Gonka: Designing a Compute-Native Decentralized Economy

2025-07-31

Gonka is a decentralized AI infrastructure designed to optimize computational power specifically for AI model training and inference, offering a competitive alternative to traditional centralized cloud providers. Centralized systems are often expensive, monopolistic, and carry risks of censorship, while existing decentralized networks frequently waste resources on non-productive tasks, such as network security.

We introduce an innovative consensus mechanism that ensures nearly 100% of computational resources are used for meaningful AI tasks, maximizing efficiency and minimizing operational costs.

The system features key roles:

- Developers build and deploy AI applications using the network's distributed power.
- Hosts contribute computational resources to the network, receiving rewards based on the amount and quality of resources they provide.

This collaboration allows Gonka network to offer AI services at significantly lower prices, making advanced AI technology more accessible to a broader audience.

This document explains the economic architecture that drives Gonka Network: how incentives are aligned, how resources are priced, and how the gonka (GNK) coin ensures sustainable growth.

It outlines the mechanisms that govern rewards, fees, and coin issuance, with a focus on predictable Developer costs and long-term fairness for Hosts. Each component is designed to maximize real utility while minimizing complexity and unnecessary overhead.

In particular, this paper details:

1. Defines the economic architecture of the network, including the GNK coin supply, its issuance schedule, and how incentive mechanisms align Hosts' compute contributions with sustainable network growth.
2. Introduces a vesting-based reward distribution system that gradually unlocks both Work Coins from Developers and Epoch-minted Reward Coins, ensuring long-term commitment and reducing short-term speculation.
3. Proposes a collateral-backed influence model, where Hosts must lock GNK coins to unlock full governance power and higher reward weight, creating a direct link between trust, contribution, and influence.
4. Presents a per-model, EIP-1559-inspired dynamic pricing system for inference, enabling cost predictability for Developers and efficient resource allocation based on network utilization.
5. Establishes a decentralized liquidity and community pool mechanism, governed by Hosts, to support early coin liquidity, bootstrap adoption, and stabilize the coin economy during network infancy.
6. Details how Hosts are compensated for both inference and training tasks via a dual reward structure, earning Work Coins from Developers and Epoch-minted Reward Coins.
7. Introduces the Decentralized AI Training Fund, which allocates 20% of inference revenue toward open-source LLM training, ensuring the network continues to produce cutting-edge models collaboratively and transparently.
8. Identifies key risk factors, including market volatility, regulation, and competitive pressures.

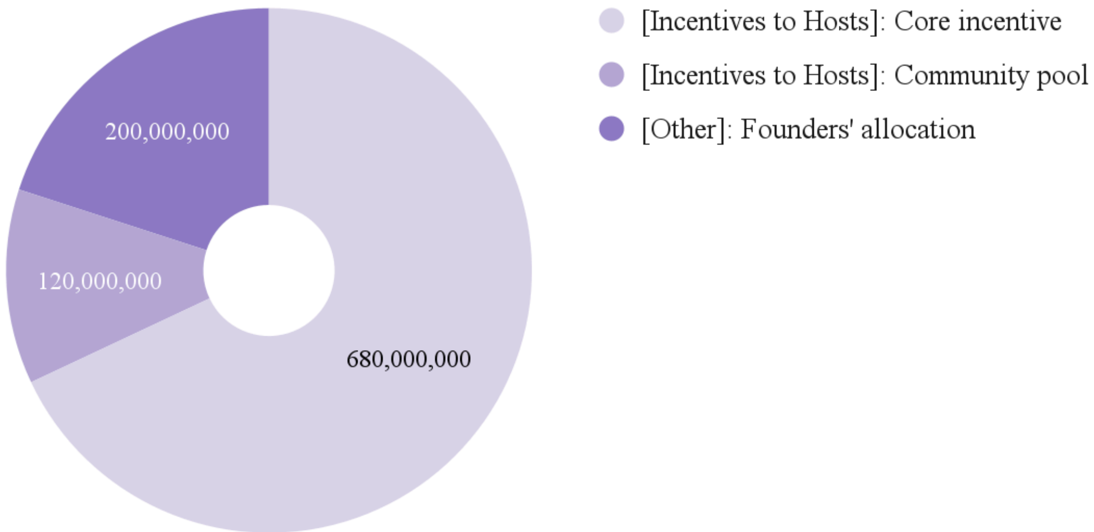
Contents

1. Tokenomics

The total supply of gonka (GNK) coins is fixed at 1 billion, allocated to incentivize Hosts, support network development, and ensure fair compensation for Hosts. Below is the detailed breakdown of the allocation. The distribution is divided into two primary categories:

1. **Incentives to Hosts** (800 million gonka (GNK) coins, 80% of total supply).
2. **Other: Founders' allocation** (200 million gonka (GNK) coins, 20% of total supply) is a portion reserved for the founding team as an allocation recognizing their ownership and long-term commitment to the project.

