0368.3311 The Theory Behind Blockchains April 3

Analysis of Nakamoto's Consensus

Lecturer: Nir Bitansky

There are several reasonable ways to model Nakamoto's protocol execution. We will follow a relatively simplified view of the protocol suggested by Pass and Shi [PS17].

The protocol Π. In Π all nodes have oracle access to a function $H : \{0,1\}^* \rightarrow \{0,1\}^n$ modeled as a random oracle, as well as access to a function $Valid_H : \{0,1\}^* \rightarrow \{0,1\}$, which tests validity of any chain (which we will define later). Every node maintains a blockchain denoted ch at any point of time. We think of ch as an array where each $ch[i]$ is a (mined) block with format

$$ch[i] := (h_{-1}, tx, nonce),$$

a hash of the previous block, transactions, and a puzzle solution for the block.

The parties proceed as follows:

- Every node starts with an initial chain containing only a special genesis block:
$$ch[1] := (\bot, \bot, \bot) \ .$$
- In every step: a node reads all incoming messages. If the longest incoming ch' such that
$$Valid_H(ch') = 1$$
is longer than its own ch, it replaces ch with ch'. (If there's more than one, it chooses arbitrarily.)
- Read an input tx.
  - Let $ch[-1] = (h_{-1}, \bot, \bot)$ be the last block in the chain.
  - Pick at random $nonce \leftarrow \{0,1\}^n$ and issue a query $h = H(h_{-1}, tx, nonce)$.
  - If $h < p \cdot 2^n$, then update the last block $ch[-1]$ to $(h_{-1}, tx, nonce)$, append $(h, \bot, \bot)$ to ch, and broadcast the updated chain. Here p is a difficulty parameter, which the protocol determines, based on the number of players and latency. We will define this later.
- Output the new ch.

Validity. We say a block $ch[i] = (h_{i-1}, tx_i, nonce_i)$ is valid with respect to (a predecessor block) $ch[i-1]$ if $h_{i-1} = H(ch[i-1])$. We say that an ℓ-block chain ch is valid if $ch[1] = (\bot, \star, \star)$ and for every $i \in [\ell - 1]$, ch_{i+1} is valid with respect to $ch[i]$. $Valid_H(ch) = 1$ iff ch is valid.

What can the adversary do? The adversary can corrupt a certain fraction of nodes and make them act arbitrarily. The only restriction is that any node can make a single query to H at every step (just like honest parties).

Where do transactions come from? In essence *we don't care*. We abstract this part out. We may think of an arbitrary efficient process TX that generates the input transactions tx to the parties at every step (which may or may not depend on the parties' outputs at every step). In particular, we are not concerned with the validity of transactions, as far as the protocol is concerned input transactions are created equal. Indeed, the validity of transactions is handles by a separate set of rules (enforced by signatures etc).

1 Theorem Statement

To state the main theorem, we will think of the protocol as a random process and argue that when sampling a random execution through this process, good properties hold with overwhelming probability.

Theorem 1.1 (Main theorem (for now, informal))**.** *Let N be the number of nodes, ρ be the fraction of malicious nodes, and Δ be the network latency (maximal number of steps it takes for a message to reach all parties). For $\rho < 0.49$, $p \approx \frac{1}{20\Delta N}$ (where 20 is a large enough constant),in a random execution of Π, with probability $1 - 2^{-\Omega(n)}$, the following properties hold:*

- **Chain growth:** *honest chains grow at rate $\approx \frac{1}{\Delta}$.*
- **Chain quality:** *for any long enough segment on an honest chain the fraction of malicious blocks is at most $\approx \frac{\rho}{1-\rho}$.*
- **Consistency:** *any two honest chains agree, except for some relatively short prefix.*

2 Proof

An ideal abstraction of the protocol. To prove the security properties of Nakamotos protocol Π, it helps to think of an abstract ideal protocol $i\Pi$ that captures the nature of the randomized process behind Π. In the ideal protocol, nodes mine blocks by interacting with an oracle $Tree$ that keeps track of all the blockchains mined thus far. We describe the oracle $Tree$ and protocol $i\Pi$ and then explain that to prove the security of Π it is sufficient to prove that of $i\Pi$.

