Where is the Light(ning) in the Taproot Dawn? Unveiling the Bitcoin Lightning (IP) Network

Pedro Casas*, Matteo Romiti*, Peter Holzer*, Sami Ben Mariem†, Benoit Donnet†, Bernhard Haslhofer*
*AIT Austrian Institute of Technology, †Université de Liege

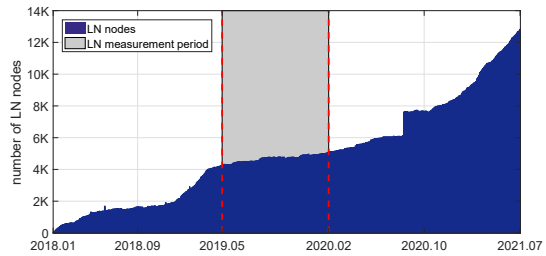
***Abstract*—Proposed in 2016 and launched in 2018, the Bitcoin (BTC) Lightning Network (LN) can scale-up the capacity of the BTC blockchain network to process a significantly higher amount of transactions, in a faster, cheaper, and more privacy preserving manner. The number of LN nodes has been significantly increasing since 2018, and today there are more than twelve thousand nodes actively participating of so-called LN payment channels. The upcoming Taproot upgrade to the Bitcoin protocol would further boost the development and adoption of the LN. Taproot is the most significant upgrade to the Bitcoin network since the block size increase of 2017, and it will make LN transactions cheaper, more flexible, and more private. We focus on the characterization of the LN network topology, using network active measurements. By crawling the underlying P2P network supporting the Bitcoin LN over a span of 10-months, we unveil the LN in terms of size and location of its nodes as well as connectivity protocols, comparing it to the P2P IP network supporting the BTC blockchain. Among our findings, we show that IP addresses exposed by LN nodes correspond mainly to customer networks, even if most BTC nodes are actually deployed at major cloud providers, and that LN nodes significantly rely on anonymized networks and protocols such as Onion, with more than 40% of LN nodes connect through Tor.**

I. Bitcoin LN 101 and Taproot

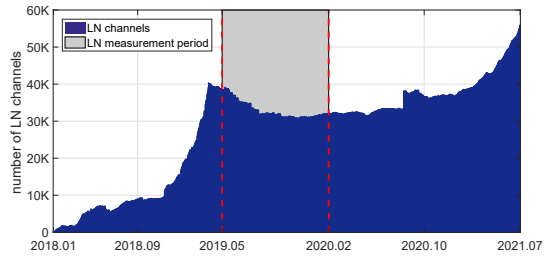
Lightning Network: one of the major limitations of the Bitcoin blockchain is its inherent lack of scalability to handle a significant number of transactions. Limited by the size and frequency of blocks, Bitcoin can currently handle around 5 to 7 transactions per second (tps), whereas mainstream centralized payment providers such as Visa or Mastercard can settle around 2.000 tps or more. Different solutions have been proposed to handle this scalability issue; in this paper, we focus on the most promising solution for scalability, namely the Lightning Network (LN) [1]. LN is a layer 2 solution that enables faster and more scalable payments that periodically anchor in aggregate form to the Bitcoin blockchain, offering eventual Bitcoin security while amplifying speed and potential throughput. A LN is a Payment Channel Network (PCN) which represents a peer-to-peer (P2P) overlay running on top of a blockchain, improving its scalability and speed without altering its properties and functioning. The main idea of the LN is to handle transactions in an off-chain manner – i.e., without registering them to the blockchain's ledger, achieving instant transaction confirmation times with negligible fees, whilst retaining the security of the underlying blockchain. To do so, the LN forms bidirectional payment channels between BTC addresses, which can be used to transact thousands of payments with essentially only two on-chain visible transactions: the opening and closing channel transactions. In a nutshell, the channel opening transaction sets the initial balances of both BTC addresses, whereas the closing one sets the last agreed settlement, distributing the total balance accordingly. The strength of the approach is that, once a channel is open, it can be used by other LN nodes to route further payments, creating as such a fast and scalable PCN with only few on-chain transactions.

