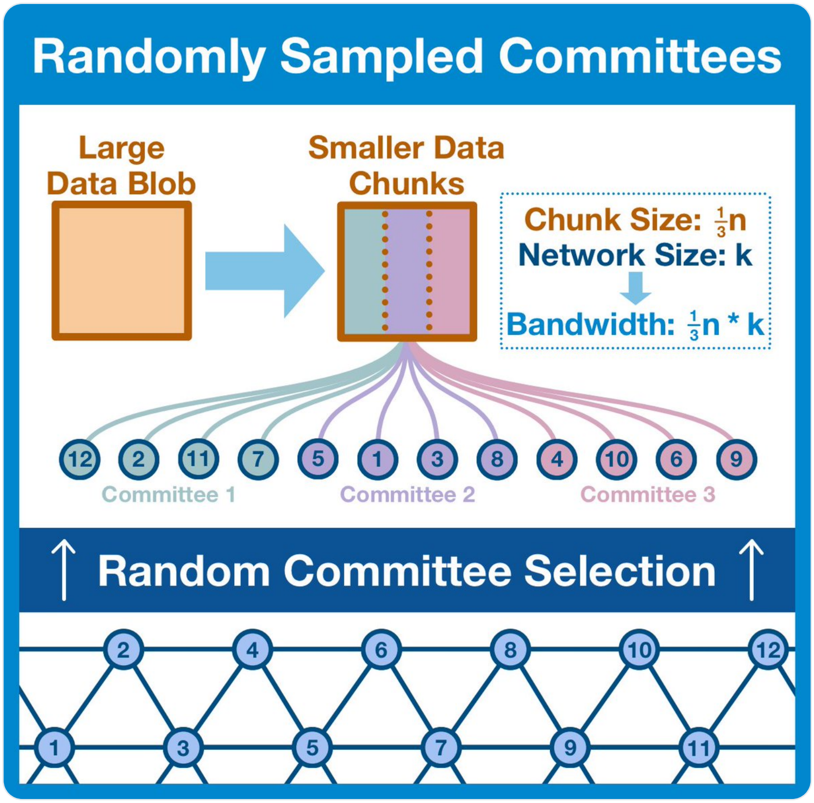
Haym @SalomonCrypto

Oct 24 • 18 tweets • SalomonCrypto/status/1584335007291613184

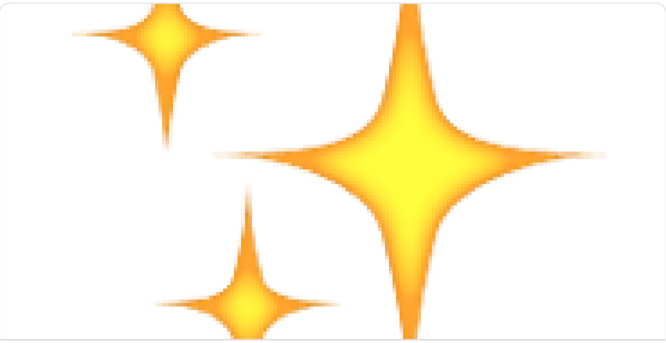
(1/17) The World Computer of tomorrow is a rollup-centric @ethereum. But rollups produce data, lots and lots of data. How can we scale a growing network without growing bandwidth?

We are going to need a couple solutions, let's start with Randomly Sampled Committees.



(2/17) @ethereum has been charting a course towards a rollup-centric future for ~2 years now.

Tl;dr Ethereum will scale by offloading execution to performance-optimized blockchains, while settlement and economic security will remain with mainnet.



