Standardizing Blockchain Layer 2 Benchmarking

by

Nathaniel Rex Cable

A Thesis

Presented to the Graduate and Research Committee of Lehigh University in Candidacy for the Degree of Master of Science

 in

Computer Science

Lehigh University May 2025 © 2025 Copyright Nathaniel Cable All Rights Reserved

Thesis Signature Sheet

Thesis is accepted and approved in partial fulfillment of the requirements for the Master of Science in Computer Science .

Thesis Title: <u>Standardizing Blockchain Layer 2 Benchmarking</u>

Name: Nathaniel Cable

LIN: 898138448

4/30/2025

Date Approved

-Signed by:

Dinecton Advisor Henry F. Korth

Co-Advisor

---- DocuSigned by:

Brian Davison Departmente Chair/Second Reader Brian Davison

Committee Member

Acknowledgments

The content of this thesis covers research I have done over the course of the latter half of my undergraduate degree and full graduate degree at Lehigh University. This research has been a part of a Blockchain Projects team including Lehigh University students David Cueva and Nicholas Carnevale, and Bucknell University student Jacob Piskadlo. Our team has also worked in collaboration with a team from the National University of Singapore, led by PhD student Kunpeng Ren and advisor Beng Chin Ooi. I would like to thank my professor and graduate advisor, Hank Korth, for facilitating and guiding our research team. I would also like to thank my fellow team members for continuously pushing forward and overcoming the many challenges we faced along the way.

Table of Contents

A	Acknowledgments		
Li	st of	Figures vii	
A	Abstract 1		
1.	Intr	oduction 3	
2.	Bloc	ckchain Benchmarking 4	
	2.1	The Blockchain Benchmarking Standardized Framework 4	
	2.2	Blockbench-v3 Results	
		2.2.1 Environment	
		2.2.2 Results	
		2.2.3 Conclusions	
	2.3	Expanding to Layer 2 Systems	
3.	laye	r 2 Solutions 11	
	3.1	Architectures	
		3.1.1 State Channels	
		3.1.2 Sidechains	
		3.1.3 Plasmas	
		3.1.4 Rollups	
	3.2	Modularity 16	
	3.3	Data Availability	

	3.4	The Finality Problem	20
	3.5	Current Ecosystem	23
	3.6	Motivation for BBSF L2 Expansion	24
4.	Bloo	ckbench-L2	27
	4.1	Workloads	27
	4.2	Metrics	30
	4.3	Driver	34
	4.4	Preliminary Results	34
5.	Con	clusion and Future Work	37
	5.1	Summary of Contributions	37
	5.2	Limitations	38
	5.3	Future Work	39
		5.3.1 Short Term	39
			40
		5.3.2 Long Term	40

List of Figures

2.1	Sui Benchmark Results	8
3.1	Modular Blockchain Architectures, from [2]	17
4.1	zkSync Benchmark Results, Partial Trust	36

Abstract

As blockchain adoption continues to grow, developers and businesses face an ever-expanding ecosystem of platforms, each offering unique trade-offs. For developers, efficient blockchain performance enables faster transaction times, lower fees, and a more responsive user experience. For businesses, the choice of blockchain platform directly impacts scalability, cost-effectiveness, and the ability to meet user demands. However, the lack of standardized benchmarking tools has made it difficult to objectively assess and compare the performance of blockchain platforms. Performance benchmarking is thus crucial to determining the suitability of a blockchain for specific use cases and applications.

To address this pressing need, the Blockchain Benchmarking Standardized Framework (BBSF) was developed to provide a consistent methodology for evaluating blockchain performance. The framework defines standardized workloads, metrics, benchmarking drivers, and reporting formats, offering a foundation for transparent and reproducible comparisons. Previous work on the BBSF focused on benchmarking layer 1 (L1) blockchains through the Blockbench-v3 implementation, using Web3-style workloads. These efforts provided valuable insights into the comparative performance of different L1 platforms, uncovering strengths and limitations that have practical implications for developers and enterprises alike.

As the blockchain ecosystem has evolved, the emergence of layer 2 (L2) solutions has introduced new opportunities—and challenges—for scaling

and cost efficiency. L2 platforms aim to address the limitations of L1 systems, such as transaction throughput bottlenecks and high fees, but their varied architectures and design choices present new complexities. These include differences in finality assumptions, decentralization models, and interactions with L1 systems. While the foundational principles of the BBSF remain applicable, directly applying an L1-oriented framework like Blockbench-v3 to L2 systems fails to capture critical performance characteristics unique to L2 architectures.

This thesis presents a comprehensive study of blockchain benchmarking, extending the BBSF to encompass the complexities of L2 solutions. It evaluates a new standardized benchmarking framework tailored to L2 systems, addressing their unique properties and challenges. Empirical testing on prominent L2 platforms, including zkSync and additional candidates, highlights the framework's effectiveness and provides actionable insights into their performance. Furthermore, this thesis explores benchmarking results for L1 platforms, such as Sui, to provide a comparative foundation for analyzing L1 and L2 performance. By integrating theoretical advancements with practical experimentation, this work seeks to establish a robust and adaptable approach to benchmarking that can guide developers, researchers, and enterprises in making informed decisions.

1. Introduction

In the rapidly evolving blockchain ecosystem, performance measure is a critical factor in both technical decision-making and business strategy. For developers, efficient blockchain performance can mean lower latency times, lower fees, and a more responsive user experience. For businesses large and small, the choice of blockchain platform directly impacts scalability, cost-effectiveness, and the ability to meet user demands. As the blockchain ecosystem grows, having standardized tools to compare and assess these platforms becomes an essential need across the board.

This thesis presents work focused on the performance benchmarking of layer 1 and layer 2 blockchains, building upon the Blockchain Benchmarking Standardized Framework (BBSF) developed in previous research[14]. The layer 1 (L1) benchmarking efforts provided valuable data and insights into the comparative strengths and weaknesses of different blockchain platforms. More recent work extends this framework to layer 2 (L2) solutions, which have gained significant traction in addressing scalability and transaction cost challenges present at L1.

In the following sections, I present performance data collected from both layer 1 and layer 2 platforms, offering an analysis of their comparative performance. This data serves as the foundation for a standardized approach to blockchain benchmarking across multiple layers, helping both technical and business teams make informed decisions in a rapidly evolving market.

2. Blockchain Benchmarking

2.1 The Blockchain Benchmarking Standardized Framework

The rapid growth of blockchain technology has brought an equally rapid expansion of platforms, each claiming unique advantages in performance, scalability, and usability. However, without standardized methodologies for evaluating performance, these claims are often difficult to verify or compare objectively. Recognizing the need for an impartial and reproducible benchmarking framework, the Blockchain Benchmarking Standardized Framework (BBSF) [14] was developed to address this gap. By providing a consistent methodology for benchmarking blockchain platforms, the BBSF facilitates fair comparisons and informed decision-making for developers and businesses alike.