Taproot: the Taproot Bitcoin upgrade will switch over to so-called Schnorr signatures, which essentially makes multi-signature transactions *unreadable*. The upgrade will mean greater transaction privacy and efficiency, and will unlock the potential for smart contracts, a key feature of its blockchain technology. The specific benefits of the upgrade include: (1) *increased privacy* – this does not refer to Bitcoin addresses or enhanced anonymity, but rather to types of transactions. Taproot will make complex transactions, such as those requiring multiple signatures or those with delayed release, indistinguishable from simple transactions in terms of on-chain footprint; (2) *lower fees* – the data size of complex Bitcoin transactions will be reduced, which will lead to lower transaction costs; (3) *more flexibility* – the new type of signature will enhance smart contract functionality in Bitcoin, making it easier and cheaper for users to set more complicated conditions for a transaction. These improvements would mean a strong boost for the LN, as LN transactions would become cheaper, more flexible, and more private. Privacy is indeed a major concern for LN transactions, as we have recently shown that security and privacy of LN payments are weaker than commonly believed [2].

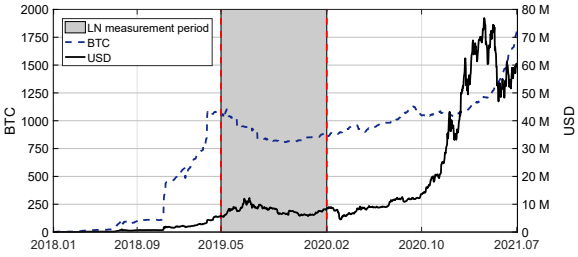
In the dawn of the Taproot upgrade, we study the P2P network behind the LN. We follow the same approach we took to characterize the P2P network supporting BTC, through active network measurements. By crawling the full P2P Bitcoin LN over a span of 10-months, we study the size of the LN as well as the location of its nodes. The crawling is realized through a modified LN node, which connects to the P2P network and gets access to the full list of available LN nodes. Recent work has also focused on the analysis of the LN [3]–[6], but generally relying either on single snapshots of the network, or from a different perspective to ours. In particular, our study locates the main Autonomous Systems (ASes) where LN nodes are deployed, and presents a comparison against BTC nodes. As we show next, an important share of BTC's infrastructure is hosted at public cloud providers in US and EU countries – mainly Germany, and major German and US ISPs provide the connection to LN nodes.



(a) LN nodes (with established channels).



(b) LN channels.



(c) LN capacity.

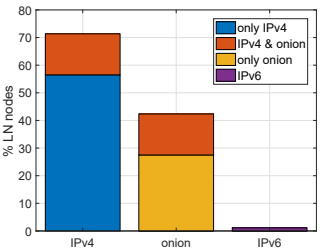
Fig. 1: Evolution of the LN network over time.

II. Unveiling the Bitcoin LN

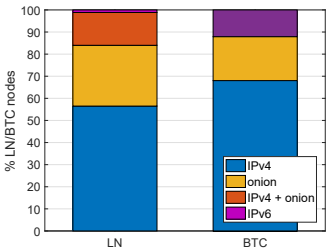
A. Data Collection and Description

To discover the nodes of the LN P2P network, we conceived a LN crawler. This LN crawler is a customized LN software client which can recursively query all the LN peers of the network, discovered by asking other nodes for the underlying IP addresses. We run two modified LN clients to interact with the LN, which periodically collect the topology of the network and keep a local copy of it. The clients are based on the Lightning Network Daemon (LND), which is a complete implementation of a LN node (https://github.com/lightningnetwork/lnd). The first client started storing data from May 2019 on, at intervals of 30 minutes, and the second one started in November 2019, at intervals of 60 minutes. The total time-span of the data collection is 10 months, from May 2019 till February 2020. The topology is retrieved using the *describegraph* command implemented by LND; for the purpose of this study, each entry on the topology description corresponds to a node ID, an (IP) address where this node is active, and an alias.

LN nodes can join the network and create payment channels using three main LN clients: LND, C-Lightning, and Eclair. A LN node typically needs to run on top of a full BTC node, but might not run on the same physical machine, and newer lightweight LN clients can even run without direct access to a

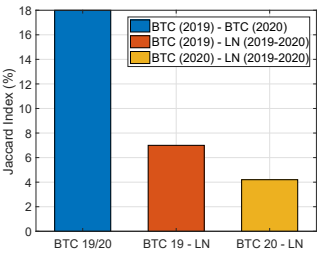


(a) Addresses in LN

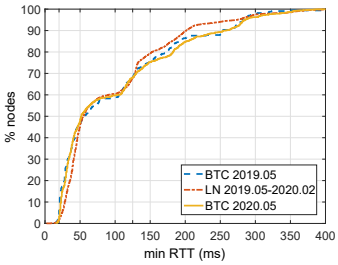


(b) Share of Addresses in LN and BTC

Fig. 2: Protocols announced by LN and BTC.