The oracle $Tree$. $Tree$ is parameterized by the difficulty parameter p, so that each mining attempt succeeds with probability p. $Tree$ keeps track of the set $Chains$ of all blockchains mined thus far. Initially, the only blockchain in the set tree is the genesis chain $\perp$. $Tree$ allows honest and corrupt nodes to mine blocks and verify the validity of blockchains through the following interfaces:

- Upon receiving a query $extend(ch, tx)$: $Tree$ checks if ch is a valid blockchain in $Chains$. If so, $Tree$ flips a coin that is 1 with probability p. If the coin flip is 1, $Tree$ adds $ch||tx$ to the set $Chains$, and returns *success*.
- Upon receiving $verify(chain)$: $Tree$ checks if $ch \in Chains$; if so, return 1; else return 0.

The ideal protocol $i\Pi$. The protocol $i\Pi$ chooses a difficulty parameter p in the same fashion as Π based on (Δ, N). The protocol proceeds as follows:

- Every node maintains the longest blockchain ch it had seen so far.
- In every round, every honest node first receives all incoming messages. For any received message ch', if $Tree(verify(ch')) = 1$ and ch' is longer than the current local chain ch, then let $ch = ch'$ and broadcast ch'.
- In every round, every honest node receives an input record tx, and then queries $Tree(extend(ch, tx))$. If this mining query is successful, the node broadcasts $ch||tx$ and appends tx to its local chain ch.

Abstract chains. Note that we now think of any chain ch simply as a chain of transactions, without any hashes h or POW solutions *nonce*.

Claim 2.1. *If the ideal protocol $i\Pi$ is secure (in the sense that it satisfies the properties given by Theorem 1.1), then so is Π.*

The essential difference between the real protocol Π and the ideal protocol $i\Pi$ is that in the real one, the possibility of hash collisions may allow the adversary to perform attacks that cannot occur in the ideal world. The claim is proved using the fact that the random oracle H is collision resistant — the probability of finding a collision using q queries is at most $q^2/2^n$. Assuming no collisions, it is not hard to see that the ideal protocol perfectly emulates the real one. We will not formally prove the claim. A proof can be found in [PSS17].

2.1 More Parameters

- We will denote by N_H the number of honest nodes and by N_M malicious nodes.
- We will denote by $\beta = pN$ the expected number of blocks mined by all parties in every step (we assume w.l.o.g that malicious nodes like honest nodes always try to mine), by $\beta_H = pN_H = (1-\rho)\beta$ the expected number of blocks mined by honest parties, and by $\beta_M = pN_M = \rho\beta$ the expected number of blocks mined by malicious parties.
- We will denote by $\Lambda = 2.01(\Delta + 1)\beta_H$ a parameter called the latency loss. We will assume throughout $\Lambda < 1$. We will eventually want to choose p so that Λ will be a small constant, say 0.01.

A simplifying assumption. We will make the simplifying assumption that N, Δ are fixed and depend only on the security parameter n. We need to know N and Δ in order to set the difficulty parameter p in a way that will guarantee security. The analysis can be extended, at the account of adding more details, to deal with the fact that the total number of nodes (or computational power) N change over time, requiring that p is updated accordingly. It also extends to show that it's enough to have reasonable approximations of Δ and N. (Recall that the actual Bitcoin protocol uses the mining histogram to derive an approximation of the total computational power. Also, we have an approximate bound on latency that can be obtained by random testing of the network — an approximate bound is about a minute.)

We will also make a simplifying assumption that $\rho > 1/10$; namely, the fraction of malicious parties is not too small. This is not needed and is only meant to simplify the asymptotics in our analysis.

2.2 Convergence Opportunities

We define a useful notion of convergence opportunities, which we will use for proving the required properties of the protocol. Intuitively, a convergence opportunity is a Δ-period of silence in which no honest node mines a block, followed by a step in which a single honest node mines a block, followed by another Δ-period of silence.

Formally, given a k-step execution, we say that $t^* \in (\Delta, k - \Delta]$ is a convergence opportunity if:

- For any $t \in [t^* - \Delta, t^*)$, no honest node mines a block in step t;
- A single honest node mines a block in step t^*;
- For any $t \in (t^*, t^* + \Delta]$, no honest node mines a block in step t.

Lemma 2.2 (Number of convergence opportunities)**.** *Consider $k(n)$-step execution. Then with probability $1 - k^2 2^{-\Omega(n)}$, for any segment $[r, r+t]$ of length $t > \frac{n}{30\beta_H}$,*

$$Conv[r, r+t] > 0.98 \cdot (1 - \Lambda) \cdot \beta_H \cdot t \ .$$

From hereon, we will always consider $k(n) = n^{O(1)}$ and thus the above bound will be $1 - 2^{-\Omega(n)}$.

