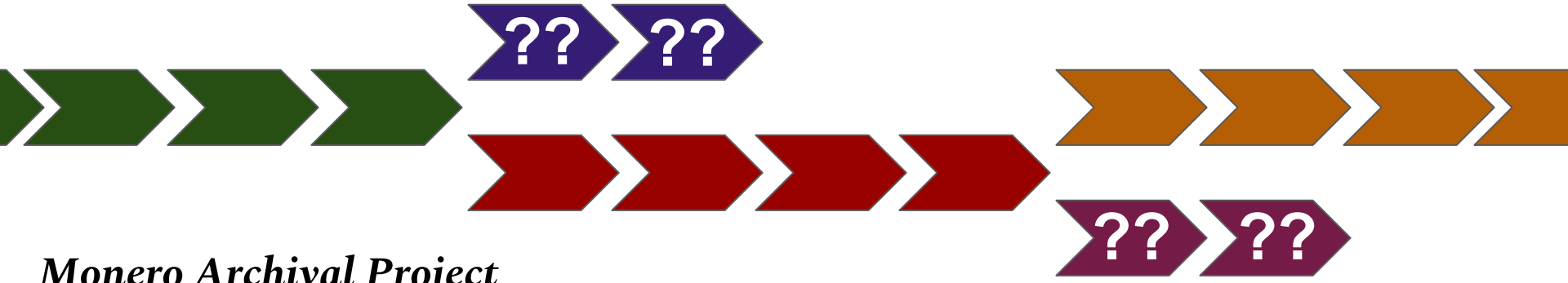
RoadMAP v 0.1

IsthmusCrypto



Monero Archival Project
https://github.com/mitchellpkt/monero_archival_project

*A product of #**NoncesenseResearchLab***
Blocks courtesy of NeptuneResearch
Analysis by IsthmusCrypto

Objective: `both_sides`

Timeline: 2018.08

What data is necessary?

If a given height is solved by multiple (2+) miners, we need to retain ALL versions of the block

In other words, any 2+ blocks A and B where:

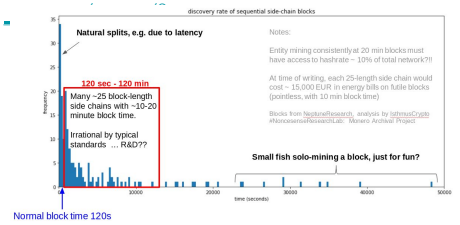
height(A) == height(B)
nonce(A) != nonce (b)

GitHub issue: *https://github.com/Mitchellpkt/monero_archival_project/issues/16*

Notes: When retaining side blocks, MAP daemon does not record main blocks, currently!!

What analysis is enabled?

- How often do side chains occur? Once per day? Once per week?
  https://github.com/Mitchellpkt/monero_archival_project/issues/12
- How long are the side chains? An orphaned block here and there? Or long side chains?
  https://github.com/Mitchellpkt/monero_archival_project/issues/13
  (result: both)
- If long side chains, how much hash power does the entity have? *(result: 50MH/s)*
  https://github.com/Mitchellpkt/monero_archival



Why is this necessary?

Right now, nobody is retaining the orphaned blocks/chains. Thus, this is important to:

- Allow R & D
- Archive for future
- Enable analyses at the left

Visualization would be new/useful.

MRL wants this as an opt-in daemon feature

Applicable to all blockchains

Objective: `MRT_and_NRT`

Timeline: 2018.08

What data is necessary?

For each instance of a received block, record the:

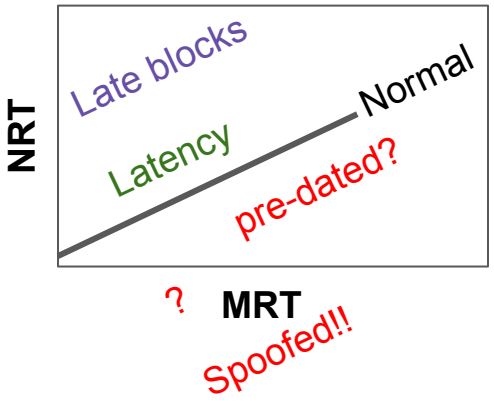
Node-received timestamp (NRT), when each MAP node hears each instance of the block

Miner-reported timestamp (MRT), extracted from the block. *(Miners only report to the second, and this data can be spoofed)*

What analysis is enabled?

Is there timestamp spoofing??

Is there selfish mining? https://bitcoinmagazine.com/articles/selfish-mining-a-25-attack-against-the-bitcoin-network-1383578440/



Why is this necessary?