The BBSF builds on principles established in other domains, such as the Transaction Processing Performance Council (TPC) [9] benchmarks for database systems, adapting these principles to the unique properties of blockchain architectures. The framework defines several key components, including standardized workloads, performance metrics, benchmarking driver design, and reporting formats. Together, these elements ensure that benchmarking data is transparent, comparable, and relevant to realworld applications.

An implementation of the BBSF, known as Blockbench-v3, serves as a

practical tool for evaluating layer 1 blockchains. Blockbench-v3 utilizes Web3-style workloads to simulate diverse transaction scenarios, providing insights into micro-metrics such as throughput and latency, as well as macro-metrics like scalability. Previous benchmarks through Blockbenchv3 focused on prominent L1 blockchains, including Ethereum and Solana, revealing critical insights into their scalability and performance trade-offs. For example, while Ethereum's established ecosystem offers a wide array of decentralized applications, its performance can be constrained by high transaction fees and latency. In contrast, newer platforms like Aptos and Sui demonstrate promising advancements in throughput and efficiency, albeit with less developed ecosystems.

Beyond evaluating individual platforms, the results from Blockbenchv3 have reinforced the viability and value of a standardized benchmarking framework. They demonstrate that the BBSF can provide developers and organizations with actionable data for choosing platforms tailored to their specific needs. The framework also encourages blockchain projects to adopt transparent and consistent performance metrics, fostering trust and comparability across the broader ecosystem.

The insights gained through L1 benchmarking also serve as a foundation for exploring new frontiers in blockchain scalability. The rise of layer 2 solutions introduces additional complexities, including varied architectures, definitions of transaction finality, and interactions with L1 systems. These challenges require adaptations to the BBSF to ensure that benchmarking remains effective in capturing the unique characteristics of L2 platforms.

2.2 Blockbench-v3 Results

At the time the BBSF paper was published, there was continuing work being done to write the smart contract workloads and drivers for both Sui and Aptos. Sui and Aptos are both layer 1 blockchains that use the Move programming language as the basis for their smart contracts, and they both claimed much higher throughput results than Ethereum. Our team believed these would be two great chains to gather benchmarking data on using the Blockbench-v3 framework, with their results being used to further back the efficacy of the BBSF as a framework design standard and Blockbench-v3 as a layer 1 implementation of said standard. As mentioned in Section 2.2.2, we did go this direction and collected preliminary data on Sui, with the Aptos driver and workloads completed and ready to run.

2.2.1 Environment

Prior benchmarking results for layer 1 blockchains like Ethereum, as shown in the BBSF paper, were gathered in collaboration with the Data System Research Group at the National University of Singapore. All benchmarks in this study were performed on a dedicated server equipped with dual Intel Xeon E5-2643 v3 processors, providing a total of 12 physical cores and 24 logical threads. The system is configured with 256 GiB of RAM, ensuring sufficient memory resources for running full blockchain nodes, benchmarking drivers, and workload generators without contention. The server architecture includes two NUMA nodes, corresponding to the two physical CPUs, which ensures high memory bandwidth and parallel processing capabilities essential for accurate performance evaluation. This hardware setup was selected to minimize resource bottlenecks and to provide a stable, high-performance environment for benchmarking activities. All components, including the blockchain node software, benchmarking framework, and workloads, were run simultaneously on this machine to maintain consistency and eliminate performance variability introduced by networked or distributed deployments. Due to the difference in benchmarking environments, layer 1 results shown here may not be directly comparable to the results provided in the BBSF paper.

2.2.2 Results

To evaluate performance, three workloads were executed on a single Sui full node: Simple Transfer, Decentralized Exchange (DEX), and NFT Marketplace. Each workload was designed to stress different aspects of the system, as described in Section 4.1.

The performance results of these workloads are summarized in Figure 2.1. This figure presents key metrics collected during the benchmarking process, including transaction throughput and latency, providing a comparative view of Sui's performance under different application scenarios. Each of the workloads were run a total of five times, and the average was taken for each measurement.

Currently, sufficient publishable benchmarking data for Aptos is not yet available. Future work will aim to address this gap and incorporate publishable Aptos results. Likewise, our Sui results do not contain data on the NFT Minting and Sports Betting contracts, or macro-metrics like Scalability obtained from running workloads with a varying number of full blockchain nodes. Future work will aim to include these results as well.



Figure 2.1: Sui Benchmark Results

2.2.3 Conclusions

Based on the results gathered and presented in this thesis, as well as the benchmarking outcomes published in the BBSF paper, the efficacy of the BBSF as a standardized framework for layer 1 blockchain evaluation has been demonstrated. The framework effectively captured key performance metrics across a diverse set of workloads and blockchain systems, enabling consistent, reproducible comparisons. These results validate the BBSF's design goals of standardization, flexibility, and applicability across different blockchain architectures, establishing it as a reliable foundation for future benchmarking efforts.

Given that the BBSF has proven effective for developing a layer 1 benchmarking framework, it provides a strong foundation for extension into the layer 2 ecosystem. The core principles of standardized workload design, metric collection, and evaluation methodology are equally relevant when assessing the performance of layer 2 solutions. Therefore, adapting and extending the BBSF to develop a standardized layer 2 benchmarking framework is a logical next step. This adaptation will allow for consistent, comparable evaluations across a rapidly growing and diverse set of layer 2 technologies, supporting both academic research and industry adoption.

2.3 Expanding to Layer 2 Systems

The Blockchain Benchmarking Standardized Framework has proven its effectiveness as a tool for evaluating layer 1 blockchain platforms, offering clear insights into performance through standardized workloads, metrics, and reporting formats. However, the continuous evolution of the blockchain ecosystem has introduced new layers of complexity with the boom of new layer 2 scaling solutions. These platforms aim to enhance scalability and reduce costs by operating as extensions to L1 systems while inheriting their security guarantees. As L2 platforms gain prominence, so does the need for a tailored benchmarking framework that accounts for their distinctive architectures, operational models, and performance characteristics, similar to what Blockbench-v3 is for L1 chains.

Layer 2 systems address critical limitations of L1 blockchains, such as high transaction fees and limited throughput, by processing transactions off-chain or within parallel execution environments. This results in significant scalability improvements. However, L2 platforms introduce a multitude of architectural and functional variations, ranging from state channels[3] and sidechains[6] to optimistic and zero-knowledge (ZK) rollups. These variations bring forth various definitions of finality, decentralization models, and data availability mechanisms, all of which complicate direct comparisons.

