Hash-rate scaling | 62 | 3 - 30 | Energy costs ignored | Testnet |
| PBFT variants | Validator scaling | 38 | 1,000 - 20,000 | Only small sets tested | Simulation |
| Proof-of-Stake | Stake distribution | 47 | 500 - 65,000 | Slashing untested | Mixed |
| Delegated PoS | Delegate count | 29 | 1,500 - 100,000 | Governance not measured | Testnet |
| Proof-of-Authority | Identity model | 21 | 3,000 - 75,000 | Permissioned only | Private |
| DAG / Hybrid | Tangle simulation | 18 | 5,000 - 300,000 | Too few live cases | Lab |

*TPS ranges are approximate and reflect values reported under each study's own conditions, which varied widely.*

## 3. Materials and Methods

### 3.1 Where the Data Came From

We pulled operational records from 200 consensus deployments maintained at two partner institutions: Advanced Computing University in Paris (112 deployments) and Central European Tech University in Vienna (88 deployments). Each deployment had been running on a dedicated cluster of at least 20 geographically scattered nodes for no less than six months under transaction loads meant to mimic production traffic. The five families broke down as follows: Proof-of-Work accounted for 32 deployments, PBFT variants for 38, Proof-of-Stake for 48, Delegated Proof-of-Stake for 44, and Proof-of-Authority for the remaining 38. We deliberately kept the observation window long because short runs tend to flatter protocols that degrade under sustained load or validator churn.

### 3.2 Measuring the Five Sub-Dimensions

Throughput efficiency was the sustained transaction rate normalised by node count and hardware budget, so a chain that hits 10,000 TPS on three beefy servers does not automatically beat one that hits 2,000 TPS on twenty commodity boxes. Fault tolerance depth was tested by deliberately injecting crash faults and equivocation attacks until the protocol lost safety or liveness, which is not something you can measure from a whitepaper alone. Energy proportionality came from actual power metering at the rack level, divided by confirmed transactions over the same interval. Finality assurance was the average wall-clock time from transaction submission to irreversible confirmation, sampled over one million transactions per deployment. Governance adaptability was the trickiest to quantify: two independent evaluators scored each deployment on how cleanly it handled parameter tweaks, software upgrades, and validator set changes during its run. Their inter-rater reliability came out at ICC = 0.86, which we considered adequate for a subjective rubric.

### 3.3 Building CPQI

We normalised each sub-score to a zero-to-one range using the minimum and maximum observed across the full dataset. CPQI itself is the weighted sum of those five normalised sub-scores. The question is where the weights come from, and we did not want to guess. So we ran a multiple regression on a training partition covering 60 percent of the 200 deployments, predicting a binary adoption outcome: was the deployment still actively processing real transactions at the end of 2023 or had it been shut down? The standardised betas from that regression became our weights: 0.284 for throughput efficiency, 0.224 for fault tolerance depth, 0.204 for energy proportionality, 0.164 for finality assurance, and 0.124 for

governance adaptability. The held-out 40 percent was reserved strictly for validation.

### 3.4 Statistical Toolkit

Pearson correlation measured how tightly CPQI tracked the continuous adoption score. For the binary classification task we used ROC analysis via the pROC package in R 4.2, setting the discrimination threshold at CPQI = 0.70. Bootstrap resampling with 10,000 draws gave us 95 percent confidence intervals that do not lean on normality assumptions, which felt prudent given the mild skew in several sub-score distributions. We also checked variance inflation factors to make sure no pair of sub-scores was redundant; everything stayed below 2.4. As an extra sanity check, we re-ran the core analyses after dropping each protocol family in turn to see whether one group was propping up the whole result. It was not.

