Nillion: Humanity's First Blind Computer

Tristan Litre[1], Miguel de Vega[1], and David Butler[1]

[1]Nillion

January 26, 2025

Abstract

Nillion is Humanity's First **Blind Computer**—a whole new category of decentralized network designed for AI and the future of the Internet. **Nillion makes new applications possible by providing storage and computation on high value, encrypted data without ever seeing it**. Whether a user, an app, or an enterprise, your data stays yours—always.

The Nillion Network is split into two parts: the Coordination Layer and the Petnet (Privacy Enhancing Technology Network). The Petnet is made up of separate Clusters of nodes that offer different technologies and products, with the long term goal of allowing anyone to come and spin up their own Cluster to fit their application's needs or to compete for users and be rewarded. The Coordination Layer is a chain that manages shared resources, such as rewards, cryptoeconomic stake, and governance, as well as enabling inter-cluster coordination.

Nillion uses a variety of different technologies to advance the goal of making it easy for developers to build with cutting edge cryptography without being experts. In this document we provide an overview of these technologies and concepts and their role in the network.

1 nilVM

nilVM is a virtual machine that allows developers to easily create programs that leverage Multiparty Computation (MPC) based on Linear Secret Sharing Schemes (LSSS). Programs are written in an open source language called Nada, which is a Python-based DSL designed to feel like an easy-to-pickup library for Python developers, and are developed with the Nillion SDK. Nada has helper libraries to make building programs easier like nada-ai, which offers a similar interface to PyTorch and scikit-learn, nada-numpy, and nada-data.

Developers can build nilVM into their applications using clients in Python, Typescript, or the CLI. They can also leverage the Storage APIs, which make it easy to quickly bootstrap projects that securely store and retrieve data using the Nillion Network.

The following is a list of links to read more about nilVM:

- Building with nilVM.
- Nada Language.

These are some examples of programs using nilVM:

- A federated learning example.
- Many community examples are available in the awesome-nillion repository.
- Interactive demo use cases are collected here.

The main builders working with nilVM are the following:

- Skillful AI has implemented private Retrieval Augmented Generation (RAG) for AI models and agents using nilVM, allowing users to upload sensitive documents and leak the minimum possible amount of information at inference time.
- Rainfall has built a self-sovereign AI platform focused on collective social intelligence generation while protecting user privacy. nilVM is used to ensure the storage of user weights while enabling private weight aggregation and inference of AI models.
- Kayra has built a dark pool program using nilVM as the backbone for order matching, providing complete execution privacy.
- Mailchain has implemented a feature that would allow users to choose to store emails in nilVM, benefiting from the secure decentralized private storage and built-in granular access control mechanisms of the network.
- Kagami has built out a trading policy engine in nilVM, enabling the storage and execution of sensitive financial trading policies in fully encrypted form.
- one.id is building a platform designed to store sensitive personal data, such as passwords in a personal vault in nilVM, creating a truly self-sovereign and private decentralized storage solution for their users.
- ChooseK is building out an order book as a service platform using nilVM to enable developers to build out private DeFi products focused on encrypted trade execution to prevent MEV.
- BattleshipX is building an interactive P2P battleship game in nilVM, enabling 1v1 gameplay of Battleship while protecting users' game strategies and positions from each other.

The following is a list of some basic facts about growth:

- The Nillion SDK has been downloaded externally 961 times in the last 2 months
- There are 104 unique contributors across Nillion public repos (excluding forks)

2 nilDB

nilDB is an encrypted, distributed NoSQL database that allows users and developers to unlock the power of sensitive data while preserving privacy. It functions by distributing trust for data storage across a cluster of nodes, with each node running an instance of the database that has a secret share of the data. Data owners can grant consent to others to run SQL-like queries over the data, including calculations and privacy-preserving aggregation, with the goal of making all of the data stored in Nillion more valuable by making it composable.