Distribution of gonka (GNK) coins (in Millions). Total supply: 1B



Incentives to Hosts consist of the following:

1. **The Core incentive for Hosts** (680 million gonka (GNK) coins, 68% of the total supply) is designed to reward Hosts for contributing computational power to the decentralized network. Under the Bitcoin-style reward System, a fixed amount of GNK is minted per epoch (initially 323,000 GNK) and distributed proportionally to each Host's Proof of Compute (PoC) weight. As the network grows and more GPUs compete for the same fixed reward, the GNK earned per GPU decreases, introducing scarcity-driven value dynamics similar to Bitcoin mining. Long-term value for Hosts comes from increasing demand for inference services, fixed emission scarcity, and rising GNK market value as mining becomes more competitive.
   The reward schedule follows a predictable exponential decay, reducing epoch rewards gradually, halving rewards approximately every 4 years.
2. **Community Pool (120 million coins).** A dedicated Community Pool is controlled exclusively by the Hosts, and its use is determined through decentralized governance via voting (see "Gonka:White Paper"). While the final allocation strategy remains entirely up to the Hosts' collective decisions, it is expected that part or all of the community pool will be allocated to strengthen early Host liquidity support.
   Host liquidity pool is designed to provide early liquidity for Hosts before gonka (GNK) coin is adopted by exchanges and P2P trading volume reaches levels comparable to similar projects. This pool will be distributed through a preprogrammed sale mechanism that allows Developers and coin holders to buy gonka (GNK) coins at any time.

Buyer can transfer USDT, Ethereum, or Bitcoin to a designated account on the network, automatically minting gonka (GNK) coins via the native bridge. Meanwhile, Hosts will be able to send their mined gonka

(GNK) coins to a dedicated smart contract on Gonka network and receive USDT, Ethereum, or Bitcoin from the pool at the latest achieved price, ensuring early liquidity.

2. Reward System of Gonka Network

2.1. Bitcoin-Style Reward System

Gonka Network introduces a Bitcoin-inspired fixed reward system, where a predetermined amount of Reward Coins is minted per epoch and distributed based on each Host's Proof of Compute (PoC) weight.

Key characteristics of the model:

- **Epoch-minted Reward Coins:** A base amount of 323,000 GNK per epoch ("Reward Coins") is minted and distributed across all active Hosts. This figure is subject to exponential decay (-0.000475 per epoch, equivalent to halving approximately every 1,460 epochs or 4 years), following the formula:

  current_epoch_reward = initial_reward × exp(decay_rate × epochs_since_genesis)

- **Scarcity-Driven Value:** As more GPUs join the network, each GPU earns fewer GNK per epoch. The increased mining cost per GNK creates intrinsic value and introduces a positive feedback loop, similar to Bitcoin:

  More Hosts → fewer GNK per GPU → higher scarcity → potential price support and growth.

- **Work Coins:** Fees paid by developers for actual inference work are also distributed based on the computational work completed.

- **Utilization bonuses** for MLNodes serving high-demand models, and model coverage incentives for Hosts supporting all governance-approved models.