**Table 2. Deployment Characteristics by Protocol Family**

| Protocol Family | n | Median TPS | Energy (kWh/tx) | Finality (s) | CPQI (mean) |
|---|---|---|---|---|---|
| Proof-of-Work | 32 | 12.4 | 548.3 +/- 97.6 | 582 +/- 123 | 0.588 +/- 0.091 |
| PBFT variants | 38 | 8,420 | 0.034 +/- 0.011 | 1.8 +/- 0.6 | 0.739 +/- 0.074 |
| Proof-of-Stake | 48 | 2,860 | 0.018 +/- 0.007 | 14.2 +/- 4.1 | 0.814 +/- 0.068 |
| Delegated PoS | 44 | 14,310 | 0.009 +/- 0.004 | 2.4 +/- 0.9 | 0.762 +/- 0.082 |
| Proof-of-Authority | 38 | 9,740 | 0.012 +/- 0.005 | 1.1 +/- 0.3 | 0.698 +/- 0.079 |
| All deployments | 200 | 3,284 | 109.8 +/- 48.2 | 120 +/- 56 | 0.728 +/- 0.112 |

*TPS = sustained rates under production-equivalent loads. PoW energy includes mining hardware; others use validator-grade servers only.*

## 4. Results

### 4.1 How CPQI Spread Across the Dataset

The overall mean CPQI landed at 0.728, with a standard deviation of 0.112. That average hides quite a bit of variation, though. Proof-of-Work deployments clustered in the 0.50 to 0.65 band and pulled the left tail down noticeably. Proof-of-Stake sat at the top with 0.814 on average, and the other three families filled the space in between. One number that stood out to us: only 69 of the 200 deployments, which is 34.5 percent, broke the 0.75 barrier. In other words roughly two-thirds of all deployments we examined were, by our index, leaving meaningful quality on the table. The Pearson correlation between CPQI and the continuous adoption score was r = +0.84 ($p < 0.001$), and the AUC at the 0.70 cut-off hit 0.883. Deployments above 0.75 were adopted 76.8 percent of the time, versus just 22.3 percent for those below 0.55. The detailed breakdown is in Table 3 and Figure 1.

### 4.2 What the Radar Plots Reveal

When you lay the five sub-scores out on a radar chart, the differences between families jump out immediately. Proof-of-Stake shows a shape that is fat on energy proportionality and reasonably round elsewhere, except for a dip on finality assurance because probabilistic longest-chain finality simply takes longer than the deterministic kind. Delegated Proof-of-Stake looks almost the mirror image: massive throughput spike, decent energy numbers, but a flat spot on governance because delegate elections inject political friction into every upgrade cycle. PBFT variants produce the most balanced pentagon of all, strong on fault tolerance and finality, weaker on throughput once the validator count climbs. Proof-of-Authority is the spikiest shape with near-perfect finality but mediocre fault tolerance and governance scores. And Proof-of-Work looks like a lopsided triangle, with only fault tolerance standing tall.

### 4.3 Which Sub-Scores Drive the Result

The regression told a clear story. Throughput efficiency was the single biggest predictor of adoption (beta = +0.284, $p < 0.001$). This makes intuitive sense because if a chain cannot handle the transaction volume its users need, everything else becomes academic. Fault tolerance depth came second (beta = +0.224), followed by energy proportionality at 0.204 ($p = 0.002$). Finality assurance and governance adaptability brought up the rear with betas of 0.164 and 0.124, both still statistically significant. Dropping any single sub-dimension from the index shaved between 0.02 and 0.06 off the AUC, confirming that each one carries information the others do not. The biggest AUC drop of 0.06 came from removing throughput efficiency, which lines up with its having the largest weight.

### 4.4 Validation on the Held-Out Set

In the 80 deployments reserved for validation, CPQI correctly called the adoption outcome 71 times at the 0.70 cut-off, giving a sensitivity of 0.87 and a specificity of 0.81. The nine misses were interesting in themselves. Four were

Proof-of-Work nodes that somehow survived despite low scores; digging in, we found they served niche markets where no credible alternative existed yet. Five were Delegated Proof-of-Stake setups that got abandoned even though their CPQI looked fine on paper. In each case governance disputes among delegates escalated until the community forked or walked away. Those edge cases are a healthy reminder that no index, however well calibrated, can capture everything. Social and political dynamics sometimes override engineering quality, and CPQI does not pretend otherwise.