Fact 2.3 (Chernoff). *Let $X_1, \ldots, X_n$ be independent Bernoulli variables with $E[X_i] \leq p$, then*

- $\Pr\left[\frac{\sum_i X_i}{pn} \in (1-\delta, 1+\delta)\right] \geq 1 - 2^{-\Omega(\delta^2 pn)}$ *(multiplicative version).*
- $\Pr\left[\sum_i X_i - pn \in (-\delta, \delta)\right] \geq 1 - 2^{-\Omega(\delta^2 n)}$ *(additive version).*

Proof of Lemma 2.2. Fix a segment $[r, r+t]$ as above. Denote by X the total number of blocks mined by honest parties during the segment, then $E[X] = pN_H t = \beta_H t$ and since all the relevant coin tosses are independent, it holds by (multiplicative) Chernoff that

$$\Pr[X < 0.99\beta_H t] \leq 2^{-\Omega(\beta_H t)} \leq 2^{-\Omega(n)} \ .$$

From hereon fix $\mu = 0.99\beta_H t$.

Now let us denote by Y_i the event that after the i-th honest block is mined, the next $(\Delta+1)N_H$ honest mining attempts result in at least one success.

$$E[Y_i] = \Pr[Y_i = 1] \leq pN_H(\Delta+1) = \beta_H \Delta$$

Define $Y = \sum_{i\in[\mu]} Y_i$ and note that Y_i are independent (Y_i doesn't care about $Y_1, \ldots, Y_{i-1}$). Thus, by (additive) Chernoff:

$$\Pr[Y \geq \beta_H(\Delta+1)\mu + \mu/200] \leq 2^{-\Omega(\mu)} \leq 2^{-\Omega(0.99\beta_H t)} \leq 2^{-\Omega(n)} \ .$$

Symmetrically, we denote by Z_i the event that before the i-th honest block is mined, the previous $(\Delta+1)N_H$ honest mining attempts result in at least one success and define $Z = \sum_{i\in[\mu]} Z_i$. We have:

$$\Pr[Z \geq \beta_H(\Delta+1)\mu + \mu/200] \leq 2^{-\Omega(n)} \ .$$

By considering the first μ added blocks, we have that with probability $1 - 2^{-\Omega(n)}$.

$$\begin{aligned} Conv[r, r+t] \geq & X - Y - Z \geq \\ & \mu - 2\beta_H(\Delta+1)\mu - \mu/100 = \\ & (0.99 - 2(\Delta+1)\beta_H) \cdot 0.99\beta_H t \geq \\ & 0.98(1-\Lambda)\beta_H t \ . \end{aligned}$$

Taking a union bound over all $[r, r+t] \subseteq [1, k]$, yields the lemma. □

2.3 Chain growth lower bound

To prove the chain growth lower bound, we observe that for any execution, whenever there is a convergence opportunity, the shortest honest chain must grow by at least one. We will then use the previous lemma to prove that honest chains grow rapidly.

Lemma 2.4 (Chain growth lower bound.). *With probability $1 - 2^{-\Omega(n)}$ over an execution, for any t_0 and $t_1 = t_0 + t$ and honest chains ch_0 and ch_1 in steps t_0 and t_1, respectively, such that $t > \frac{n}{20\beta_H}$,*

$$|ch_1| - |ch_0| \geq 0.97(1-\Lambda)\beta_H t \ .$$

Proof. First, we claim that

$$|ch_1| - |ch_0| \geq Conv[t_0 + \Delta, t_1 - \Delta] \ .$$

Indeed, if t is a convergence opportunity, then the shortest honest chain at the end of round $t+\Delta$ must be longer than the longest honest chain at the beginning of round $t-\Delta$, which now implies the above, by taking the contribution from all converges opportunities in $[t_0, t_1]$.

By Lemma 2.2,

$$Conv[t_0 + \Delta, t_1 - \Delta] \geq 0.98(1-\Lambda)\beta_H(t - 2\Delta) \geq 0.97(1-\Lambda)\beta_H t \ .$$

where we used the fact that $2\Delta < \frac{1}{\beta_H} < \frac{t}{n} \leq \frac{t}{100}$ for large n. □

2.4 Chain quality

We will prove chain quality by comparing the number of malicious blocks with the honest chain growth lower bound. If malicious nodes mine fewer blocks than the minimum honest chain growth, we will conclude that there cannot be too many malicious blocks in an honest nodes chain.