Right now we only record the MRT, which can be spoofed. NRT will allow us to see when blocks are REALLY mined.

Without NRT, all other analysis have to use MRT, which makes the results uncertain (since timestamps necessarily accurate)

Applicable to all blockchains

GitHub issue:
https://github.com/Mitchellpkt/monero_archival_project/issues/7

Notes:

Objective: `subsecond_NRT`

Timeline: 2018.08

What data is necessary?	*What analysis is enabled?*	*Why is this necessary?*
Node-receipt timestamp (NRT) with subsecond accuracy We may receive several copies of the exact same block (currently, these are distinguished by the `random` field), and we'll want to keep a precise timestamp for when each copy arrives	When running on a single node, this enables study of latency. What's the timing on the shortest route for a txn/block to arrive at MAP node? What's the timing on the longest route? What is the scale of the time difference? milliseconds? Seconds? When running on global nodes, can then start studying overall network topology/latency	Without subsecond NRT, all copies of the blocks may appear with the exact same timestamp, which makes it impossible t study differences in the timing. Since I have not had access to NRT thus far, I do not know what the scale of the timing should be.

GitHub issue:
https://github.com/Mitchellpkt/monero_archival_project/issues/17

Applicable to all blockchains

Notes:

Objective: `transactions_archive`

Timeline: 2018.08

What data is necessary?

Each transaction received by the MAP node can be identified by its transaction hash. Some of these transactions are later mined in blocks, showing up in the "tx_hashes" field.

We want to retain content and timing on ALL transactions announced, whether or not they are ever mined.

Especially the key images!

What analysis is enabled?

- Double-spend attack detection!! Studying *likely* causes of forks. A single orphaned block with a bunch of tx_hashes that match with the main chain is probably benign latency split. If it contains totally different transactions, then it is more likely to be a double-spend attack.
- Similar analyses regarding timing, as described in other

Why is this necessary?

Currently it is impossible to compare transactions between main and side chains

Currently it is not possible to see how often a key image is reused (indicating that an attacker is trying to spend a spent output)

Currently, transactions that are not mined are discarded (the same way side blocks have been). This leaves Monero with a blind spot.

Applicable to all blockchains

GitHub issue:
https://github.com/Mitchellpkt/monero_archival_project/issues/14

Notes: Unlocks next level of analyses

Objective: `retain_IP_addresses`

Timeline: 2018.09

What data is necessary?

When a block or transaction is received, record the IP address from which it originated.

This serves two purposes:

- Cross-referencing block and txn broadcasts
- Approximate geolocation

What analysis is enabled?

Some blocks we get lots of copies, some blocks only one. Why is this? Do we have 1 or 2 reliable connections and then a bunch of spotty ones? Or do we have 0 reliable connections and just pick up data randomly?

{Once objective `global_operations` is underway, then combining this data from multiple nodes will open up way more analyse}

Why is this necessary?

Right now duplicate copies of blocks are recorded, but we don't keep any track of which copies come from which nodes.

This makes it impossible to do any analysis of how nodes pass around data, the stability of connections, etc.

GitHub issue:

https://github.com/Mitchellpkt/monero_archival_project/issues/18

Notes:

Preserve privacy when publishing!!!!

Obviously given VPNs, IP != user location. However is valuable for studying network

Applicable to all blockchains

Objective: **Timeline: 2018.10**

`global_operations`

What data is necessary?	*What analysis is enabled?*	*Why is this necessary?*
MAP data as described above, being recorded and collected from archival nodes VPS's hosted around the world	When a transaction or block is first announced to the network, which node hears it first How does it propagate across the globe? How quickly does it take that path? This is data that can be presented in realtime on a MAP dashboard. This conveniently also strengthens the Monero Network :-)	Right now, our single MAP archival daemon only collects data from one location. Thus, it is unknown how representative our data is (for some analyses) Multiple nodes will give larger sample sizes to all analyses, which is crucial. Big picture: In a year, this project may scale beyond Monero, and maintain VPSs collecting this important archival information across multiple different blockchains. This could be a generalized specialty service.

GitHub issue: TBF

Notes:

Objective:
`template_slide`

Timeline: 2018.xx

What data is necessary?	*What analysis is enabled?*	*Why is this necessary?*
The data should	The data should	The data should
Look like:	Look like:	Look like:
Lorem Ipsum	Lorem Ipsum	Lorem Ipsum

GitHub issue: TBF

Notes:

Monero specific

- or -

Applicable to all blockchains