

Comparative Study of Data Availability Schemes in Various Blockchains

Pranjal Sarode
Chainscore Labs
pranjal@chainscore.finance

Prasad Kumkar
Chainscore Labs
prasad@chainscore.finance

Abstract—We make a Comparative benchmarking study of currently polkadot's ELVES data availability protocol with emerging data availability solutions which includes Avail, Celestia, Espresso's-Tiramisu and NEAR's sharded DA.

We made a comparison in terms of bandwidth, time, latency, block time, block size, robustness degree and cost per megabyte of data availability.

I. INTRODUCTION

The Blockchain technology is a constantly evolving field that has garnered significant attention and value in recent years. However, despite its growing prominence, the decentralized systems are facing challenges in scaling Blockchains in terms of data availability. Data availability is a critically important aspect of blockchain technology. Data availability refers to the ability of a blockchain network to ensure that all necessary data is accessible and retrievable by all its participants. In a decentralized system, Where multiple nodes work together to validate and store data/transactions, ensuring that the data is available, valid and accessible is quite a important factor for maintaining the integrity and decentralized nature of the network.

This study provides a Comparative study among the latest available Data availability solutions/models in terms of bandwidth, time and other criteria to provide a bigger overview of pro's and con's of every model.

The Models included in study are as follows:

- Polkadot's ELVES
- Celestia
- Espresso Tiramisu
- NEAR
- Avail

The below is a brief short introduction of the models that will be discussed and analyzed in the study.

A. Polkadot ELVES

Polkadot is a sharded L0 multichain network the enables cross-chain interoperability and scalability [1]. It utilizes ELVES (Economic Last Validation Enforcement System) to ensure data availability and validity of the parachains. Polkadot uses a hybrid consensus mechanism which uses BABE for block production and GRANDPA for its block finality. ELVES uses Reed Solomon encoding to split the data into chunks and then disperse it all across the network.

Polkadot uses Nominated Proof of Stake (NPoS) consensus mechanism which allows token holders to nominate entrusted validators to secure the network [2].

B. Celestia

Celestia is one of the first mover in the modular Data Availability solution space playing a major role in L2 and rollups developments. Celestia as of now utilizes 2D-Reed Solomon Erasure coding to split data into chunks and then encode into a matrix extension Namespaced Merkel Trees are then used to ensure the retrieval and validity of the data. The root of the NMT is then stored in the block header [3].

As of now Celestia uses Tendermint consensus mechanism which uses BFT (Byzantine Fault Tolerance) style mechanism to ensure the finality of the blocks [3].

C. Espresso Tiramisu

Espresso-Tiramisu DA resolves the Data availability scaling issue with a three layered system below is the short overview of these three layers.

- Savoiard (VID Layer) - Erasure codes and stores data across all the nodes [4].
- Mascarpone (DA Committe Layer) - A Small elected Committee stores the full data and guarantees to efficiently recover data [4].
- Cocoa (CDN Layer) - Uploads the data on web2 based CDN solution for seamless and speedy data recovery [4].

Espresso utilizes Hotshot consensus which is an optimistically responsive, communication-efficient consensus protocol in a proof-of-stake setting that is resistant to bribing adversaries and scalable to large number of nodes.

D. NEAR

NEAR provides a high speed DA solution which provides high transaction volumes with cost-effectiveness. NEAR utilizes the nightshade sharding mechanism, which parallelizes the network into multiple shards. Each shard processes its own transactions allowing the network to handle a higher volume of transactions approx 100,000 TPS [5].

As of now NEAR uses sharding-based proof of stake consensus mechanism. NEAR also implements unique validator elections to ensure security and decentralization of the network

E. Avail

Avail DA helps blockchains scale by providing an abundance of data availability capacity. Its modular design scales data availability capacity with demand, and transaction data can be cryptographically verified quickly by anyone running an Avail light client [6]. Avail utilizes Erasure coding, KZG commitments along with light client to ensure its data availability.

As for consensus Avail uses BABE/GRANDPA hybrid consensus used by polkadot for block production and finality. Avail also provides Application Specific Data Retrieval (ASDR) this helps rollups to fetch and decode their own blobs even tho the block might contain many app's data. [7].