Lemma 2.5 (Chain quality)**.** *With probability $1 - 2^{-\Omega(n)}$, for any $\ell \geq n$, for any honest chain ch and ℓ blocks $ch[j+1], \ldots, ch[j+\ell]$, at most $\gamma\ell$ blocks are malicious, where*

$$\gamma \leq \frac{1.05}{(1-\Lambda)} \frac{\rho}{1-\rho} \quad .$$

Remark: One could expect that the bound would be ρ rather than the above. A reason for the loss is the selfish mining attack, where malicious nodes wait to hear about the next honest block before publishing their own block. Intuitively since they can do it for any block they mine, the total fraction of work in the system that is not wasted is $1 - \rho$ and not 1.

Befroe we prove the lemma, we state two simple lemmas that we will use in the proof and also later on.

Lemma 2.6 (Total block upper bound)**.** *With probability $1 - 2^{-\Omega(n)}$ over an execution, for any r and $t > \frac{n}{20\beta}$, the total number of blocks mined in $[r, r+t]$ is bounded by $1.01\beta t$.*

Proof. By Chernoff and union bound. □

Lemma 2.7 (A bound on adversarial blocks)**.** *Let $Adv[r, r+t]$ denote the number of blocks mined by malicious nodes during $[r, r+t]$. With probability $1 - 2^{-\Omega(n)}$, for any r and $t > \frac{n}{20\beta_M}$,*

$$Adv[r, r+t] \leq 1.01 \cdot \beta_M \cdot t.$$

Proof. By Chernoff and union bound. □

Proof. Consider an execution and let us assume that Lemmas 2.4,2.6,2.7 hold for this execution (which occurs with probability $1 - 2^{-\Omega(n)}$). Let ch be any honest chain output by some party during the execution and consider any ℓ-block sub-chain $ch[j+1], \ldots, ch[j+\ell]$. We will assume that $ch[j]$ and $ch[j+\ell+1]$ are both blocks that were mined by honest parties (or the genesis or end-of-chain blocks). This is w.l.o.g as we can always further extend the sub-chain until we reach such a chain, this only decreases the chain's quality.

Let r be the step in which $ch[j]$ was mined and let $r+t$ be the step in which $ch[1, j+\ell]$ first appeared as an honest party's chain, and note that by the definition of the ideal protocol all ℓ blocks in the subchain were mined during $[r, r+t]$. We'll denote by $ch_r = ch[1, j]$ and $ch_{r+t} = ch[1, j+\ell]$ to two corresponding honest chains.

Roughly speaking, we will use chain growth to show that t is not much longer than ℓ. Then we will bound the number of adversarial blocks mined in these t steps to get our bound.

First, we will argue that t is large enough for the guarantees of our lemmas to kick in.

Claim 2.8. $t > \frac{n}{20\beta_M} > \frac{n}{20\beta_H}$.

Proof. Let $t' = \frac{n}{2\beta}$. Then, by the total block upper bound (Lemma 2.6), any t'-step segment has at most $1.01 \cdot \beta t' < n$ blocks. Since at least $\ell \geq n$ blocks are present in the segment $[r, r+t]$, it follows that

$$t > t' = \frac{n}{2\beta} > \frac{n}{20\beta_M} \quad ,$$

where we use our simplifying assumption that $10\beta_M > \beta$. □

We can now apply the chain growth lower bound to the chains ch_0 and ch_1 to deduce

$$\ell = |ch_{r+t}| - |ch_r| \geq 0.97(1-\Lambda)\beta_H t \ ,$$

which gives us an upper bound on t.

Now, we can apply the bound on malicious blocks:

$$Adv[r, r+t] \leq 1.01 \cdot \beta_M \cdot t \leq \frac{1.05}{1-\Lambda} \cdot \frac{\beta_M}{\beta_H} \cdot \ell = \frac{1.05}{1-\Lambda} \cdot \frac{\rho}{1-\rho} \cdot \ell \ .$$

□

2.5 Consistency

We will now prove consistency, which means that any two honest chains agree with one another, except for the last n blocks. This also implies that once you see your block deep enough in the chain it is going to stay there. The idea behind the proof is that honest chains have many convergence opportunities, and in order to prevent consistency the adversary must intervene — spend work — in all of them.