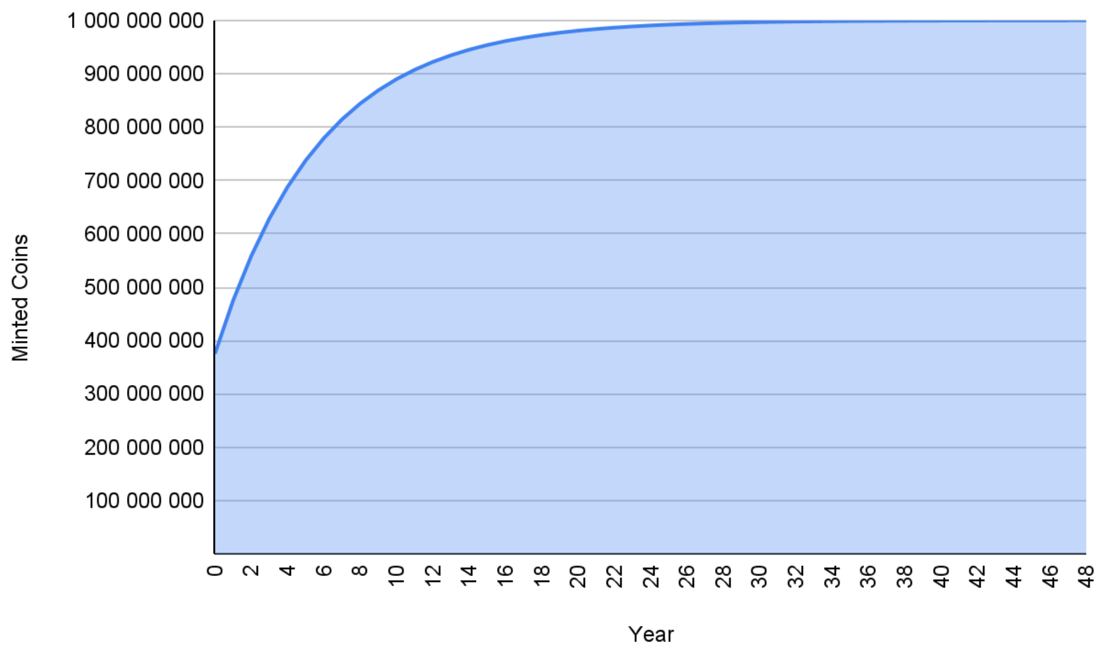


Chart 1. Cumulative Minted Coins vs. Year

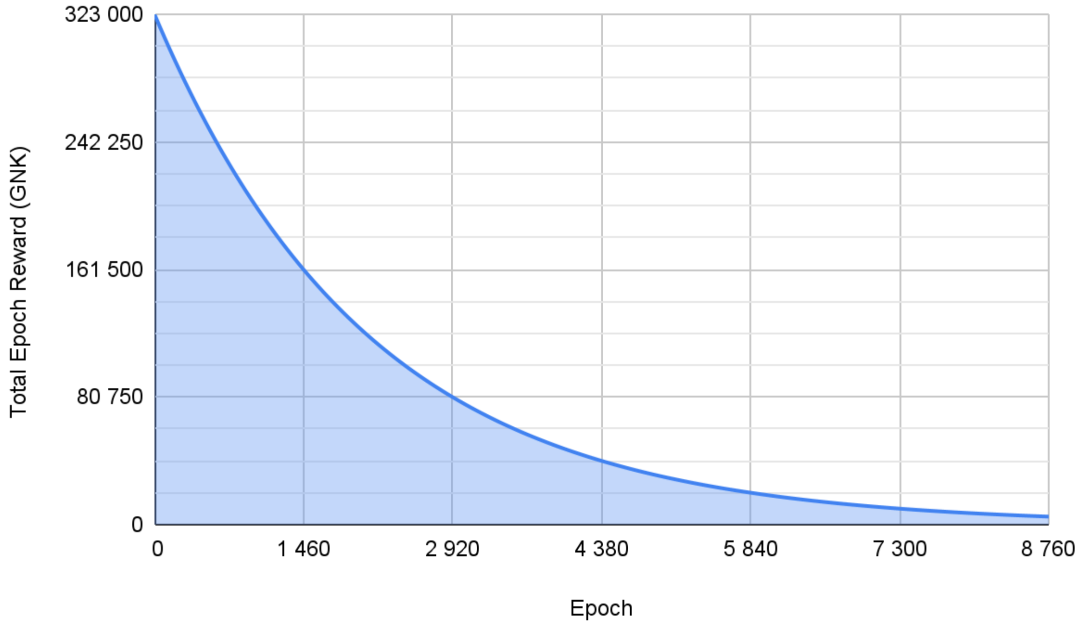


Chart 2. Total Epoch Reward (GNK) vs. Epoch

2.2. Reward Vesting

All rewards distributed by the Gonka Network, including both Epoch-minted Reward Coins and usage-based Work Coins, are subject to an intelligent vesting schedule managed by the dedicated vesting system:

- **Personalized Scheduling:** The network maintains a personalized release schedule for each participant, tracking their rewards on a daily basis to ensure fair and predictable distributions.
- **Efficient Processing:** When participants earn new rewards, the system spreads the total amount evenly across the vesting period. Any small fractional amounts are added on the first day to ensure that no coins are lost due to rounding.
- **Automatic Management:** The system automatically handles all vesting operations without requiring participant intervention, maintaining efficiency while preventing system bloat.
- **Daily Releases:** Every day, the oldest vesting entry for each participant is automatically released to their available balance, creating a steady stream of unlocked coins.