For instance, optimistic rollups rely on probabilistic finality during their challenge periods, whereas zk-rollups achieve deterministic finality through succinct proofs. Similarly, decentralized data availability networks may provide varying levels of performance and trustworthiness compared to centralized solutions. A benchmarking framework designed for L1 systems is not equipped to capture these nuances, necessitating the extension of the BBSF to L2 platforms.

3. Layer 2 Solutions

With the acceleration of widespread blockchain adaptation, the limitations of layer 1 systems have driven the emergence of numerous layer 2 solutions. Operating as extensions of L1 networks, L2 platforms aim to optimize scalability and reduce costs while inheriting the security and ecosystem compatibility of their underlying L1 chains. Due to it's dominance in the space, Ethereum serves as the base layer for a majority of layer 2 solutions in the current ecosystem. In recent years, the growth of L2 architectures has reshaped the blockchain landscape, with rollup-centric solutions dominating the space.

3.1 Architectures

Layer 2 solutions encompass a diverse array of architectures, each tailored to specific use cases and performance goals. In a broad sense, L2 systems can currently be categorized into state channels, sidechains, plasmas, and rollups. Among these, rollups, such as optimistic rollups and zeroknowledge rollups, have emerged as the dominant approach due to their scalability and compatibility with Ethereum-based applications.

3.1.1 State Channels

State Channels facilitate off-chain transaction execution by enabling participants to transact directly within a private, pre-established channel, significantly reducing the computational burden on the underlying layer 1 blockchain. A well-known example is the Lightning Network [13], which operates as a state channel solution for Bitcoin, allowing users to conduct rapid, low-cost transactions by aggregating multiple exchanges off-chain before settling final balances on the main chain. These channels require parties to lock assets in a multi-signature contract on-chain before engaging in transactions off-chain, enabling near-instantaneous execution without network congestion. The final state of the transactions is periodically committed to the main chain, ensuring security through cryptographic signatures while minimizing on-chain interaction. While state channels offer considerable advantages in speed and cost efficiency, they are best suited for applications with predefined participants, such as micropayments or gaming ecosystems, due to their reliance on cooperative exit strategies and the necessity for participants to remain vigilant against fraud. Their limited applicability in broader decentralized finance (DeFi) contexts has constrained their adoption compared to rollup-based solutions, which maintain broader interoperability with smart contracts and existing blockchain ecosystems.

3.1.2 Sidechains

Sidechains^[6] operate as independent blockchains that run parallel to a layer 1 network, utilizing a pegged currency mechanism to facilitate asset transfers between the two chains. Participants lock assets on the main chain, which then releases an equivalent amount on the sidechain, enabling transactions to occur with greater speed and lower cost compared to layer 1 execution. Unlike state channels, sidechains do not require predefined participants, allowing users to engage in transactions without prior setup beyond asset bridging. The primary advantage of sidechains lies in their ability to partition assets between chains, creating parallelism in transaction processing while reducing congestion on the primary blockchain. However, security is a critical consideration, as sidechains operate with their own consensus mechanisms, making them susceptible to attacks if validator pools are insufficiently decentralized. Additionally, sidechains rely on periodic asset transfers back to layer 1 for users to regain full security assurances. Despite these trade-offs, sidechains have proven effective for scalability, particularly in scenarios where transaction data does not need to be immediately secured on the base chain, ensuring efficient execution without compromising usability. Their flexibility in bridging mechanisms and asset management continues to make them a viable layer 2 option within the blockchain ecosystem, though from a benchmarking perspective, align closer to a layer 1-tailored framework.

3.1.3 Plasmas

Plasmas[12] function as hierarchical chains that extend the layer 1 blockchain by processing transactions off-chain while maintaining security through fraud-proof mechanisms. In this architecture, individual assets, often represented as fixed units of currency, are treated similarly to NFTs, with each transaction being tracked separately. Plasma chains periodically submit aggregated transaction data to the layer 1 chain, where fraud proofs allow invalid transactions to be challenged, ensuring overall integrity. However, this structure introduces latency concerns, particularly in withdrawals, as users must wait through an extended challenge period before assets can be safely returned to layer 1. Additionally, plasma chains require asset owners to actively monitor their holdings, introducing user attention requirements that may limit widespread adoption. While plasmas effectively reduce congestion on the main chain and enable scalability, their reliance on asset tracking and prolonged exit periods has restricted their application primarily to specialized financial use cases rather than broad DeFi solutions. Newer layer 2 approaches, particularly rollups, have largely replaced plasmas in terms of general usability and integration with existing smart contract platforms.

3.1.4 Rollups

Rollups[15] have become the most widely adopted layer 2 scaling solution due to their ability to enhance performance while maintaining the security guarantees of the underlying layer 1 blockchain. Unlike sidechains or state channels, rollups batch multiple transactions off-chain before submitting a compressed representation to layer 1, significantly reducing gas costs and improving throughput. The two primary categories of rollups, Optimistic and Zero-Knowledge (ZK), differ in their approach to transaction validation. Optimistic rollups, such as Arbitrum[10], assume transactions are valid unless proven otherwise, requiring a designated challenge period where fraudulent transactions can be disputed through fault proofs. This approach introduces latency in finalizing transactions but allows rollups to execute standard Ethereum smart contracts without modification. On the other hand, zk-rollups, such as zkSync[18], rely on cryptographic zeroknowledge proofs to almost instantly verify the correctness of a transaction batch[7]. While zk-rollups achieve faster finality and greater security, they incur additional computational costs in generating proofs, introducing overhead that may be difficult for individual users to run.

Beyond these core designs, rollups have expanded into modular architectures, which separate execution, settlement, consensus, and data availability into distinct layers, optimizing for performance and interoperability. Validiums[4], a subset of zk-rollups, store data off-chain rather than on layer 1, increasing scalability but introducing trust assumptions regarding off-chain storage providers. The Dencun upgrade on Ethereum further improves rollup efficiency with proto-danksharding[5], enabling rollups to leverage on-chain data blobs for temporary storage. Additionally, emerging *layer 3* architectures envision rollups as multi-layered systems, facilitating application-specific scalability by allowing L3 chains to commit to an L2 before final settlement on L1. As rollups evolve, their role in blockchain scalability extends beyond transaction batching to broader ecosystem integration, including cross-chain rollups, such as ZKM's entangled rollup[17], which seeks to unify multiple layer 1 networks under a common layer 2 framework. These advancements illustrate the dynamic competition within layer 2 solutions, where rollups continuously refine their architectures to optimize decentralization, security, and performance.