Lemma 2.9. *Assume that $\frac{\rho}{1-\rho} < 0.97(1-\Lambda)$. Then with probability $1 - 2^{-\Omega(n)}$, for any two steps $r, r+t$ and two honest chains ch_r, ch_{r+t} in these steps,*

$$ch_r[1, \ell - n] = ch_{r+t}[1, \ell - n] \ ,$$

where ℓ is the length of the shortest chain amongst the two.

Proof. We focus on the case that $t \leq \Delta$. This is actually enough to prove to general case (homework).

Let i be the maximal index such that $ch_r[1, i] = ch_{r+1}[1, i]$ and the i-th block is honest. Let $s - 1$ be the step in which i was mined and note that by the definition of the ideal protocol all blocks at positions $> i$ on ch_r and ch_{r+t} must have been mined after step $s - 1$. Now consider any convergence opportunity $t^* \in [s, r - \Delta]$, then the corresponding honestly mined block must be at position $i_{t^*} > i$ and all such convergence opportunities occur at distinct positions.

We now claim that for every such convergence opportunity the adversary must soon enough mine a block of its own in position i^{t^*}, or that ch_r, ch_{r+t} would have converged there.

Claim 2.10. *For every such t^*, a malicious block must be mined in position i_{t^*} between $[s, r+t]$.*

Proof. By the definition of a convergence opportunity, there is a single honest block mined in $[t^* - \Delta, t^* + \Delta]$, which is mined in time t^* in position i_{t^*}. If no malicious block is mined in position i_{t^*} during $[s, r+t]$, for any two honest chains ch, ch' at times $[t^* + \Delta, r+t]$, and in particular for ch_r, ch_{r+t}, it holds that $ch[1, i_{t^*}] = ch'[1, i_{t^*}]$. This contradicts the fact that i is maximal. □

By Claim 2.10,

$$Adv[s, r+t] \geq Conv[s, r - \Delta] \ .$$

We show that this implies that $i \geq \ell - n$, which will conclude the proof. Assume toward contradiction that $i < \ell - n$.

Claim 2.11. $t' := r - \Delta - s > \frac{n}{\min\{20\beta_M, 30\beta_H\}}$.

The proof is similar to that of Claim 2.8 (taking into account the fact that $\Delta < \frac{1}{\beta_H}$) and is omitted.

Now, we apply the bounds on convergence opportunities and adversarial blocks, to deduce

$$\begin{aligned} Adv[s, r+t] \leq & 1.01\beta_M(t' + t + \Delta) \leq \\ & 1.01\beta_M(t' + 2\Delta) < \\ & 0.98 \cdot (1-\Lambda) \cdot \beta_H \cdot t' \leq Conv[s, r-\Delta] \ , \end{aligned}$$

where we use the assumption that $\frac{\beta_M}{\beta_H} = \frac{\rho}{1-\rho} < 0.97(1-\Lambda)$ and the fact that $2\Delta < \frac{1}{\beta_H} < \frac{t}{n}$.

2.6 Setting the Parameters

We briefly discuss how one can set the parameters to yield the properties claimed in Theorem 1.1. If we set $p \approx \frac{1}{200N\Delta}$, then it is guaranteed that $\beta_H = pN_H > pN/2 \approx \frac{1}{400\Delta}$ and $\Lambda = 2.01(\Delta + 1)\beta_H \leq 0.01$. We accordingly get

- chain growth at rate $\approx \frac{1}{400\Delta}$,
- chain quality $\approx \frac{\rho}{1-\rho}$
- and consistency for all but the last trailing n blocks.

□

References

[PS17] Rafael Pass and Elaine Shi. Rethinking large-scale consensus. In *30th IEEE Computer Security Foundations Symposium, CSF 2017, Santa Barbara, CA, USA, August 21-25, 2017*, pages 115–129, 2017.

[PSS17] Rafael Pass, Lior Seeman, and Abhi Shelat. Analysis of the blockchain protocol in asynchronous networks. In *Advances in Cryptology - EUROCRYPT 2017 - 36th Annual International Conference on the Theory and Applications of Cryptographic Techniques, Paris, France, April 30 - May 4, 2017, Proceedings, Part II*, pages 643–673, 2017.