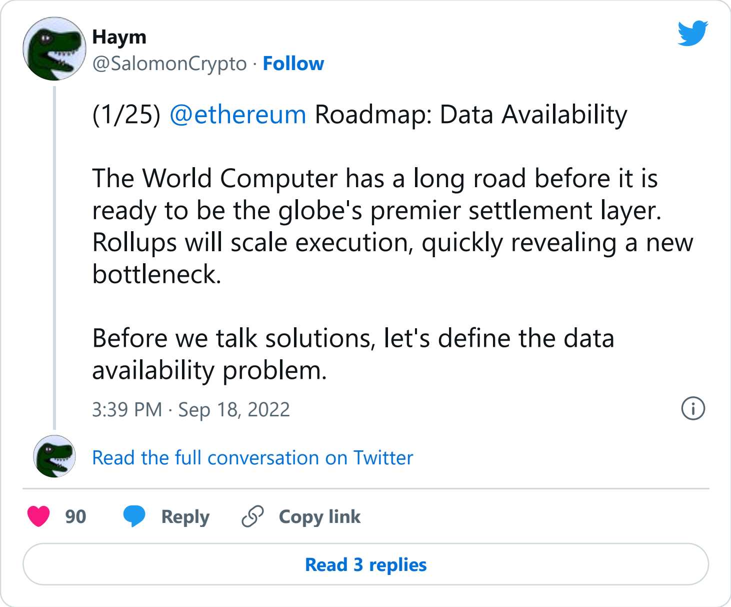
A rollup-centric ethereum roadmap

What would a rollup-centric ethereum roadmap look like? Last week the Optimism team announced the launch of the first stage of their testnet, and the roadmap to mainnet. They are not the only ones; F…

https://ethereum-magicians.org/t/a-rollup-centric-ethereum-roadmap/4698

(3/17) The rollup-centric paradigm is based around quickly and cheaply executing transactions in a more centralized environment, and then posting a (compressed) record back to @ethereum.

In the rollup-centric future, it becomes critical to ensure that data is available.



Haym
@SalomonCrypto · **Follow**

(1/25) @ethereum Roadmap: Data Availability

The World Computer has a long road before it is ready to be the globe's premier settlement layer. Rollups will scale execution, quickly revealing a new bottleneck.

Before we talk solutions, let's define the data availability problem.

3:39 PM · Sep 18, 2022

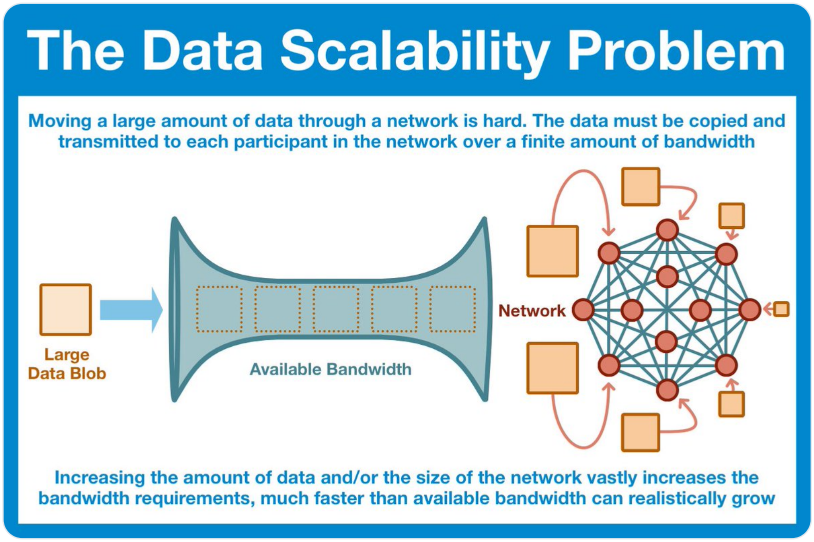
Read the full conversation on Twitter

(5/17) Now look, I am an @ethereum zealot, a true believer in the inevitable dominance of the World Computer. I believe that billions and billions of transactions will happen on rollups every single day.

And that is A LOT of data circulate through the Etheruem network.

(6/17) The goal: verify the availability of high volumes of data without requiring any single @ethereum node to download and personally verify ALL of the data.

Even if we were ok with forcing every node to download the data, the reality is that the network could not handle it.



The Data Scalability Problem

Moving a large amount of data through a network is hard. The data must be copied and transmitted to each participant in the network over a finite amount of bandwidth

Increasing the amount of data and/or the size of the network vastly increases the bandwidth requirements, much faster than available bandwidth can realistically grow

(7/17) Before we continue we need to be clear on what kind of data we are talking about.