This mechanism aligns Hosts' reward with the long-term stability and growth of the network, preventing short-term speculation and ensuring predictable income streams for reliable Hosts. Hosts can query:

- The total amount of coins to be vested.
- A detailed breakdown of their vesting schedule (the array of future unlocks).
- The total amount of coins that have already been released to them.

2.3. Collateral-Backed Influence

By default, only 20% of a Host's compute-derived voting weight is activated automatically. The remaining 80% becomes available only when the Host locks GNK coins as collateral. This ensures that participants with significant governance power also bear proportional economic responsibility. To strengthen network security and ensure incentive alignment, the following rules apply:

- **Potential Weight Calculation:** First, based on a participant's Proof of Compute activities (work done, nonces delivered, etc.), the system calculates their total Potential Weight.

- **Base Weight:** A portion of this Potential Weight is granted unconditionally as a base share. This is determined by a Base Weight Ratio, a governance-votable parameter representing the percentage of weight that is collateral-free. It is proposed to have a default value of 0.2 (20%). The formula is:

  *Base Weight = Potential Weight * Base Weight Ratio*

- **Collateral-Eligible Weight:** The remaining portion of the Potential Weight is the Collateral-Eligible Weight:

  *Collateral-Eligible Weight = Potential Weight * (1 - Base Weight Ratio)*

- **Enable Full Governance Weight:** To enable the portion of governance weight exceeding the base 20%, a participant must lock GNK coins as collateral. This additional share (called Collateral-Eligible Weight) is activated in proportion to the amount of collateral provided. The system uses a governance-defined parameter, Collateral Per Weight Unit, to determine how much collateral is required per unit of Collateral-Eligible voting weight. If a portion of the required collateral is locked, a corresponding fraction of the Collateral-Eligible Weight will be enabled. Full governance weight is enabled only when the entire required collateral is locked.
- **Final Effective Weight:** The participant's final, effective weight used in governance and other network functions is the sum of their Base Weight and the Activated Weight backed by their collateral.
- **Slashing:** Protocol violation triggers partial or total loss of collateral, ensuring malicious actors cannot cheaply influence network decisions.
- **Grace Period:** For the first 180 epochs (approximately 6 months), new participants can participate in governance and earn voting weight through PoC alone, without collateral requirements. During this period, the full governance rights are available, while voting weight remains tied to verified compute activity.

Voting power is never derived solely from holding coins. GNK coins serve as economic collateral, not as a source of influence. Influence is earned through continuous computational contribution, while locking GNK collateral is required to secure participation in governance and enforce accountability.

All network-level decisions, including block finalization, model registrations, unit pricing, protocol upgrades, and Community Pool allocations, are determined by PoC-weighted votes. All these parameters are defined in the Genesis Code and can be modified via governance proposals, allowing the network to dynamically adjust decision-making rules over time (see “Gonka: Whitepaper”).

2.4. Collateral Requirements for Proof of Compute

The weight a participant gains from Proof of Compute is directly proportional to the number of valid nonces they successfully submit during Sprint. To ensure that this influence is economically grounded, only a predefined Base Weight is activated by default. Any weight beyond this baseline, referred to as Collateral-Eligible Weight, requires locked GNK collateral.

The Collateral Per Weight Unit parameter, governed on-chain, defines how much GNK must be locked per unit of Collateral-Eligible Weight. Because a participant’s Potential Weight is a function of their nonce output, this parameter effectively establishes the required collateral per valid nonce.

For example, assume an H100-class GPU produces approximately 1600 nonces per epoch. To activate the full 80% Collateral-Eligible Weight, a Host may be required to lock 100 GNK coins. This results in a Collateral Per Weight Unit of 0.0625 GNK per nonce (100 GNK / 1600 nonces). In this way, as a participant’s network influence grows via Proof of Compute, their financial accountability scales proportionally, aligning governance power with both real-world compute and economic commitment.