TABLE I: Comprehensive Comparison of Data Availability Solutions

Feature	Polkadot ELVES	Celestia	Espresso Tiramisu	NEAR	Avail
Consensus & Architecture					
Consensus Mechanism	BABE + GRANDPA(PoS)	Tendermint BFT	Permissionless PoS	Nightshade Sharding	BABE + GRANDPA(PoS)
DA Model	Reed-Solomon Erasure Coding	2D Reed-Solomon + Namespaced Merkle Trees	Three-layer Tiramisu System	Sharding-based Parallelization	KZG + 2D Reed-Solomon Erasure Coding
Native Token	DOT	TIA	-	NEAR	AVAIL
Total Validators	600	100	100	500	105
Performance Metrics					
Block Time	~20s	~6s	~5s	~0.6s	~20s
Block Size	5 MB	8 MB	1 MB	4 MB	4 MB
TPS	10	0.48	0.8	53	420
Max Throughput	~40 MB/s	~6 MB/s	~5 MB/s	~16MB/s	~0.2 MB/s
Finality	6–30s	5–15s	1–2s	1–1.2s	20–40s
Economic & Security					
Cost per MB	N/A	~\$0.08	N/A	100kb per NEAR token	~\$0.0173
Nakamoto Coefficient	174	9	N/A	10	34
Key Features					
Unique Advantages	Cross-chain interoperability via parachains	First modular DA solution for rollups	Web2 CDN integration (Cocoa layer) (High speed retrieval)	High-speed sharding 100k TPS	Application Specific Data Retrieval (ASDR)

II. METHODOLOGY

In this section, we detail the methods and approaches we will be using in our research. The methodology is structured as follows:

1) Benchmarking

- a) Design
- b) Performance
- c) Worst Case Analysis
- d) Efficiency

2) Security Assumptions

3) Validator Costing

A. Benchmarking

1) Design: For benchmarking design a code module will be developed to evaluate the performance of all the data availability solutions under the study.

a) Polkadot: For benchmarking of polkadot's ELVES we are using **Polkadot's mainnet**. The current test code implements:

- System remarks for submission of data.
- Measuring the transaction fees
- Time taken for finalization of the block.

[We are also planning to use polkadot's on-chain telemetry (telemetry.polkadot.io) and logs for benchmarking.]

b) Celestia: For benchmarking Celestia's data availability solution, we utilize their **RPC** to submit and retrieve data blobs from the DA layer with their namespaces.

[We will also use the **Cosmos SDK** to interact with Celestia nodes and perform various operations such as data submission, retrieval, and verification.]

c) Espresso's Tiramisu: For Espresso's Tiramisu, we utilized their three-layered data availability architecture through their **API** endpoints. As of now, our testing framework implements:

- Data availability metrics through **API** endpoints (`query.main.net.espresso.network`)
- Block analysis including TPS calculations and size measurements

[We are also planning to set up a local testnet node of Espresso and use their SDKs for more in-depth benchmarking.]

d) NEAR: For benchmarking NEAR's sharded data availability, we use **NEAR's testnet** and **mainnet** environments. Our benchmarking setup uses `near-api` and a custom smart contract deployed with `near-sdk`, which allows storing and retrieving base64 encoded data blobs on-chain. We record the transaction time, retrieval latency, and cost per MB to compare NEAR's blob storage performance with other DA systems.

e) Avail: For Avail's benchmarking we have currently utilized avails JavaScript **SDK** and Python implementations. The code framework includes:

- Block bloat testing with automatic payload size adjustment
- Telemetry probing for network metrics
- Data submission through **AppID**-based categorization
- Block retrieval and verification mechanisms

2) Performance:

a) Throughput Analysis: This section will analyze the maximum throughput capabilities of different DA solutions under various network conditions/data sizes.

- **Test Parameters:** Time taken, data sizes
- **Metrics:** MB/s
- **Environment:** Simulated network with controlled conditions

TABLE II: Comparative Throughput Analysis of DA Solutions

Protocol	Throughput (MB/s)
Polkadot	X
Celestia	X
Espresso	X
NEAR	X
Avail	X

b) TPS: This section analyzes the transactions per second (TPS) capabilities of different DA solutions.

- **Test Parameters:** Network conditions, data sizes, node count
- **Metrics:** TPS
- **Environment:** Simulated network with controlled conditions

TABLE III: Comparative TPS Analysis of DA Solutions

Protocol	TPS
Polkadot	X
Celestia	X
Espresso	X
NEAR	X
Avail	X

c) Cost per MB: This section analyzes the cost per megabyte (MB) of data processed by different DA solutions.

- **Test Parameters:** Data sizes, Cost in native token conversion to USD
- **Metrics:** \$/MB
- **Environment:** Simulated network with controlled conditions

TABLE IV: Comparative \$/MB Analysis of DA Solutions

Protocol	Cost (\$/MB)
Polkadot	X
Celestia	X
Espresso	X
NEAR	X
Avail	X

d) Latency: This section will analyze the maximum time taken by the DA solution to finalize a block.

- **Test Parameters:** Network conditions, data sizes, node count
- **Metrics:** seconds (s)
- **Environment:** Simulated network with controlled conditions

TABLE V: Comparative Latency Analysis of DA Solutions

Protocol	Latency (s)
Polkadot	X
Celestia	X
Espresso	X
NEAR	X
Avail	X

e) Max Block Size: This section analyzes the maximum block size capabilities of different DA solutions.