Unlike the other L2 architectures, rollups preserve Ethereum's security model while significantly improving transaction throughput and cost efficiency, currently making them the most viable scaling approach for decentralized applications. Additionally, the variation between optimistic and ZK, as well as modularly-designed rollups, presents unique benchmarking challenges, such as latency measurement, data availability, and differing finality assumptions, which must be systematically addressed. Given their dominance in the L2 ecosystem and ongoing innovations, benchmarking rollups provides meaningful insights into layer 2 performance while ensuring comparability across different rollup implementations. Due to their widespread adoption, scalability benefits, and alignment with Ethereum's evolving infrastructure, rollup-based solutions will be the primary focus of our layer 2 benchmarking framework.

3.2 Modularity

Modular blockchain architectures redefine scalability by decoupling key components into distinct layers, optimizing performance, security, and flexibility. Unlike monolithic blockchains, which bundle execution, consensus, settlement, and data availability within a single framework, modular architectures assign these responsibilities to separate systems, allowing each to specialize in its function while interacting cohesively. These various *layers* and their functions are as follows:

- **Consensus layer:** Ensures the validity and ordering of transactions through a distributed network of nodes. It determines the current state of the blockchain by implementing cryptographic protocols that enable agreement among participants. Modular designs frequently rely on existing layer 1 networks for consensus, leveraging their security while outsourcing execution to layer 2 solutions.
- Settlement layer: Finalizes transactions by ensuring they are immutable and enforceable within the blockchain framework. This layer is crucial for rollups, which commit batched transactions to layer 1 for security guarantees. In some cases, rollups handle settlement within

their own architecture, introducing alternative mechanisms for dispute resolution and transaction validation.

- Data Availability layer: Governs the storage and accessibility of transaction data. While traditional blockchains retain all data onchain, modular systems introduce off-chain solutions such as Celestia and EigenDA, which reduce costs and enhance scalability. Optimistic and zk-rollups depend on this layer to ensure verifiable access to transaction data, with varying degrees of decentralization and trust assumptions. Section 3.3 provides more on data availability .
- Execution layer: Responsible for processing transactions and smart contracts. In rollups, this occurs off-chain, enabling high-speed execution before finalization on the settlement layer. Different rollup architectures implement execution uniquely, balancing performance against cryptographic verification requirements.



Figure 3.1: Modular Blockchain Architectures, from [2]

By distributing these functions across specialized components, modularity enhances blockchain scalability and interoperability while preserving decentralization principles. The ability to select best-in-class solutions for each layer provides a tailored approach to layer 2 design, improving efficiency without compromising security. This layered structure is well illustrated in Figure 3.1, which outlines the roles of execution, settlement, consensus, and data availability in a modular, rollup-centric blockchain design [2]. As both established and newer blockchains integrate modular solutions, this architecture is poised to shape the next generation of decentralized systems.

3.3 Data Availability

Data availability plays a critical role in the security, scalability, and reliability of layer 2 solutions, ensuring that transaction data remains accessible for validation and dispute resolution. Unlike layer 1 systems, where all transaction data is stored on-chain, L2 solutions implement varied approaches to data availability depending on their architecture, balancing cost-efficiency with decentralization. These models directly impact performance benchmarks, as data availability mechanisms influence transaction finality, throughput, and trust assumptions. Layer 2 solutions generally fall into two broad categories regarding data storage:

• On-Chain Data Availability: Some L2 solutions store transaction data directly on the L1 chain, leveraging the security guarantees of Ethereum (or another base layer). This approach, commonly used in zk-rollups, ensures that transaction data is publicly accessible and verifiable. The Ethereum proto-danksharding upgrade, introduced via EIP-4844, optimizes on-chain data storage with "data blobs" that remain available for a fixed period, reducing cost while maintaining accessibility. Additionally, many systems adopt a calldata approach, embedding transaction data in the calldata field of L1 transactions to minimize storage overhead and reduce gas fees, while still ensuring that the data remains transparent and verifiable.

• Off-Chain Data Availability: Other L2 solutions, such as validiums, store transaction data off-chain while committing cryptographic proofs of correctness to L1. These systems rely on external data availability committees (DACs) or decentralized storage networks (e.g., Celestia, EigenDA), which must be trusted to retain and serve data upon request. While off-chain models improve scalability and reduce costs, they introduce trust assumptions and potential risks related to data retrieval.

The data availability model adopted by any L2 system directly influences benchmarking outcomes, requiring careful consideration of how finality, throughput, and decentralization are assessed. When it comes to finality assumptions, off-chain data availability can extend perceived finality beyond traditional L1 settlement, as transactions may be considered "final" within L2 before full data availability is confirmed. Benchmarks must account for varying notions of finality based on trust assumptions. For throughput comparisons, L2 solutions relying on external data availability layers may achieve higher throughput by reducing on-chain storage requirements, but their performance depends on the responsiveness and integrity of these off-chain providers. Benchmarking must measure both raw transaction processing speed and data retrieval latency to offer an accurate performance evaluation. For a macro-metric like decentralization, systems utilizing off-chain data availability introduce an additional centralization variable, as trust in a specific provider or committee affects overall security. Benchmarks should consider the reliability, number of participants, and fault tolerance of the off-chain data layer.

The evolution of modular rollups and cross-chain data availability solutions will further complicate benchmarking methodologies. Rollups integrating independent data-availability networks (rather than relying on L1 storage) may require new micro-metrics to assess data integrity, redundancy, and retrieval efficiency. Additionally, multi-layered architectures blending L2 and L3 solutions will introduce transaction workflows where data availability expectations vary based on cross-chain interactions.

To ensure a fair and standardized benchmarking framework, defining uniform metrics that accommodate diverse data availability models is essential. This includes separate benchmarks for data latency, failure recovery, and trust assumptions, allowing users and developers to evaluate L2 performance in the context of real-world applications.

As L2 ecosystems continue to evolve, the role of data availability will remain central to performance evaluations, shaping how blockchain scalability is measured and optimized.

3.4 The Finality Problem

Finality is the assurance that a transaction is immutable and permanently recorded, and forms the foundation of blockchain reliability. While layer 1 networks, such as Ethereum and Bitcoin, define finality through consensus mechanisms, layer 2 solutions introduce more nuanced interpretations due to their varied architectures and reliance on off-chain execution. The lack of

a universal definition of finality across L2 platforms presents challenges for benchmarking, particularly in measuring transaction latency, throughput, and security guarantees.

