

Rating System Proposal

TLDR on proposal

We need to figure out how to score operators if we're going to optimize our set of validators. With an effective scoring system we can then use the two levers

we have to improve Lido performance;

1. which operators get a new validator, and
2. which operators get validators exited when liquidity is necessary.

I've spent a lot of time on this already, but the scoring component itself needs significant additional research. This proposal is to conduct that research to understand the data available for use in a rating system, outline the availability of the data, features of the data itself and figure out how this data might be used in a rating system.

Importantly, this proposal has been crafted around the existing and present work from the Nethermind team

. I've been on regular calls with the Lido and Nethermind team for a couple of months now, and so have a good understanding of where this work will fit into the overall system being worked on both by Lido and Nethermind.

Summary of existing work

Following discussions on the forum around a system for rating operators and subsequently altering the number of validators they operate, I started working on a new system.

The new system involved a couple thousand lines of code and was accompanied by a 15,000+ word thesis. This thesis explained the rationale behind each parameter in the model, as well as why other metrics were ignored. [@lizzy](#), [@vsh](#) and the Nethermind team ([@mpzajac](#)) have reviewed this thesis in detail.

While I'm very happy with the model and thesis, it's not production ready. There's still a decent amount of work to do to really understand precisely what data should be used for rating operators. Then, we need to decide where we get the data from, and how it's used.

Synopsis

Here's a very high-level synopsis of the approach I took:

Every day we get the data for the last 30-days and calculate the Node Operator Score (simply score going forward) for each operator. The NOS is a single number between 0 and infinity.

The system budgets a certain amount of validators assigned to new operators in a given period. If there is budget remaining, we run the full model. If there is no budget, we only consider existing operators. We can think of this as a graduation system, where new operators must graduate from level 1, to level 2, to level 3 etc, whereby the maximum number of validators an operator can run will increase by level.

To decide which operator gets assigned a new validator, we simply randomize but with their normalized NOS as a weighting.

Hence, if a score is above the average, the operator will have a higher chance of being selected, while a lower scoring operator will have a lower chance. Everyone with a score above zero has a chance of being selected.

When users request a redemption of their funds, we must decide which validators to exit. To do this, we get the reciprocal of the score and multiply it by the number of validators that an operator is running and then multiply it by a damping variable.

This damping variable is set dynamically to a value that makes the largest operator twice as likely to be exited as the smallest active operator, if they both have the same score. The resulting number is used as the weight when randomly selecting an operator to have their validator exited.

This approach means that operators with more validators have a higher chance of being exited. However, the scores still have a significant impact on the odds.

The above is a very brief explanation of the core scoring system. Clearly, the most important factor in the scoring is actually what goes into creating the score, which I've spent a good amount of time on, but for which further research is required.

That is the topic of this future grant!

Proposal

Objective:

Understand the data available for use in a rating system to score node operators on Lido. Outline the availability of the data, features of the data itself and how this data might be used in a rating system.

Purpose:

With greater understanding of the potential options for inclusion in a rating system, we can narrow our focus to a subset of data. By doing so, we can create a rigorous rating system to monitor the health of the Lido operators. Eventually, this system could be used for influencing the staking router and the exiting of validators.

Proposed research:

1. Research existing rating systems, including Rated.Network and Observation.Zone. (10-hours)
2. Summarize and critique these existing systems, giving an executive summary of where they succeed and opportunities for improvement. (10-hours)
3. Research and list the data available to us for use in a rating system. Excluding those that are difficult to use or unrealistic to obtain for a rating system, for example, personal information better suited for a whitelabel-resistance module. (20-hours)
4. Critically examine the data and its features to decide whether it could be utilized in a rating system. (30-hours)
5. Comment on the usefulness of each dataset for a rating system, pitfalls to avoid and how it might be used for rating. (30-hours)

Output:

The final output will be a report detailing the proposed research above and recommending a final set of inputs to be used for a rating system. Data sources tested will be stored, but neither the data nor any code used for examining the data will be included in the output to avoid excess hours on this project, without meaningful benefit to Lido (I don't want to spend hours polishing clean code, clean data and presentations, that's not helpful).

Deadline:

15th August, 2023

Expected hours:

100-hours more + 150-200 hours already invested

Total cost:

\$30,000

Cost breakdown:

\$15,000 USD for work so far + \$15,000 USD upon delivery and successful LEGO review of pending work described in this proposal. To be paid in fiat or USD stablecoins.