The initial cohort of applications building on nilDB are focused on Data Marketplaces and Insights, unlocking the huge amounts of data users generate daily through connected devices that sit on centralized servers owned by private web2 companies. This cohort of web3 and web2 companies are building on Nillion to unlock the utility of this personal data by not only storing it securely, but also allowing it to be used and analyzed by buyers that need the data (e.g. AI models for training) or users that want aggregated analytics. Nillion generates usage as the amount of high value data stored in the network grows and compute queries are run.

This is the main link to read more about Nillion's architecture.

The following is a list of the main builders working with nilDB:

- ZAP Labs: Founded by former Starkware developers and a co-founder of Kakarot zkEVM, ZAP is building decentralized infrastructure for a new data-as-an-asset economy by leveraging zkTLS. They aim to store user data in 'Data Pools' and provide analytics from users' data to paying third party customers. Nillion will serve as ZAP's storage and computation backend, having chosen it over their original Filecoin-based design because of Nillion's superior security and programmability.

- Fulcra: Founded by the Founder of HUMAN, the world's most successful bot defense platform (one of TIME Magazine's 100 Most Influential Companies 2023 reaching 3 billion devices). Fulcra creates a data vault across a range of smart devices to generate new insights. Fulcra will use the Nillion Network to establish new collective insights of people's data in a data-to-earn marketplace.

- Reclaim Protocol: Reclaim is a zkTLS infrastructure platform enabling users to trustlessly prove identity and reputation through off-chain platforms. For instance, users could login into their bank and prove their bank balance or login into Uber and prove that they've taken more than a certain number of rides this year. The Nillion Network is used as a storage and processing platform for the proofs generated. It enables proofs to be stored in a decentralized network, offering private decentralized storage and composability of data while facilitating secure privacy-preserving data processing across user's verified data, unlocking a new paradigm of use cases.

- Healthblocks: HealthBlocks enables users to collect data across different sources and devices into a single place all while ensuring ownership and control remain firmly in the users' hands. The collected data can then be used in Nillion to generate insights to third parties without leaking anything about the specific details of the data.

- Humanity: Users have gathered significant data and social graphs over years of extensive usage. Although they have access to the data, they can't own or monetize it. Humanity is developing a digital identity and analytics platform that integrates data from social, health, and entertainment platforms to establish proof of humanity while ensuring privacy and secure identity verification. Furthermore, with the collected data, it enables cross-platform behavior analysis, macro-trend identification, and influence mapping. Nillion is the infrastructure platform to securely store and process user data in a fully privacy-preserving manner.

- Dwinity: Dwinity's mission is to empower data owners and establish a thriving data economy by eliminating informational asymmetries that have long hindered fair and transparent data usage. To achieve this, Dwinity is building a comprehensive, data-ownership-driven ecosystem designed to revolutionize the storage, analysis, exchange, and commercialization of sensitive personal data. They use Nillion to store and process user data in a fully privacy-preserving manner.

- Verida: Verida is developing a personalized AI platform, facilitating users to chat with a personalized chatbot trained on a range of sensitive documents, such as their messages. nilDB will be used to store the personal user information, and AIVM to run inference in encrypted form.

3 nilAI

nilAI is Nillion's suite of AI-optimized privacy technologies. The initial products that make up nilAI are **AIVM**, **nada-AI**, and **nilTEE** – as well as the intersection of those technologies.

AIVM is a secure AI inference platform for Deep Learning models based on Multiparty Computation (MPC). It enables private model inference and custom model deployment while maintaining data confidentiality throughout the entire computation process. Individual nodes are blind to user prompts and the results of the model, while the model can be kept private from users if desired. AIVM speeds up inference using Discrete Wavelet Transform (DWT), leveraging [3][1], a technique coauthored between Nillion and a researcher from Meta's AI Research team, and is built on top of Meta's CrypTen secure machine learning framework.

nada-AI is a library for nilVM that makes it easy to use nada for AI purposes. Users can run small models with a PyTorch-like interface. The current library supports Neural Networks (NN), Convolutional Neural Networks (CNN), and Linear Regression models among others. The library also offers easy-to-use Colab notebooks for developers to bootstrap their own uses for the models.

nilTEE is Nillion's offering leveraging Trusted Execution Environments (TEEs). The initial utility in the network is informed by requests from application builders, who want to be able to run larger LLMs performantly inside of a TEE. The recommendation of Nillion is not to rely on TEEs for long term storage of any user information, but rather

[1]See also related work [1, 2] coauthored between Nillion and other institutions.

limit exposure of the trust model only to inference time. Active research is also being developed in leveraging both AIVM and nilTEE in a split inference setup to harden the security and performance of both technologies.