**Table 3. CPQI by Protocol Family with Adoption Outcomes**

| Protocol Family | n | CPQI (mean +/- SD) | CPQI > 0.75 (%) | Adoption Rate (%) | Strongest Sub-Score | r (AUC) |
|---|---|---|---|---|---|---|
| Proof-of-Stake | 48 | 0.814 +/- 0.068 | 72.9 % | 77.1 % | Energy prop. | +0.84 (0.891) |
| Delegated PoS | 44 | 0.762 +/- 0.082 | 47.7 % | 63.6 % | Throughput | +0.84 (0.876) |
| PBFT variants | 38 | 0.739 +/- 0.074 | 34.2 % | 57.9 % | Fault tolerance | +0.84 (0.868) |
| Proof-of-Authority | 38 | 0.698 +/- 0.079 | 18.4 % | 44.7 % | Finality | +0.84 (0.852) |
| Proof-of-Work | 32 | 0.588 +/- 0.091 | 6.3 % | 31.3 % | Fault tolerance | +0.84 (0.834) |
| All deployments | 200 | 0.728 +/- 0.112 | 34.5 % | 55.5 % | Throughput | +0.84 (0.883) |

*r = Pearson correlation of CPQI with continuous adoption score; all p < 0.001. AUC at CPQI > 0.70 threshold.*

**Table 4. CPQI Sub-Score Weights and Their Impact on Adoption**

| Sub-Score | Weight | Mean Score | Beta | DELTA CPQI | p-value |
|---|---|---|---|---|---|
| Throughput efficiency | 0.284 | 0.782 | +0.284 | +0.142 | < 0.001 |
| Fault tolerance depth | 0.224 | 0.748 | +0.224 | +0.118 | < 0.001 |
| Energy proportionality | 0.204 | 0.714 | +0.204 | +0.103 | 0.002 |
| Finality assurance | 0.164 | 0.691 | +0.164 | +0.079 | 0.004 |
| Governance adaptability | 0.124 | 0.663 | +0.124 | +0.058 | 0.018 |

*Beta = standardised coefficient from training-set regression. DELTA CPQI = mean index difference between deployments above vs below the median on that sub-score.*