2.5. Managing Collateral

To prevent abuse, withdrawals are not immediate. They are subject to an unbonding period that is tied to the network's epoch lifecycle. This ensures that collateral, which was used to gain influence, remains slashable for a period after the decision to withdraw is made.

The network is governed by its participants through adjustable parameters that can be modified through democratic voting:

Setting	Default Value	Purpose
Base Influence Ratio	20%	How much influence is guaranteed without locking collateral
Collateral Per Unit	0.0625 GNK	How much GNK must be locked as collateral to obtain one unit of influence
Withdrawal Period	1 network's epoch lifecycle	How long coins remain at risk after collateral is unlocked but before they become withdrawable
Malicious Behavior Penalty	20%	Penalty for cheating or providing bad results
Poor Performance Penalty	10%	Penalty for missing too much work
Performance Threshold	5% missed	How much work can be missed before penalties apply

2.6. Transaction cost management

Transaction costs within the network are designed to be flexible and predictable, following a dynamic, EIP-1559-inspired per-coin pricing mechanism. Users specify a maximum cost they are willing to pay per AI token, ensuring that tasks are performed within this predefined limit.
If the allocated coins for computation are exhausted before task completion, the transaction is canceled, and the spent coins are not refunded, similar to Ethereum's handling of exceeded Gas limits. The rewards generated from these transactions are distributed between the Host executing the task and the one verifying it, ensuring fair compensation across Hosts.

2.7. Long-term sustainability

The long-term strategy focuses on attracting Hosts willing to dedicate their hardware to the network, positioning the project as a cost-competitive alternative to traditional cloud providers. Leveraging decentralization, the network is designed to offer significantly lower prices for computational tasks, making it an attractive option for Developers and businesses seeking cost-effective solutions.
In conclusion, the tokenomics strategy is crafted to ensure the network's sustainability, support early adopters, and maintain competitive pricing.

3. Financial benefits for Developers and Hosts

3.1. For Hosts

1. Early growth and high reward potential:
    a. Gonka (GNK) coins are in their early lifecycle, with significant network growth ahead.

    b. Hosts benefit from low valuation, accumulating more coins before demand scales.
    c. As inference requests increase, rewards will rise, directly tied to network adoption.
2. Gonka Tokenomics is developed with Hosts' interests at its core:
    a. Direct coin rewards for computational contribution:
        i. 80% of total gonka (GNK) coins (800M) allocated for Host incentives.
        ii. Compensation is based on actual computational work, ensuring fair rewards.
            - Bitcoin-Style Epoch Rewards: A fixed amount of GNK is minted every epoch and distributed proportionally to Proof of Compute (PoC) weight. As more GPUs join the network, individual rewards decrease, creating scarcity and potential long-term price appreciation for GNK, similar to Bitcoin's mining dynamics.
            - Hosts earn Work Coins directly from developers for every inference task processed.
    b. Community pool and Host liquidity support:
        i. 120M gonka (GNK) coins are set aside for an early liquidity pool, governed exclusively by Hosts, to provide collective resources for network initiatives.
        ii. Hosts can exchange mined coins for stable assets (USDT).
3. Optimal use of computational power and Hosts' contribution to the AI future:
    a. Gonka doesn't use Hosts' computational power (and electricity) in the absence of tasks. Computational power can be used on other platforms, while there is nothing to compute on Gonka.
    b. Mining of gonka (GNK) coins is a way for individual hardware owners to contribute to and benefit from AI computational markets.
    c. Gonka (GNK) coins help to diversify revenue streams with less speculative, therefore more stable demand for compute power.
    d. GPU crypto mining clusters have compute power comparable to the top AI companies and cloud providers. However, many crypto resources are used for coin security rather than for useful computation. Gonka (GNK) coins are optimised for computational outcome (not compromising security), the majority of computational power contributes directly to the AI future.