(a) Intersection of BTC and LN IP sets.



(b) Min RTT to LN and BTC nodes.

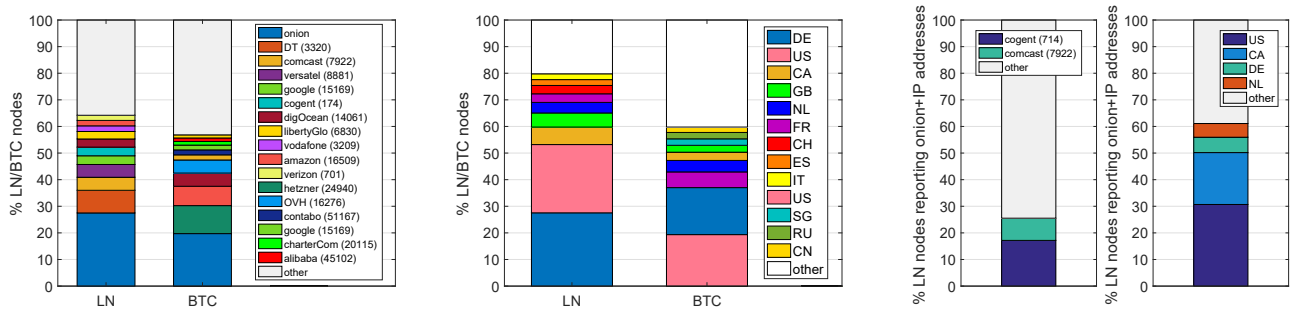
Fig. 3: Co-location of LN and BTC nodes.

local BTC node. For example, Eclair clients off-load trust to third-party servers through the usage of Simplified Payment Verification (SPV) technology. Therefore, to better understand the location of LN nodes, we additionally take snapshots of the underlying BTC P2P network, and compare the results. To do so, we follow a similar approach to our previous work on BTC [7], also based on active crawling. We take two snapshots of the BTC network for comparison, one at the start of the LN measurement campaign (May 2019), and a second snapshot one year later, by May 2020.

We complement the network measurements with LN statistics obtained through on-line APIs available at https://lopp.net/lightning-information.html. This open project provides an index of publicly-available curated data and resources on the LN.

B. Evolution of the LN Size and Relevance

Figure 1 depicts the evolution of the LN network over time in terms of (a) number of *active* nodes (i.e., with established payment channels), (b) number of payment channels, and (c) cumulative transaction capacity of the overall payment network, both in BTC and USD. For reference, our LN crawling campaign is marked in the figures (May 2019 to February 2020). The LN size and adoption have been steadily increasing since its inception in 2018, and has seen an outstanding growth in the last six months, matching the BTC price outbreak in late 2020. The network has today more than 12.000 active nodes and more than 55.000 established payment channels, with a total transaction capacity of roughly 1750 BTCs, equivalent to about 60 million USD.



(a) ASes hosting LN/BTC nodes. (b) Countries. (c) *Hidden* LN nodes.

Fig. 4: BTC nodes are mainly deployed at major cloud providers, whereas LN node IPs correspond to customer networks.

C. Node Location Analysis

We detected a total of 8.267 LN nodes active in the network during the 10-month measurement campaign. These nodes are not necessarily unique, as they were detected over the total span of the measurements, i.e., they were not necessarily simultaneously active. For comparison to the BTC network, we take the May 2020 snapshot of the BTC P2P network. The BTC network consisted of 10.355 reachable nodes at the time of this snapshot.

Figure 2 depicts the share of protocols announced by LN and BTC nodes. Most of LN nodes announce IPv4 addresses, and the usage of IPv6 in LN is negligible, which is significantly different from BTC, where it is used by more than 10% of the nodes. As we see next, this is coherent with the fact that LN nodes are mostly located at customer ISP networks, whereas BTC infrastructure is mostly hosted at public cloud networks, which offer IPv6 service. LN nodes significantly rely on anonymized networks – e.g., Tor (onion). Indeed, more than 40% of the LN nodes connect through onion. Interestingly, about 15% of the LN nodes expose both IPv4 addresses and onion IDs, which can be used to understand where anonymized nodes connect from. The prevalence of anonymized connections is significantly higher in the LN.