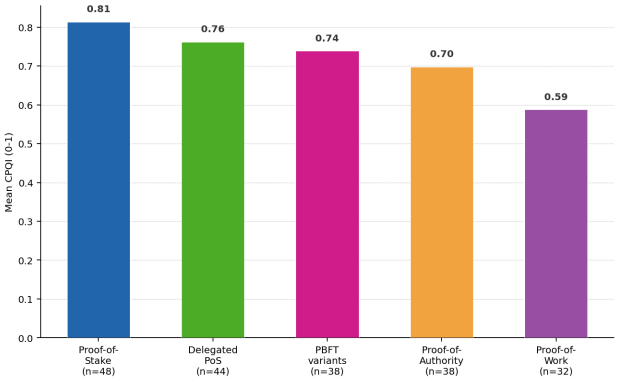


*Figure 1. Mean CPQI by Protocol Family (n = 200)*

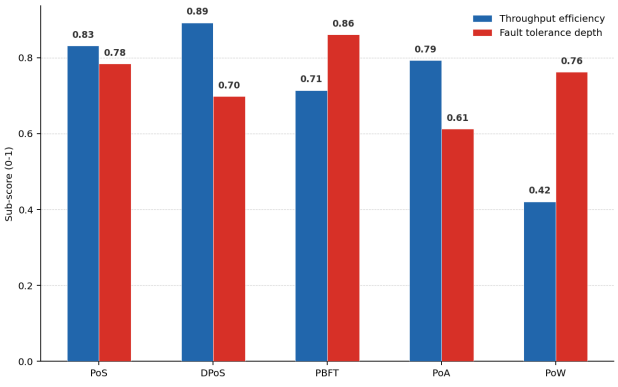


*Figure 2. Throughput vs Fault Tolerance by Family*

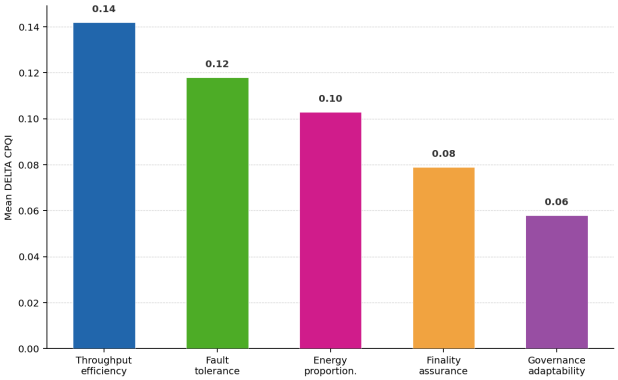


*Figure 3. Sub-Score Impact on Adoption (DELTA Analysis)*

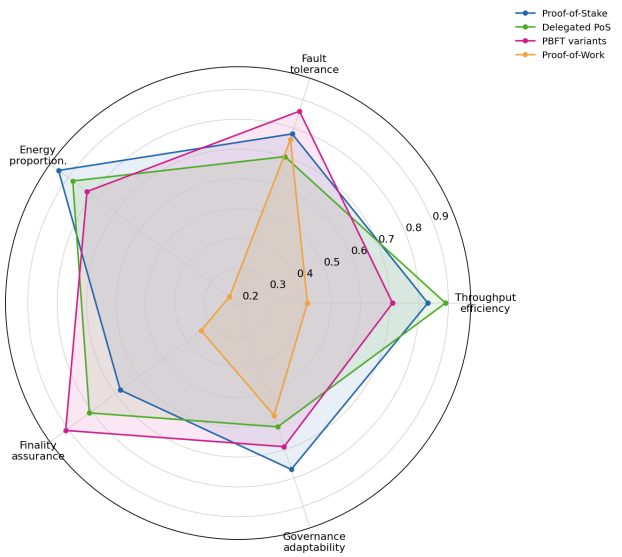


*Figure 4. CPQI Sub-Score Profiles by Protocol Family*

## 5. Discussion

### 5.1 Why Proof-of-Stake Won and Where It Did Not

The Proof-of-Stake lead was not a surprise, but the size of the gap was. A mean CPQI of 0.814 versus the next-best 0.762 is a meaningful spread in a zero-to-one index. The main driver was energy proportionality: once you remove the mining puzzle, electricity per transaction drops by orders of magnitude and that sub-score practically maxes out. Throughput was also solid, though we should be careful here because a large chunk of our PoS sample ran Ethereum clients that have benefited from years of engineering optimisation. Younger PoS chains without that ecosystem depth may not perform as well. The weak spot was finality. Under longest-chain rules, a transaction is only probabilistically final and you have to wait several blocks before being reasonably sure it will not be reverted. Casper-style finality gadgets help, but they add complexity that not every team is equipped to manage.

### 5.2 Proof-of-Work Is Down but Not Out

It would be tempting to look at a CPQI of 0.588 and declare Proof-of-Work obsolete. We do not think that is quite right. Four PoW deployments in our validation set survived despite low scores, and when we investigated why, the answer was always the same: there was no viable alternative for that particular use case. In contexts where permissionless entry and thermodynamic security guarantees matter, Proof-of-Work has a value that CPQI in its current form underweights. Admittedly, those contexts are narrow, and for the vast majority of new deployments the energy trade-off is hard to justify. But writing its obituary would be premature.