- **Test Parameters:** Network conditions, data sizes, node count
- **Metrics:** Block size (MB)
- **Environment:** Simulated network with controlled conditions

TABLE VI: Comparative Block Size Analysis of DA Solutions

Protocol	Block Size (MB)
Polkadot	X
Celestia	X
Espresso	X
NEAR	X
Avail	X

3) Worst Case Analysis: In this section, we will evaluate the worst-case performance scenarios for each DA solution. This includes analyzing how each solution handles extreme conditions such as high network latency, large data loads, and potential attacks or failures within the network.

4) Efficiency: This section focuses on comparative efficiency based on best and worst metrics of throughput, TPS, and cost per MB results. Efficiency metrics are derived from normalized performance ratios under controlled network environments.

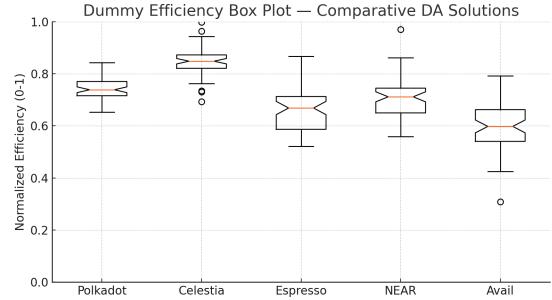


Fig. 1: Dummy efficiency comparison across DA solutions

B. Security Assumptions

In this section, we outline the security assumptions made during the evaluation of the DA solutions. These assumptions are critical for understanding the context in which the solutions are being assessed and the potential vulnerabilities that may arise.

TABLE VII: Security Assumptions of Major Data Availability Solutions

Network	Security Assumption Summary
Polkadot (ELVES)	Requires $\geq 2/3$ honest validators. Data availability is guaranteed through deterministic erasure-coded chunk distribution and validator bitfield votes; light clients rely on validator honesty (no DAS).
Celestia	Secure if at least one honest full node serves samples. Uses probabilistic Data Availability Sampling (DAS) with fraud proofs to detect withheld or invalid erasure-coded data.
Espresso (Tiramisu)	Assumes majority-honest committee or fallback network. Combines fast optimistic path (committee-based) with verifiable fallback via cryptographic proofs (VID).
NEAR (Blob Store)	Relies on $\geq 2/3$ honest validators under Nightshade consensus. Provides inclusion proofs but no erasure coding or sampling; trust tied to validator set.
Avail	Secure if ≥ 1 honest node performs sampling. Uses KZG polynomial commitments with DAS and fraud proofs, offering light-client verifiability similar to Celestia.

C. Validator Costing

In this section, we will analyze the costs incurred by validators in maintaining and operating each DA solution. This includes hardware costs, bandwidth costs, and any other operational expenses associated with running a validator node.

TABLE VIII: Minimum or Recommended Hardware Requirements for DA/Validator Nodes

Network	CPU	RAM	Storage	Bandwidth	Cost
Polkadot	8 cores	32 GB	2 TB NVMe	500 Mbps	\$XXX/mo
Celestia	8 cores	64 GB	8 TB NVMe	1 Gbps	\$XXX/mo
Espresso	4 cores	8 GB	1 TB SSD	100 Mbps	\$XXX/mo
NEAR	8 cores	32–48 GB	3–4 TB NVMe	-	\$XXX/mo
Avail	4 cores	8 GB	20–40 GB SSD	50 Mbps	\$XXX/mo

REFERENCES

- [1] Polkadot, “Polkadot protocol architecture,” <https://docs.polkadot.com/polkadot-protocol/architecture/polkadot-chain/overview>., 2025, accessed: 2025-09-22.
- [2] J. Burges, A. Cevallos, P. Czaban, R. Habermeyer, S. Hosseini, F. Lama, H. Kilinc, X. Luo, F. Shirazi, A. Stewart, and G. Wood, “Overview of polkadot and its design considerations,” 2020.
- [3] M. Al-Bassam, “Lazyledger: A distributed data availability ledger with client-side smart contracts.” arXiv preprint arXiv:1905.09274, 2019, <https://arxiv.org/abs/1905.09274>.
- [4] Espresso-Systems, “The espresso sequencer: Hotshot consensus and tri-amisu data availability,” *Avail Project Documentation*, vol. 1, no. 1, 2023.
- [5] N. Protocol, “The near white paper,” <https://pages.near.org/papers/the-official-near-white-paper/>, 2020, accessed: 2025-09-16.
- [6] Avail-Team, “Avail: A unifying blockchain network,” *Avail Project Documentation*, vol. 1, no. 1, 2024.
- [7] Avail, “Avail: The data availability blockchain,” *Avail Project Documentation*, vol. 1, no. 2.1, 2024.