Main links to learn more:

- Nillion AIVM: A cutting-edge framework designed for privacy-preserving inference using advanced cryptographic protocols.
- Nada AI: A Python library designed for performing ML/AI on top of Nada DSL and the Nillion Network.
- A new wave of privacy-preserving large language models: An article on Nillion's pioneering work on the use of the DWT to accelerate private inference on LLMs [3] done in collaboration with Meta and the University of California Irvine.
- Personalized AI: An educational video featuring the relationship between personalized AI and privacy.

4 nilChain

nilChain is the Nillion Network's Coordination Layer, managing rewards, payments, cryptoeconomic security, and inter-cluster coordination for the network. It is built using the Cosmos SDK and is IBC-compatible for interoperability. nilChain is purely for coordination at this stage and has no smart contract execution platform. It is accessible using the Keplr or Leap wallets and has been running on testnet since May 2024.

The following are the main links to learn more about nilChain:

- nilchain: The Coordination Layer for the Nillion Network.
- Nillion documentation about network configuration.

5 Orchestration Layer Vision

The Orchestration Layer is Nillion's vision for the future of Privacy Enhancing Technologies (PETs). Beginning with nilVM's compiler, which abstracts away many different MPC protocols for users, the Orchestration Layer seeks to make it simple to leverage PETs to build new applications without being an expert in cryptography. This vision brings together all the pieces mentioned above, as well as Nillion's unreleased products that leverage Trusted Execution Environments (TEEs) for AI inference and Fully Homomorphic Encryption (FHE) for arbitrary computation.

The current focus for Orchestration is unifying the different PETs and products available on the Nillion Network, allowing builders to interoperate easily between them. One of the key research focuses here is orchestrating together the AIVM with LLM inference

in a TEE using Split Inference, hardening the TEE even in the case of attacks. Another workstream is focused on integrating private Retrieval Augmented Generation (RAG) with nilDB, allowing for AI inference using the sensitive data as context.

Main links to learn more about Orchestration:

- The Orchestration Layer: Nillion's Vision for a Better Internet: A news article articulating Nillion's vision for the Orchestration Layer.
- The Orchestration Layer: An educational video featuring Nillion's Orchestration Layer.

6 Blockchain Partner Integrations

The Nillion Network will make Blind Computation accessible across web3 and has partnered with an initial set of blockchains as design partners to integrate with, including Arbitrum, Near, Aptos, Sei, and more. Although many builders are building using Nillion as a backend directly, Nillion's blockchain integrations make it possible for smart contracts and users on other chains to leverage blind computations directly from their chains in a seamless and native way – using the partner chain's native gas token for fees and without downloading a new wallet. This enables dApps across ecosystems to build features seamlessly into the whitespace Nillion enables.

The main link to learn more about the blockchain integration architecture is the Chain Integration Design Document.

The following is a list of some of the partner blockchain announcements:

- The Block: Nillion integrates with NEAR Protocol, enhancing privacy tools for developers.
- The Block: 'Blind computing' platform Nillion integrates with Arbitrum to enhance privacy on Ethereum.
- Cointelegraph: Aptos network integrates Nillion for privacy-preserving apps.

7 Nucleus Builders Program and Ecosystem

Nucleus is the Nillion ecosystem support program for creators and visionaries building on the Nillion Network. Nucleus offers ongoing, flexible support designed to meet the unique needs of builders at every stage of their journey building across different verticals and horizontals. Whether they are a company building in traditional Web2, or a crypto start up, Nucleus equips them with the tools, mentorship, and resources to succeed in adopting Nillion's novel technology.

Since its inception, Nucleus has attracted close to 50 developers, start ups, and infrastructure platforms across more than 10 verticals in building out their solutions on the