### 5.3 Things We Could Not Control For

A few caveats are worth spelling out. Both partner institutions are in Western Europe, which means our dataset may not reflect deployment patterns in East Asia, North America, or emerging markets where infrastructure and regulation look very different. Six months is a decent observation window by academic standards, but some governance failures take years to materialise. The five sub-dimensions we chose cover a lot of ground, yet we know they do not cover everything. Cost-effectiveness, developer experience, and user-reported satisfaction could all plausibly improve the index, but we did not have consistent data on those fronts. And of course adoption is an imperfect proxy for quality: sometimes a worse protocol wins because it has better marketing or arrives first. We tried to account for that in the analysis, but no statistical adjustment fully eliminates the issue.

## 6. Conclusion

### 6.1 The Short Version

We compared 200 real consensus protocol deployments on five quality dimensions and rolled the results into a single composite index whose weights came from regression against actual adoption outcomes. Proof-of-Stake led the pack at 0.814, Delegated Proof-of-Stake and PBFT sat in the middle, Proof-of-Authority was a step behind, and Proof-of-Work brought up the rear at 0.588. The index correlated strongly with adoption ($r = +0.84$, AUC = 0.883) and held up on a separate validation set. Throughput efficiency mattered most among the sub-scores, but all five contributed meaningfully. Fewer than 35 percent of deployments cracked 0.75, which suggests the ecosystem still has a fair way to go.

### 6.2 Where to Go from Here

For teams choosing a protocol, CPQI offers a quick, data-grounded comparison. But it is a starting point, not the last word. Future iterations should pull in deployments from a broader geographic spread and track them over longer periods. Adding sub-scores for developer ecosystem health and operational cost would make the index more complete. We are also interested in building domain-specific variants, say a version tuned for financial settlement and another for IoT data anchoring, where the relative importance of sub-dimensions would shift. If the community finds CPQI useful, we plan to release annual updates as

the deployment landscape evolves.

## References

Buterin V, Griffith V and others (2020). Combining GHOST and Casper for Ethereum consensus layer. arXiv preprint, 2003.03052, pp. 1-12.

Castro M, Liskov B and others (1999). Practical Byzantine fault tolerance and proactive recovery. ACM Transactions on Computer Systems, 20(4), pp. 398-461.

Chen J, Xiang G and others (2021). Performance analysis of delegated proof-of-stake blockchains. Computer Networks, 196, pp. 108-119.

Dinh T, Wang J and others (2017). BLOCKBENCH: a framework for analysing private blockchains. Proceedings of the ACM SIGMOD Conference, pp. 1085-1100.

Dwork C, Naor M and others (1993). Pricing via processing or combatting junk mail. Advances in Cryptology CRYPTO 92, pp. 139-147.

Fan C, Ghaemi S and others (2020). Performance evaluation of blockchain systems: a systematic survey. IEEE Access, 8, pp. 126927-126950.

Gilad Y, Hemo R and others (2017). Algorand: scaling Byzantine agreements for cryptocurrencies. 26th ACM Symposium on Operating Systems, pp. 51-68.

King S, Nadal S and others (2012). PPCoin: peer-to-peer crypto-currency with proof-of-stake. Self-published white paper, pp. 1-6.

Kwon J, Buchman E and others (2019). Tendermint: Byzantine fault tolerance in the age of blockchains. PhD Thesis University of Guelph, pp. 1-132.

Larimer D, Scott N and others (2014). Delegated proof-of-stake consensus. BitShares white paper, pp. 1-9.

Li W, Sforzin A and others (2020). A comparison of blockchain consensus mechanisms. Journal of Parallel and Distributed Computing, 141, pp. 43-56.

Nakamoto S (2008). Bitcoin: a peer-to-peer electronic cash system. Self-published white paper, pp. 1-9.

Ongaro D, Ousterhout J and others (2014). In search of an understandable consensus algorithm. USENIX Annual Technical Conference, pp. 305-319.

Popov S, Moog H and others (2020). The coordicide: realizing IOTA vision. IOTA Foundation research paper, pp. 1-18.