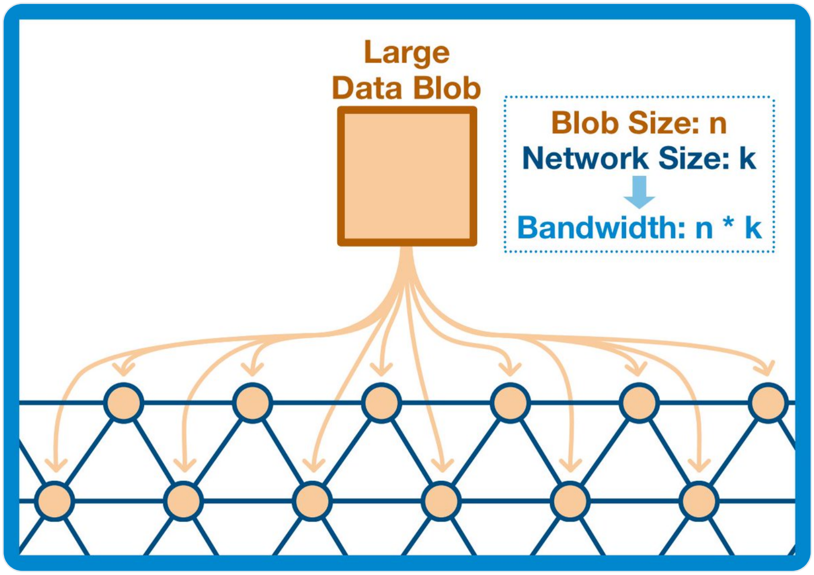
Any data within the EVM (the computing environment within @ethereum) needs to be communicated across the entire network. ALL nodes needs a copy of EVERY EVM state change.

(8/17) Rollups exists OUTSIDE of the EVM; they are independent blockchains. Yes, they post a copy of the transaction to @ethereum but that's just to ensure immutability.

We only care that a copy is posted to the World Computer, not that each node gets a copy.

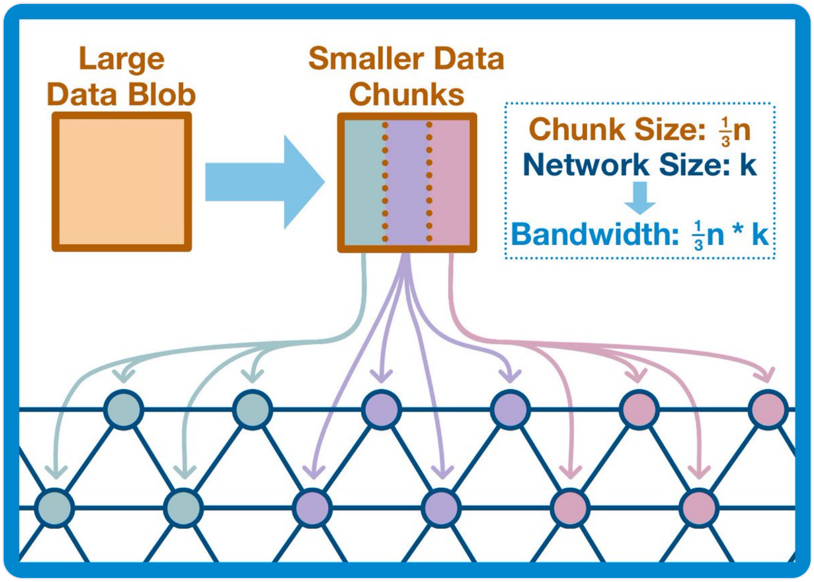
(9/17) The most simple and obvious way to ensure that the data is posted to the World Computer is just to have every node in the network download it. Yes, the nodes don't NEED it, but at least we KNOW it's there.

But, as we've already discussed, this just isn't scalable.



(10/17) One idea is to split up the work: simply divide the nodes in your network and your data into an equal number of groups.