Unlike L1 chains, where finality is determined by the blockchain's native consensus rules, L2 finality depends on trust assumptions related to transaction commitment and validation. Different L2 architectures approach finality in distinct ways:

- Optimistic Rollups: Transactions are assumed valid unless challenged within a dispute window. While the official finality occurs when the challenge period ends, applications may consider earlier probabilistic finality based on historical challenge success rates.
- Zero-Knowledge Rollups: Transactions achieve near-instant finality once a cryptographic proof verifies the correctness of a rollup batch. However, verifying individual transactions within a batch depends on separate data availability mechanisms, introducing potential latency.
- State Channels and Plasmas: Finality depends on participants signing off on the latest state in state channels, while in plasmas, transaction validity relies on fraud-proof mechanisms where asset owners may need to actively monitor for disputes.

The choice of finality definition impacts performance metrics such as transaction latency and throughput, influencing how different L2s are compared. A simple benchmarking framework that assumes only L1 finality would significantly underestimate the practical usability of certain L2 solutions, while failing to capture their efficiency benefits. For example, an optimistic rollup with a week-long challenge period would not be able to compare to the finality of a zk-rollup, even if the batch is probabilistically certain to go through.

Beyond the theoretical definitions of finality, an essential consideration is the ability of participants to know when a transaction is final. In public L1 networks, finality is observable directly from the blockchain. In L2 systems, however, information asymmetry arises due to differences in data availability and trust assumptions. If transaction finality occurs but remains unknown to users until queried from an off-chain data provider, the perceived latency of the system increases, as users must wait for data retrieval to confirm their transaction status. Additionally, some applications rely on trusted sequencer confirmations, treating transactions as finalized long before official layer 1 settlement, which can create discrepancies in measuring the actual speed and security of layer 2s. These inconsistencies highlight the need for benchmarking methodologies to distinguish between absolute finality, which refers to irreversible settlement recorded on L1, and *practical finality*, where users or applications assume finality based on their trust in L2 mechanisms. This differentiation is crucial in evaluating the real-world performance of L2 architectures, as trust-based assumptions may significantly alter how finality is perceived and utilized. To create an unbiased benchmarking framework, metrics must define points of finality at multiple stages, including transaction acceptance, batch formation, challenge expiration, and L1 commitment in order to reflect how different L2 architectures optimize for speed versus security.

3.5 Current Ecosystem

The rapid expansion of layer 2 scaling solutions has been driven by the growing demand for high-throughput, low-cost blockchain transactions without compromising the security of layer 1 chains like Ethereum. As transaction fees surged on Ethereum, L2 systems emerged as the dominant strategy for improving scalability while retaining decentralization and interoperability. The expansion of state channels, sidechains, plasmas, and rollups has led to intense competition among L2 providers, each striving to position their architecture as the most efficient, cost-effective, and secure solution.

To gain an edge, many L2 projects market performance metrics such as transaction-per-second (TPS), time to finality, and gas savings, often emphasizing highly optimistic figures. This trend has mirrored previous benchmarking tactics seen at the L1 level, where networks selectively highlight throughput under ideal conditions rather than practical workloads. For example, rollups can claim to achieve thousands of transactions per second, but these numbers frequently rely on assumptions about trust models, optimal transaction batching, or favorable network conditions. Similarly, L2s using off-chain data availability models often advertise faster transaction speeds, yet their security and decentralization trade-offs remain underexplored in mainstream discussions. These issues highlight the need for a standardized benchmarking framework that objectively measures L2 performance across varying trust assumptions, transaction types, and realworld usage scenarios, ensuring transparent comparisons that cut through marketing-driven claims.

Among the diverse L2 designs, rollup-based solutions have emerged as

the dominant scaling approach, primarily due to their compatibility with Ethereum, strong security guarantees, and flexibility in execution models. On Jan 1, 2024, rollups had just under \$19 Billion total locked value (TLV). In December 2024, rollups peaked at just over \$55 Billion, a nearly 300% increase in under a year [11].

The two primary rollup types, optimistic and zero-knowledge, have shaped the L2 landscape, offering distinct trade-offs in latency, security, and cost efficiency. These two distinct rollup mechanisms were discussed in more detail back in Section 3.1.4. The recent rise of modular rollups, which separate execution, consensus, and data availability layers, further pushes rollups toward greater scalability and interoperability.

Due to their strong adoption across DeFi, gaming, and tokenization applications, rollups have set the standard for L2 scalability. Their evolving architectures, coupled with Ethereum's continuous improvements, ensure that rollups remain the most viable solution for blockchain expansion, reinforcing their dominance in the L2 space. As L2 competition intensifies, rollups will continue to define scalability standards, influencing benchmarks that measure performance, decentralization, and security across blockchain networks.

3.6 Motivation for BBSF L2 Expansion

The Blockchain Benchmarking Standardized Framework [14] was originally designed to provide a structured methodology for evaluating blockchain systems, ensuring fair performance comparisons across various architectures. From there, Blockbench-v3, a layer 1 BBSF-based benchmark, was developed and tested, as mentioned earlier. However, as layer 2 solutions have become the dominant scaling mechanism for Ethereum, the need for a tailored benchmarking system for L2 rollups has grown significantly. The unique characteristics of Ethereum-based L2s introduce complexities that standard L1 benchmarking methodologies fail to address. Moreover, the current ecosystem reveals an interesting dynamic between Ethereum-based L2s and newer, performance-focused L1 platforms. While Ethereum enjoys a large market cap and widespread usage, these benefits must be balanced against performance trade-offs, necessitating fair and comprehensive comparisons that extend beyond throughput and latency metrics.

One of the driving motivations for adapting the BBSF to L2 benchmarking is the increasing competition among rollups, where projects often present selective or self-serving performance metrics to market their solutions more favorably. Many L2 platforms report high throughput numbers without detailing transaction type, trust assumptions, data availability dependencies, or realistic network factors, leading to inconsistent or misleading comparisons. By extending the BBSF to cover L2-specific concerns such as finality variations and the impact of trust, this framework aims to establish standardized benchmarks that accurately reflect real-world performance while ensuring transparency.

Given that the majority of rollup-based L2 solutions operate on Ethereum, our benchmarking implementation is specifically designed to measure performance, security, and decentralization within Ethereum's layer 2 rollup ecosystem. This system will define precise measurement points for key L2 rollup components across the varying points of finality, and gain insight on more macro-level metrics such as decentralization and scalability. In doing so, our extended framework not only facilitates comparisons between Ethereum-based L2s and high-performance L1 alternatives, but also empowers application developers to make decisions that account for both technical performance and broader ecosystem factors.

