TLDR

: This is a mashup of thehonest-majority availability post and @vbuterin’s erasure coding note, with the aim to achieve scalable fork-free sharding.

Construction

Suppose there are n validators and that we have $\textrm{SHARD\_COUNT}$ shards. For every collation body B to be pushed to the VMC, the proposer prepares an erasure coding with n pieces $B_1, ..., B_n$, one for each validator. The erasure coding is such that any $\textrm{SHARD\_COUNT}$ pieces are sufficient to reconstruct B. The proposer also needs to prove that the Merkle root r of the pieces B_i corresponds to a faithful erasure coding of B (the hard part!).

We now require a BLS threshold signature of r from $(n + \textrm{SHARD\_COUNT})/2$ validators before B can pushed to the VMC. This can be guaranteed if the honest majority assumption is strengthened from $n/2$ to $(n + \textrm{SHARD\_COUNT})/2$.

If collation sizes are capped at 1MB and $\textrm{SHARD\_COUNT} = 100$ then each honest validator only has to download 10kB (the size of a piece) per collation per shard. That is, we can have honest-majority availability votes across all shards for just the bandwidth cost of a single shard, thereby achieving scalability.