Now that Casper is moving from a prepare/commit scheme to a simpler single vote scheme, the incentive structure can change (and hopefully be simplified!) as well.

At a high level, the protocol has rounds of voting (called “epochs”) where two consecutive rounds with \geq 2/3

votes (by weight) lead to finality. In each round, we know which validators voted and which didn’t, and we can summarize the liveness strategies for validators as “Vote. No seriously, vote!”

For each epoch, let us consider a couple of parameters that may be interesting to our analysis of incentives:

F_v

: the fraction of validators who voted, by weight.

TSF

: number of epochs, since the last finalized epoch, that we have failed to finalize. For example, if the epoch previous to the current one was finalized, then TSF = 0

.

TD

: the total sum of all the deposits of the all the validators.

We can now define two functions (names a WIP

):

Y(F_v, TSF, TD)

N(F_v, TSF, TD)

In some epoch, each validator who votes gets Y(F_v, TSF, TD)

applied to their balance, and the same for a validator who did not vote and N(F_v, TSF, TD)

. For example, if a validators votes during an epoch, and Y(F_v, TSF, TD)

return .001, then the validators balance grows during that epoch by .001.

From here, our goal is to specify reasonable constraints on Y

and N

so the incentives encourage what we want. Many thanks to Vitalik for these initial thoughts and encouraging me to start a thread here

Here are some basic examples of possible constraints:

1. Validators who are not voting should not make a profit. So, N(F_v, TSF, TD) \leq 0

.

1. If finality is consistently reached (which means F_v = 2/3

and TSF = 0

), a validator should be net punished if they are voting less than 2/3

of the time. Thus, as this validator will be rewarded with Y(2/3, 0, TD)

two-thirds of the time, and N(2/3, 0, TD)

one-third of the time, thus we can say that Y(2/3, 0, TD) * 2/3 + N(2/3, 0, TD) * 1/3 = 0

.

1. Validators who are online and voting should never be penalized to much.
2. As TSF

increases, the punishments on the offline validators should increase as well. This implements the “leak” mechanism, where if some coalition of validators goes offline (or is censored), their deposits will decrease until finality can be achieved again.

1. We should bound greifing factors wherever possible and explore the trade-offs between them. For example, Y(x, y, z) < Y(1, y, x), for all x < 1. Otherwise, a coalition of validators can make a profit if they censor other validators, which is not good! In general, a goal should be to characterize the tradeoffs with these greifing factors.

As an example of exploring the tradeoffs, consider that the “leak” mechanism affects the “length” of the weak subjectivity synchrony assumption. For example, if 2/3

of offline weight can “leak away” in 2 weeks, then we could end up with two chains, one with the 2/3

of the original weight and the other with 1/3

the original weight, where both chains are reaching finality, and no validators were slashed. On the other hand, this mechanism stops the FFG from getting stuck if some large amount of validators go offline, and also make it possible for a censored minority of validators to start their own chain!