3.2. For Developers

1. Transparent Fee Model: Developers pay exactly for the AI tokens they process — with the same price and equal access to resources for everyone, no matter the size of your project.
2. Predictable and gradual cost adjustments:
    a. Dynamic pricing allocates compute efficiently, preventing congestion on popular models and avoiding unfair queueing effects. Developers can choose models based on performance and cost trade-offs.
    b. Avoids sudden price hikes seen in centralized cloud services.
3. Developers who fund or contribute to model training earn a share of future inference revenue.
4. Early participation rewards:
    a. During the first 90 epochs, inference pricing is set to zero, enabling experimentation, rapid prototyping, and onboarding without cost barriers.
    b. Developers using the network in its early phases benefit from lower-cost computational resources.
5. Easy-to-use API with no DevOps (similar to platform.openai.com)
6. Mitigates the risk of large centralized AI platforms, their potential for censorship, and anti-competitive practices.

4. Risks and Considerations

While the proposed issuance model offers significant earning potential for Hosts and investors, particularly in early stages, its success depends on achieving adoption milestones, managing coin value stability, and fostering trust through transparent governance.

Hosts must carefully weigh the potential benefits of early participation against the following risks:

1. Market volatility risk - the hypothesis relies on achieving significant Developer adoption and sustaining demand for inference services. Fluctuations in the coin's market value may reduce the real-world value of rewards earned by Hosts, and unpredictable earnings may deter long-term investments in infrastructure. Transparent communication and governance can reduce speculation and foster coin stability.
2. Regulatory uncertainty risk* - coins issued as rewards could face regulatory scrutiny, especially in jurisdictions with strict cryptocurrency laws. The model assumes compliance with decentralized network principles and thus, no classification as a security in the United States and other jurisdictions with similar approaches to crypto asset regulation. As such, all figures and growth projections are theoretical and should not be construed as investment promises. Legal challenges may impact coin issuance and create unforeseen compliance burdens for Hosts. Adherence to decentralized principles and proactive monitoring of legal frameworks may, but is not guaranteed to, minimize risk exposure**.
   **Recipients of GNK coin rewards—whether through staking, mining, or participation—may incur tax obligations depending on their jurisdiction. The tax treatment of such rewards is evolving and can vary significantly across countries (e.g., income vs. capital gains treatment, timing of recognition, valuation issues). Gonka does not provide tax advice, and coin holders are solely responsible for reporting and paying any applicable taxes.*
   ***The legal and regulatory environment surrounding cryptocurrencies and blockchain-based assets is rapidly evolving and highly unpredictable. There is no guarantee that regulators, now or in the future, will agree with the positions taken by the Gonka project regarding the classification, treatment, or operation of GNK coins.*
3. Technological and market competition risk - centralized providers or competing networks may develop more efficient systems, reducing demand for decentralized alternatives. Providers could see reduced utilization of their infrastructure, however, differentiation through decentralization, transparency, and competitive rewards is key and can minimize risk exposure.

5. Dynamic Pricing for AI Inference

Automatic, real-time price inspired by Ethereum’s EIP-1559, applied on a per-model basis:

- **Per-Model Pricing:** Each AI model has an independent per-token price, dynamically adjusted every block based on recent network utilization and available capacity.
- **Stability Zone:** Prices remain unchanged when utilization is between 40% and 60% of model capacity. Outside this zone:
  - **Below 40%:** Prices decrease to encourage usage.
  - **Above 60%:** Prices increase to moderate demand.
- **Linear Elasticity:** Price changes are proportional to deviation from the optimal utilization range, capped at a 2% change per block.
- **Price Floor:** A minimum of 1 nicoin per AI token prevents zero-cost scenarios, maintaining economic incentives even during low demand periods.
- **Grace Period:** During the first 90 epochs, inference prices are set to zero, enabling early adoption and experimentation without cost barriers.

Decentralized Model Registration Process
Developers submit a model proposal specifying the model type, expected AI token throughput capacity, and inference characteristics. Governance requires a deposit to prevent spam, the proposal is reviewed by Hosts, and once approved, the model is listed with a dynamically adjusted per-token price based on real-time utilization and available capacity.

6. Hosts' compensation model for training new Large Language Models

The decentralized AI infrastructure is built on a foundational commitment to open-source development and accessibility. Advancing AI should not be controlled by centralized entities only. Instead, it must remain collaborative, transparent, and widely available to Developers and users worldwide. To uphold this commitment, the network guarantees that all models trained using its resources will remain open-source, ensuring that AI innovation remains decentralized and equitable. The network actively commits to continuously contributing to open-source AI development, setting it apart from centralized entities like

Meta, which do not provide such guarantees. This commitment fosters motivation among contributors and ensures that innovation remains in the hands of a global community rather than a select few.