The nodes in group one can download and validate the first chunk, the nodes in group two can check the second, etc.



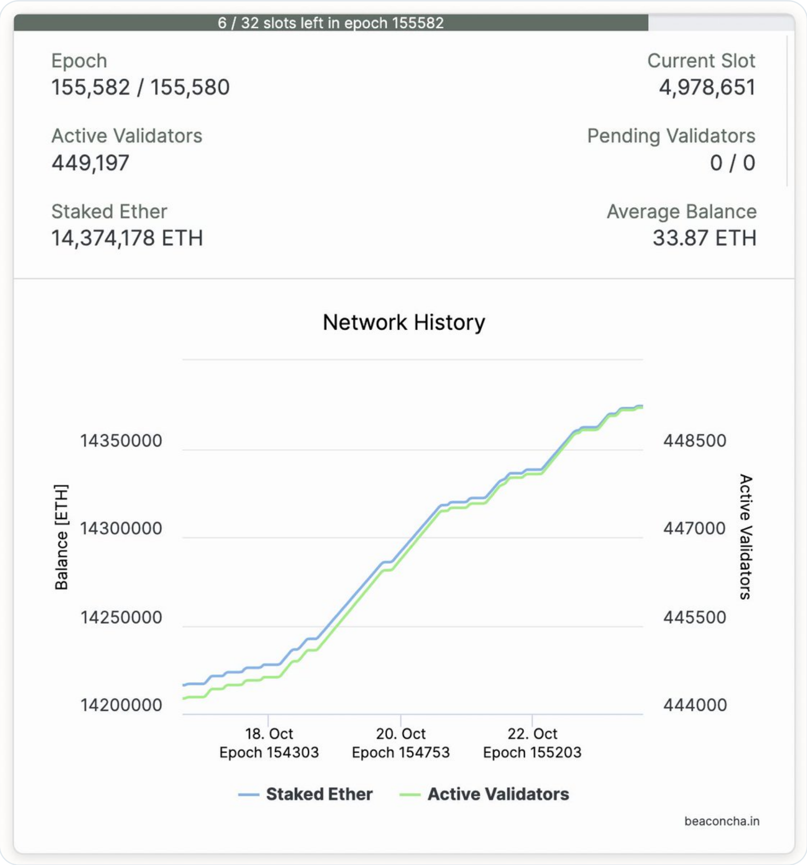
(11/17) Each group (committee) is responsible for downloading their assigned chunk of data.

If all/most of the committee can download it, the network can consider the chunk available.

If all/most of the chunks are available, the network can consider the blob available.

(12/17) Go check out @beaconcha_in; you'll see there are currently ~450k @ethereum validators. Let's decide that if 500 validators can successful download a chunk of data, we can call it available.

So let's just give ourselves committees of 1000, just in case.



6 / 32 slots left in epoch 155582

Epoch
155,582 / 155,580

Current Slot
4,978,651

Active Validators
449,197

Pending Validators
0 / 0

Staked Ether
14,374,178 ETH

Average Balance
33.87 ETH

Network History

(13/17) 450k validators / 1000 validators per group = 450 committees

450 committees means we can split our data into 450 pieces, and therefore our bandwidth requirements are 450x less!

Change the parameters all you want, this committee approach will save a lot of bandwidth.

(14/17) But we have an issue. Let's say a malicious node operator saved up his $ETH until he had enough to deploy enough validators to control an entire committee.