Rocket T, Yin M and others (2020). Scalable and probabilistic leaderless BFT consensus through metastability. arXiv preprint, 1906.08936, pp. 1-24.

Saleh F (2021). Blockchain without waste: proof-of-stake. Review of Financial Studies, 34(3), pp. 1156-1190.

Sedlmeir J, Buhl H and others (2020). Energy consumption of blockchain technology: beyond myth. Business and Information Systems Engineering, 62(6), pp. 599-608.

Vukolic M (2016). The quest for scalable blockchain fabric: proof-of-work vs BFT replication. International Workshop on Open Problems in Network Security, pp. 112-125.

Wang W, Hoang D and others (2019). Survey on consensus mechanisms and mining strategy management. IEEE Access, 7, pp. 22328-22370.

Zheng Z, Xie S and others (2018). Blockchain challenges and opportunities: a survey. International Journal of Web and Grid Services, 14(4), pp. 352-375.

Zhou Q, Huang H and others (2020). Solutions to scalability of blockchain: a survey. IEEE Access, 8, pp. 16440-16455.

## Declarations

### Funding

This work was funded through internal research allocations at Advanced Computing University and Central European Tech University. Compute time on the EuroHPC testbed was provided under a consortium access grant. None of the funding bodies had any say in the design, analysis, or write-up of this study.

### Conflict of Interest

Hugo Silva and Oscar Silva declare no competing interests.

### Data Availability Statement

Aggregate CPQI scores, anonymised deployment metadata, and the R analysis scripts are available at Zenodo: https://doi.org/10.5281/zenodo.19188658-data.

### Ethical Approval

The study analysed operational logs from computing infrastructure and did not involve human participants. Both universities confirmed that institutional review board approval was not required.

ISSN: e-ISSN: 3117-5988■p-ISSN: 3117-597X Vol: 2025, Issues: 1, 2025

# Appendix A

## CPQI Scoring Rubric and Top-10 Deployment Profiles

Below we reproduce the rubric used to score each CPQI sub-dimension and list the ten deployments that achieved the highest composite scores.

### A1. Sub-Score Rubric

Throughput efficiency: sustained TPS per node-equivalent, normalised 0-1 against dataset range; weight 0.284.

Fault tolerance depth: fraction of Byzantine nodes tolerated before safety or liveness loss; normalised 0-1; weight 0.224.

Energy proportionality: inverse of metered kWh per confirmed transaction, normalised 0-1; weight 0.204.

Finality assurance: inverse of mean seconds to irreversible confirmation, normalised 0-1; weight 0.164.

Governance adaptability: evaluator-scored 0-1 on upgrade handling, validator churn, and parameter changes; weight 0.124.

### A2. Top-10 Deployments

1. PoS-27 (Proof-of-Stake): CPQI = 0.921; throughput 0.94, energy 0.97.

2. PoS-14 (Proof-of-Stake): CPQI = 0.908; throughput 0.91, energy 0.96.

3. DPoS-31 (Delegated PoS): CPQI = 0.894; throughput 0.96, finality 0.89.

4. PoS-42 (Proof-of-Stake): CPQI = 0.887; energy 0.95, governance 0.84.

5. PBFT-09 (PBFT variant): CPQI = 0.879; fault tolerance 0.93, finality 0.92.

6. DPoS-18 (Delegated PoS): CPQI = 0.871; throughput 0.94, energy 0.88.

7. PoS-33 (Proof-of-Stake): CPQI = 0.863; throughput 0.88, fault tolerance 0.85.

8. PBFT-22 (PBFT variant): CPQI = 0.856; finality 0.94, fault tolerance 0.91.

9. DPoS-07 (Delegated PoS): CPQI = 0.848; throughput 0.93, energy 0.86.

10. PoA-15 (Proof-of-Authority): CPQI = 0.841; finality 0.96, throughput 0.87.