However, training new Large Language Models is one of the most computationally intensive processes in AI. Fair and sustainable compensation for Hosts is essential to keeping the decentralized AI network competitive and encouraging long-term participation.

To achieve this, the network adopts a Network-Financed approach, where 20% of all inference revenue (including transaction fees and mining rewards) will be allocated to fund the training of new models. This allocation will specifically cover the cost of the unit-of-compute used during training, ensuring that resources are efficiently utilized to support model development. Additionally, a portion of this 20% can be allocated as grants, awarded through a voting process, to support the most promising and innovative training procedures proposed by the community. This mechanism ensures that the network continues to support large-scale AI advancements, even after the initial subsidies phase out. Moreover, as the network grows, this approach allows for the collection of substantial funding, making it feasible to train massive AI models that compete with centralized alternatives.

The network will implement a consensus-based governance mechanism to maintain flexibility in adjusting the revenue allocation percentage. If new models significantly boost network adoption, consensus may choose to maintain or increase the 20% allocation. If the allocation impacts network competitiveness, consensus can reduce the percentage, ensuring a balance between affordability and innovation. This mechanism will only become available after Year 5, ensuring the network has demonstrated sufficient model training capabilities before modifications are allowed.

The decision on how to conduct the next training procedure follows a clear yet flexible process to ensure the most effective training approaches receive support. Rather than focusing solely on which model to train, the emphasis is on refining the training methodology. Contributors, including Hosts and Developers, can propose code changes (such as pull requests in a repository) that introduce new training strategies. These proposals undergo thorough discussion, with extensive presentations and debates to assess their feasibility and effectiveness.

Once a training approach gains consensus, it is first approved as a code update. Following this, the same Host can formally propose a full training experiment, detailing the parameters, dataset, and allocated funding. This proposal then enters a governance voting phase, where it is likely to undergo multiple iterations, including rejections and modifications, before reaching final approval. The network ensures that the best AI research and development teams worldwide can propose and experiment with promising training techniques, even if they lack their own hardware. This ensures that innovative research is not limited by computational constraints while maintaining full transparency (everyone can still access and evaluate R&D results, fostering continuous innovation). Priority is given to training experiments that demonstrate strong potential to maximize impact, ensuring resources are directed toward AI advancements that offer the greatest value. This iterative process fosters innovation, aligns the network's growth with user needs, and ensures that training methodologies continue to evolve in the most effective way possible.

Why the Decentralized AI Training Fund? Various approaches were considered for financing LLM training. The Decentralized AI Training Fund was chosen due to its sustainability, fairness, and ability to encourage meaningful participation from Hosts.

Financing Model	Advantages	Challenges
Decentralized AI Training Fund (chosen)	● Ensures stable, long-term funding; ● Predictable rewards for Hosts;	● Requires governance to adjust revenue allocation

	• Scalable as network revenue grows	
New Coin Issuance	• Immediate incentives for Host; Allows early growth	• May lead to inflation and reduced coin value; Unclear incentive alignment
Shareholding Structure	• Hosts receive compensation through additional fees paid by Developers, along with a proportional share of the network reward. • Hosts receive a stake in the model and are compensated whenever it is used. • Creates long-term incentives to support training.	• Developers using trained models may pay an additional fee. • Additional usage fees may be at a disadvantage against free alternatives like LLaMa (funded by Meta). • Network may subsidize the cost, though this could result in slightly lower earnings for Hosts. • Technical challenges related to the rights for the model weights, which could lead to security threats and potential attacks targeting model ownership.

The Decentralized AI Training Fund ensures long-term funding for LLM training, balancing sustainability with competitive pricing. By allocating 20% of inference revenue to model training, the system secures billions in funding while allowing for governance-based adjustments in later years. This approach benefits Hosts and Developers, fostering a scalable, decentralized AI ecosystem that can compete with centralized cloud solutions while maintaining fairness and efficiency.

Host-financed training relies on direct funding from Hosts and Developers, creating a more demand-driven approach. While this model requires users to actively contribute to financing, potentially limiting participation, it remains a viable option for those who choose to fund training independently. Additionally, while there is a risk of model duplication and free redistribution, this approach offers flexibility for individuals or organizations willing to invest directly in model development.