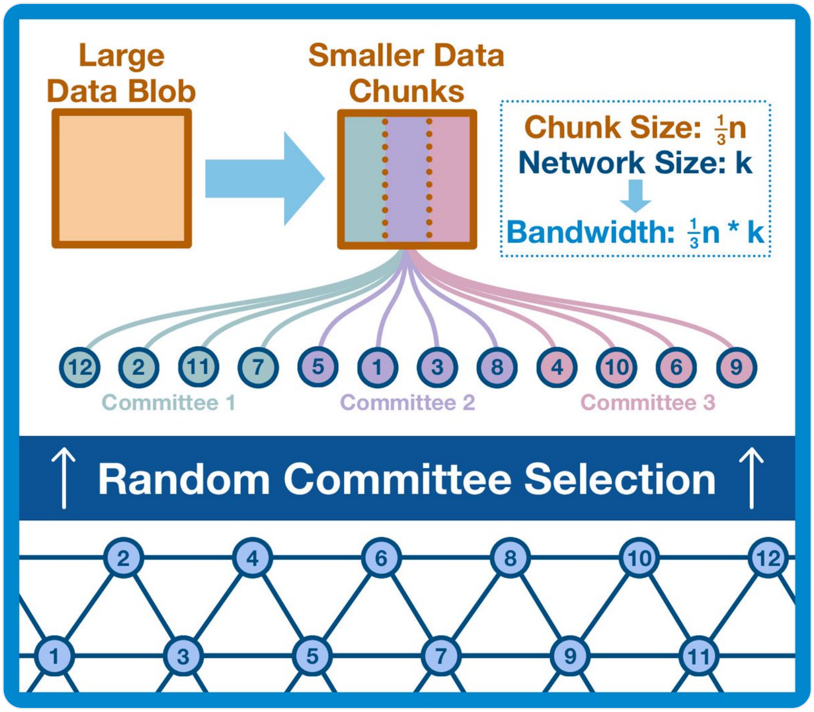
With control of a committee, he could easily confirm an invalid or unavailable chunk.

(15/17) In our example above, controlling .2% of the validators would give you 100% control over an entire committee. That committee would only have control over one chunk, but one chunk is enough.

What if it contains a transaction to mint 100,000,000,000 $USDC?

(16/17) Fortunately, the solution is simple: we shuffle the validator set before we assign the data chunks.

Using a random seed, we assign each validator to a committee arbitrarily. A malicious node operator has no way to control which committees his validators are assigned to.



(16/17) With enough validators, it's totally possible that an attacker might get lucky and gain temporary control over a single control over a committee.

However, this situation is only mathematically realistic if they control >1/3 of all validators (compromising @ethereum PoS).

(17/17) As @ethereum continues to build towards a rollup-centric future, it becomes critical to design for increased data requirements of a robust rollup ecosystem.

Randomly sampled committees are a step forward in ensuring data is available without crushing the network.

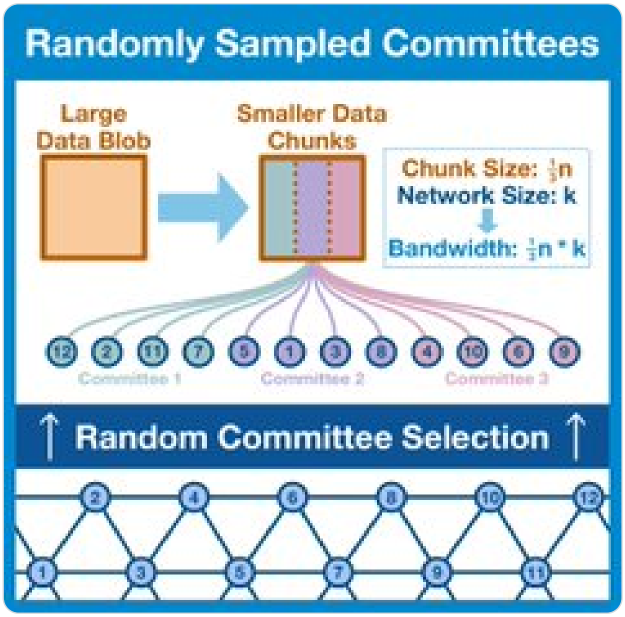
Like what you read? Help me spread the word by retweeting the thread (linked below).

Follow me for more explainers and as much alpha as I can possibly serve.

Haym
@SalomonCrypto · **Follow**

(1/17) The World Computer of tomorrow is a rollup-centric @ethereum. But rollups produce data, lots and lots of data. How can we scale a growing network without growing bandwidth?

We are going to need a couple solutions, let's start with Randomly Sampled Committees.



• • •