4. Blockbench-L2

Blockbench-L2, our adaptation of the Blockchain Benchmarking Standardized Framework for layer 2 systems, is specifically designed for Ethereumbased L2 rollup solutions, which currently dominate the scaling ecosystem. This framework incorporates Web3-style workloads, similar to its L1 counterpart Blockbench-v3, capturing a diverse range of transaction types, execution complexities, and traffic patterns that are representative of realworld decentralized applications.

4.1 Workloads

Layer 2 transactions fundamentally mirror layer 1 transactions, as they are aggregated into a single batch before being finalized on the main chain. Despite this structural similarity, L2 solutions introduce their own distinct transaction processes, which are subsequently submitted to L1, creating subtle differences between L1 and L2 workloads. Given the widespread use of Decentralized Exchanges (DEXs), NFT Marketplaces, and NFT Minting, which remain integral components across both layers, these workloads will be retained from Blockbench-v3 for our L2 benchmarking framework to ensure consistency in evaluation.

However, the sports betting workload from Blockbench-v3, while relevant, represents a more niche transaction pattern and does not align as closely with the broader scope of L2 applications. To better reflect the demands and utility of L2 scalability, we replace this workload with one that captures the unique aspects of Web3 gaming, an area where L2 solutions provide meaningful performance improvements. Blockchain-based gaming naturally incorporates all three of our key benchmarking metrics, making it an ideal candidate for evaluation. Instead of developing an entire L2 game, we simulate the transaction workload, mimicking the functionalities such as in-game token earnings, purchases, and item exchanges. This approach ensures that the benchmarking process remains focused on performance analysis while reflecting real-world usage scenarios within the L2 ecosystem.

Therefore, our workloads for Blockbench-L2 are as follows:

- Simple Transfer: The introduction of layer 2s presents a new possibility for digital currencies to fully replace the traditional dollar. Since a stablecoin requires a lot more nuance and technicalities, we aim to replicate the functionality by unleashing a very high volume of wallet-to-wallet transfers of tokens. This workload will allow developers and users alike to gauge the performance capabilities of different L2s on the matter of high volume, simplistic transactions.
- Decentralized Exchange (DEX): The DEX workload will be nearly identical to the DEX workload outlined by our Blockbench-v3 implementation for L1s. Uniswap v3 and dYdX are two extremely popular layer 2 DEX implementations with many others close behind that leverage layer 2's small gas fees. Adding to this competitive land-scape, Uniswap Labs recently announced Unichain[1], an innovative layer 2 solution designed to further reduce transaction costs and improve execution speed, thereby expanding the range of options available to decentralized exchange developers. The DEX workload will

involve 2 tokens, a liquidity pool, and functions that allow the providing and removing of liquidity as well as swap operations.

- NFT Marketplace: The NFT Marketplace remains nearly identical to the Blockbench-v3 NFT Marketplace. NFTs are a big part of the social blockchain space, and Marketplaces are the single place to buy, sell, and trade these tokens. The workload will consist of a marketplace with functions that allow users to buy and list NFTs.
- NFT Minting: NFT Minting is another Blockbench-v3 imported workload that focuses on an overlooked aspect of the NFT space and the strain that it has on the blockchain network. NFT Minting consists on a simple contract that mints 10,000 NFTs (the standard) all at once, which can be stressful for a layer 1. Some layer 2 networks like zkSync have NFT Minting built in as a key supported feature.
- Web3 Gaming: The Gaming workload is a new workload from the Blockbench-v3 implementation, taking the place of the Sports Betting contract. L2 gaming has become a large part of the social blockchain ecosystem, where games utilize these chains to handle ingame transactions and currency. Our workload will not involve the use of an actual game, but instead mimic the functionalities that a game would have with a blockchain. This includes providing players with tokens, allowing players to purchase tokens, allowing players to use tokens to purchase items (NFTs), and perhaps allowing players to trade or sell their items.

4.2 Metrics

Evaluating the performance of layer 2s requires a structured approach to benchmarking that accounts for finality, execution speed, and cost efficiency. The core micrometrics of throughput, latency, and gas consumption serve as fundamental indicators of how efficiently an L2 system processes transactions while maintaining security guarantees.

Throughput, measured in transactions per second (TPS), reflects the overall capacity of an L2 system to handle transaction volume under various conditions. Higher throughput suggests improved scalability, but the definition of what constitutes a "processed transaction" varies depending on the finality model. Latency measures the time elapsed between transaction initiation and confirmation, impacting user experience and application responsiveness. Different L2 architectures introduce distinct latency considerations depending on whether transactions are immediately acknowledged, internally processed, or fully settled on layer 1. Gas consumption, recorded as gas cost per transaction, provides insight into the efficiency of an L2 system in reducing fees relative to Ethereum's base layer.

Since finality assumptions significantly influence throughput and latency measurements, this framework categorizes finality into three trustbased models: Full Trust, Partial Trust, and No Trust Finality. By structuring throughput and latency measurements across three different finality models, this benchmarking framework provides transparent insights into L2 performance, allowing comparisons that account for trust assumptions, security guarantees, and decentralization trade-offs. Our three finality models are as follows:

- Full Trust Finality measures throughput and latency from the moment a transaction is received by the L2 sequencer, assuming users trust the system to process and eventually settle their transaction correctly. This approach results in the highest perceived transactions per second (TPS), as it counts all transactions entering the sequencer queue without waiting for additional validation. Latency under full trust finality is minimal, often sub-millisecond to a few seconds, since the sequencer provides immediate acknowledgment. This model benefits applications requiring rapid execution, such as high-frequency trading and gaming, but relies on centralized infrastructure, making transactions vulnerable to censorship or manipulation.
- Partial Trust Finality considers transactions final only once they have been fully processed within the L2 rollup and are scheduled for inclusion in a layer 1 batch. While this reduces TPS compared to full trust finality, it ensures that only validated transactions are counted, offering a more balanced metric that excludes unprocessed transactions. Latency here increases to seconds or minutes, depending on batching mechanisms, proof generation, and rollup architecture. Optimistic rollups tend to exhibit shorter latency, while zk-rollups experience longer delays due to computational proof generation. This model provides a realistic performance assessment without incorporating layer 1 confirmation delays.
- **Trustless Finality** defines throughput and latency based on full layer 1 settlement, meaning a transaction is only considered finalized once it is permanently recorded on Ethereum. Since rollups batch transactions before posting them to layer 1, this metric reports the

lowest TPS, as many transactions remain pending before final inclusion. Latency under this model is highly variable, with zk-rollups finalizing within minutes, whereas optimistic rollups may take hours or days due to challenge periods. While this ensures absolute security, it is impractical for real-time applications, as most users and developers operate under earlier finality assumptions rather than waiting for full L1 settlement.