We now take a closer look to the commonalities between IP address sets between LN and BTC. Given the negligible usage of IPv6 in LN, we just consider the IPv4 sets for both LN and BTC, which in both cases corresponds to roughly 70% of the detected nodes. Recall that a LN node may run on top of a full BTC node, but it could also run as an independent wallet at a different device, thus IP addresses for LN and BTC nodes do not necessarily overlap. In fact, the overlap between addresses in the considered sets is negligible. Figure 3(a) reports the overlap between different IPv4 addresses sets, through a standard Jaccard Index (JI). For two given sets, a JI $= 1$ means total overlap, and a JI $= 0$ means total separation. For reference, we consider both BTC snapshots at May 2019 (BTC-19) and May 2020 (BTC-20). The overlap between LN and both BTC-19 and BTC-20 results in a JI $= 7\%$ (892 IPv4 addresses) and JI $= 4.2\%$ (509 IPv4 addresses), respectively. Naturally, the lack of overlap might also be exacerbated by the usage of dynamic IP addresses, and the instantiation and disconnection of nodes. For reference, the overlap between BTC-19 and BTC-20 results in a JI $= 18\%$ (2244 IPv4 addresses). Based on these results, one might hypothesize that the usage of the LN is becoming more user-centric, with a smaller fraction of LN nodes corresponding to full BTC nodes. Still, as we show next, geo-localization of LN and BTC nodes remains stable over time.

To dig deeper into the co-localization of nodes, we consider the distribution of minimal RTT (i.e., a reference to propagation latency) from a vantage-point located in EU, towards the identified LN and BTC nodes. Figure 3(b) depicts the corresponding CDFs. The overlap among CDFs shows not only how stable seems to be the BTC network over time in terms of geo-localization of its nodes, but also that only slight differences between LN and BTC nodes are visible, which might suggest that the co-location of LN and BTC nodes is actually much higher than what we observe through the JI indexes.

Finally, Figure 4 reports the (a) hosting ASes and (b) locations (country) of the nodes. Again for comparison purposes, we take the BTC-2020 snapshot. We obtain geo-localization information from publicly available geo-referenced IP databases. While locations tend to be similar - both the BTC and the LN networks are mainly located in western countries, being US and Germany the dominant hosting countries, there is a significant difference in the type of ASes where BTC and LN nodes are deployed. Active BTC nodes are mainly hosted by major cloud providers in EU and US, such as Hetzner, Amazon, Google, DigitalOcean, and OVH; however, LN nodes expose connections on mayor ISPs corresponding to the location countries, such as Deutsche Telekom (Germany) and Comcast (US). A confirmed interpretation of these results would require a deeper and joint topological view of both LN and BTC layers, but a priori one could hypothesize that the topology shows LN nodes deployed at customer networks, connecting to the local BTC P2P infrastructure, deployed at the cloud. Figure 4(c) shows the location of those LN nodes using anonymized connections, but also exposing IPv4 addresses. These *hidden* LN nodes are mostly located in North America – US and Canada, hosted by customer ASes (Cogent, Comcast, etc.), with notable EU presence in Germany and the Netherlands.

REFERENCES

[1] J. Poon, et al., "The Bitcoin Lightning Network: Scalable Off-chain Instant Payments," online available at https://lightning.network/lightning-network-paper.pdf, 2016.

[2] M. Romitti, et al., "Cross-Layer Deanonymization Methods in the Lightning Protocol,", in *Financial Cryptography and Data Security – FC*, 2021.

[3] E. Roher, et al., "Discharged Payment Channels: Quantifying the Lightning Network's Resilience to Topology-based Attacks," in *IEEE Euro S&P Workshops*, 2019.

[4] I.A.. Seres, et al., "Topological Analysis of Bitcoin's Lightning Network," in *Mathematical Research for Blockchain Economy*, pp. 1–12, 2020.

[5] S. Martinazzi, et al., "The Evolving Topology of the Lightning Network: Centralization, Efficiency, Robustness, Synchronization, and Anonymity," in *PLOS ONE*, vol. 15(1), 2020.

[6] P. Zabka, et al., "Node Classification and Geographical Analysis of the Lightning Cryptocurrency Network," in *ICDCN Conference*, 2021.

[7] S. Ben Mariem, et al., "All that Glitters is not Bitcoin – Unveiling the Centralized Nature of the BTC (IP) Network,", in *IEEE/IFIP NOMS Symposium*, 2020.