Outside of these micro-metrics, we have various macro-metrics that can be observed. These system-level measurements capture broader architectural and performance trade-offs that shape the real-world viability of layer 2 solutions. There are three primary macro-metrics that Blockbench-L2 aims to quantify and evaluate: scalability, decentralization, and modularity.

In a layer 1 context, scalability is measured by varying the number of full nodes and observing how throughput and latency respond. For layer 2 systems, the natural equivalent is the size and configuration of the sequencer network. To assess scalability, we could deploy a varying number of sequencer instances and measure transaction throughput and latency under a certain workload. A truly scalable L2 design will exhibit close to linear improvements in TPS as sequencer capacity grows, with latency remaining within a certain range. By plotting throughput versus sequencer node count, we can identify the point at which adding sequencers gives diminishing returns. This differs from layer 1 measures of scalability, where we expect throughput to decline as the node network grows due to more consensus overhead.

Decentralization at layer 2 can be deconstructed into multiple dimen-

sions: sequencer type, data availability method, and operator diversity. Sequencers can be classified on a spectrum from fully centralized (single operator) to fully permissionless (multi-party consensus). For data availability, we can distinguish between on-chain storage, off-chain data blobs with fraud proofs, and specialized DA networks, assigning each some qualitative score. Finally, we can measure operator diversity by counting distinct validator or node operators, assuming this data is publicly available. This specific factor for decentralization level is notably fully independent from the design of the system and would therefore be measured outside of the context of the benchmarking setup. Together, these metrics can produce a decentralization index that highlights where an L2 design leans toward efficiency at the expense of trust, or vice versa.

A modularity metric gauges the degree to which an L2 architecture separates execution, settlement, consensus, and data availability layers into interchangeable components. We could define a modularity score based on the number of independently deployable modules and the ease with which a component can be swapped without changes in the core protocol. High modularity not only accelerates innovation by allowing best-in-class solutions for each layer, but also simplifies benchmarking cross-layer interactions. By comparing modularity scores across L2 designs, we can assess how readily each network can evolve or integrate new technologies without sacrificing its security or performance guarantees.

4.3 Driver

The Blockbench-L2 driver extends the original BBSF driver design to meet the additional demands of benchmarking layer 2 solutions. Unlike layer 1 systems, where finality is relatively uniform, layer 2 introduces multiple variations of finality as detailed in Section 4.2. As a result, the Blockbench-L2 driver must be able to track and differentiate between these three types of finality for every workload it executes. This adds complexity compared to the original BBSF design, requiring more detailed instrumentation and more granular event tracking throughout the benchmarking process.

In addition, because our focused layer 2 solutions ultimately rely on Ethereum for settlement and finality guarantees, the driver must also interact with the Ethereum L1 to accurately measure trustless finality. This means the driver not only monitors the layer 2 system under test, but also observes relevant contract events, transaction finalization, and settlement proofs on Ethereum. These cross-layer interactions are necessary to provide an accurate, end-to-end picture of transaction finality from a trustless perspective, ensuring that the benchmarking results reflect the true guarantees offered by the system. As such, while the Blockbench-L2 Driver is built upon the BBSF design principles, it evolves them to address the more intricate nature of layer 2 architecture.

4.4 Preliminary Results

The development and testing of Blockbench-L2 is still very much in progress. The real contribution of this work is the framework provided to fairly and accurately benchmark and compare layer 2 scaling solutions. Work is actively being done to develop the driver to be able to gather all of our micro-metrics, and to further be compatible with scaling efforts for our scalability macro-metric. To provide a proof of concept of our framework, our team has been able to develop and run the Simple Transfer workload on the zk-rollup based layer 2 called zkSync, and measured using partial-trust finality.

We ran the following workloads on the zkSync release *core-v26.8.3*. The machine used is the same as the once described in Section 2.2.1. For these benchmarks, we configured an environment of 4 sequencers and a single Ethereum node. Our current testing environment does not come equipped with GPUs, so we are unable to run the zkSync prover there. Without the prover, we are only able to measure the raw throughput of the sequencer, and our measurements will not reflect the time it would take for proof generation. As mentioned before, this framework is still a work in progress, and these results are a preliminary proof-of-concept.

As seen in Figure 4.1, we were able to collect metrics on zkSync using the partial trust definition of finality. As mentioned in Section 4.2, we measure partial trust finality at the time in which the sequencer confirms the transaction is batched and ready to be submitted to the layer 1. Results were collected over five individual runs of each workload and averaged. Due to the benchmark not including the prover, it is to be noted that these numbers do not reflect the overall system as a whole. Future works will address this issue and include more holistic benchmarking results. However, as a proof-of-concept benchmarking trial, these results show great promise for the Blockbench-L2 framework. There is a clear variation be-



Figure 4.1: zkSync Benchmark Results, Partial Trust

tween throughput and latency results between the various workloads, and it will be interesting to view how these results change with varying environmental configurations. It will also be nice to view these results in context with the calculated macro-metrics once the framework is finalized and processes for determining these are well defined.

5. Conclusion and Future Work

5.1 Summary of Contributions

The blockchain industry has long struggled with the challenge of unreliable performance evaluation. In the absence of standardized methodologies, various performance claims often lack the transparency and consistency needed for independent verification and meaningful comparisons. Previous work establishing the Blockchain Benchmarking Standardized Framework (BBSF)[14] and implementing Blockbench-v3 successfully addressed these issues for layer 1 systems by providing a comprehensive, standardized benchmark that detailed every aspect of workload execution, from transaction specifications to micro and macro-metric definitions. Building on this foundation, our work extends these principles to the more complex layer 2 landscape through the development of Blockbench-L2.

Blockbench-L2 not only builds on the framework established in layer 1 benchmarking but also introduces metrics that reflect the complex characteristics of layer 2 systems. These include unique measures of throughput and latency under different finality assumptions, such as full trust, partial trust, and trustless models. Additionally, our driver component standardizes workload execution across varying runtime environments, providing insights into more macro-level such as scalability, decentralization, and modularity. By incorporating a diverse set of workloads such as simple transfers, decentralized exchange functionalities, NFT marketplaces and minting processes, as well as Web3 gaming simulations, our framework goes beyond traditional benchmarks, grounding performance evaluation in scenarios that mirror real-world applications.

The implications of this work for the blockchain space are far-reaching. A standardized layer 2 benchmarking framework like Blockbench-L2 paves the way for greater industry transparency, ensuring that performance claims are not only comparable but also reproducible by independent third parties. This advancement fosters trust among developers, researchers, and enterprise users alike, allowing them to make more informed decisions when selecting and optimizing layer 2 solutions. By delivering a fair, adaptable, and transparent benchmarking framework, our work contributes to setting a new standard in blockchain performance evaluation, which is crucial for driving innovation and scaling decentralized systems in a rapidly evolving digital landscape.

5.2 Limitations

In our ongoing work, we continue to refine our benchmarking framework by addressing several active areas for improvement. One key area involves our experimental hardware setup. Currently, our system does not include a GPU-enabled machine, which is necessary for running GPU-based provers on zk-rollup systems, such as zkSync. We shall be moving to a system with those resources so that we can fully support the benchmarking of GPUdependent components and provide a comprehensive evaluation of rollup performance.

Another aspect involves the configuration and deployment of certain

layer 2 systems. The rapidly evolving nature of the layer 2 ecosystem means that platforms and protocols frequently update their core configurations, presenting an ongoing challenge in maintaining accurate and up-to-date benchmarks. We view this problem as an opportunity to drive further standardization in the industry. Our long-term vision is for blockchain companies to benchmark their systems openly and consistently using our framework, fostering transparent documentation and facilitating easier, community-validated comparisons across both layer 2 solutions and emerging performance-focused layer 1 platforms.

5.3 Future Work

As previously stated, this work is currently in progress with my team at Lehigh University. The current team, as well as new students to come in the future, will continue this layer 2 work and likely expand to newer areas in the greater blockchain space.

5.3.1 Short Term

In the near future, I expect that the layer 2 extension of the BBSF will be finalized and published. Along with this publication, I expect we will have complete benchmarks available from Blockbench-L2, ranging across a variety of Ethereum-based layer 2 rollups. These benchmarks will likely contain valuable performance data for these systems, and make noticeable the large difference trust assumptions play when gathering key metrics.

5.3.2 Long Term

In the long term, I hope to see many layer 2 systems successfully benchmarked through Blockbench-L2, with widespread adoption of this standardized framework within the blockchain space. Future work should also extend beyond Ethereum-based solutions. For instance, exploring benchmarking methodologies for emerging layer 3 and layer n+1 protocols would further refine our understanding of scalability and interoperability across next-generation blockchain infrastructures. Bitcoin-based L2 solutions such as GOAT Network[8] would also be interesting to explore. Such extensions would prepare the framework to handle additional abstractions of consensus, execution, and data availability beyond what is offered by current layer 2 solutions.

Additionally, there is great potential for adapting the framework to other architectures, including layerZero[16] and specific bridging mechanisms. With the introduction of platforms such as UniChain[1], as well as the growing interest in utilizing Bitcoin and other layer 1 blockchains as bases for layer 2 solutions, a comprehensive benchmarking suite that spans these layers becomes essential. These directions promise to facilitate fair comparisons across a diverse ecosystem, ultimately aiding application developers and stakeholders in making informed decisions about the trade-offs between security, scalability, and performance.

Bibliography

 H. Adams, M. Toda, A. Karys, X. Wan, D. Gretzke, E. Zhong, Z. Wong, D. Marzec, R. Miller, Hasu, K. Floersch, and D. Robinson. Unichain: An optimistic rollup optimized for efficient markets. Web document, October 2024.

https://docs.unichain.org/whitepaper.pdf.

[2] M. Bedawala. Monolithic vs. modular blockchain. Web document, 2024.

https://usa.visa.com/solutions/crypto/ monolithic-vs-modular-blockchain.html.

- [3] Ethereum.org. State channels. Web document, 2023. https://ethereum.org/en/developers/docs/scaling/ state-channels/.
- [4] Ethereum.org. Validium. Web document, 2023. https://ethereum. org/developers/docs/scaling/validium.
- [5] Ethereum.org. Danksharding. Web document, 2024. https:// ethereum.org/roadmap/danksharding.
- [6] Ethereum.org. Sidechains. Web document, 2024. https://ethereum. org/developers/docs/scaling/sidechains.
- [7] Ethereum.org. Zero-knowledge rollups. Web document, 2024. https: //ethereum.org/en/developers/docs/scaling/zk-rollups/.
- [8] GOAT Network Research Group. GOAT Network: Natively Extending Bitcoin using Entangled Rollup. Web document, 2025. https://www.

goat.network/whitepaper.

- [9] J. Gray and A. Reuter. Transaction Processing: Concepts and Techniques. Morgan Kaufmann, 1993.
- [10] H. Kalodner, S. Goldfeder, X. Chen, S. Weinberg, and E. Felten. Arbitrum: Scalable, private smart contracts. In *Proc. 27th USENIX Security Symposium*, page 1353–1370, 2018.
- [11] L2Beat.com. Layer 2 Total Value Secured. Website, 2025. https: //l2beat.com/scaling/tvs?tab=rollups.
- [12] J. Poon and V. Buterin. Plasma: Scalable autonomous smart contracts. Web document, 2017. https://plasma.io.
- [13] J. Poon and T. Dryja. The Bitcoin Lightning Network, scalable offchain instant payments. Web document, 2016. https://lightning. network/lightning-network-paper.pdf.
- [14] K. Ren, J. F. V. Buskirk, Z. Y. Ang, S. Hou, N. R. Cable, M. Monares, H. F. Korth, and D. Loghin. BBSF: blockchain benchmarking standardized framework. In VDBS '23: Proceedings of the 1st Workshop on Verifiable Database Systems, New York, NY, USA, 2023. Association for Computing Machinery.
- [15] A. Tzionis. A "literature review" on rollups and validium. Web document, 2023. https://ethresear.ch/t/ a-literature-review-on-rollups-and-validium/16370/1.
- [16] R. Zarick, B. Pellegrino, I. Zhang, T. Kim, and C. Banister. Layerzero: An omnichain interoperability protocol. Whitepaper, 2024. https://layerzero.network/publications/LayerZero_ Whitepaper_V2.0.pdf.
- [17] ZKM Research. Entangled rollups: Multi-chain interoperability

without bridges. Web document, 2024.

https://whitepaper.zkm.io/entangled_rollup_light_paper.

pdf.

[18] zkSync. Cryptography used, 2022.

Vita

Nathaniel Rex Cable grew up in the Wyoming Valley region of Pennsylvania with his parents Philip and Melissa Cable. He attended Lehigh University from 2020 to 2024 where he studied Computer Science in the P.C. Rossin College of Engineering and Applied Science. Following his graduation in 2024 with a B.S. in Computer Science, Cable returned to Lehigh University for his Masters in Computer Science, continuing his research on Blockchain Benchmarking. Cable has been a part of the Blockchain Projects research group, a subset of the larger Scalable Systems and Software research group at Lehigh University, under Professor Henry F. Korth. Following graduate school, Cable will be joining Oracle as a Member of Technical Staff on their blockchain platform team.