Infrastructure Concentration and the Operational Independence of the Bitcoin Peer-to-Peer Network: A Multi-Source Measurement Study

Tuna Kılıç <tuna@tunakilic.com>

Measurements: 3–6 October 2026

Abstract

Bitcoin's protocol does not depend on any single operator, but its nodes still run on machines and networks that belong to companies in particular countries. We measure whether that infrastructure is concentrated, and if so, whether the concentration matters. Our data come from three crawlers observed on 3–4 October 2026 (our own crawler run from two vantage points, Bitnodes, and the KIT DSN monitor), with two of them repeated on 6 October, complemented by an estimate of non-listening nodes and offline attribution with CAIDA, RIR, PeeringDB and APNIC data. Only 34.5% of reachable hosts use IPv4 or IPv6 and can be attributed at all. Among those, the largest provider (Hetzner) hosts 11.0% of nodes, the provider-level HHI is 267, and roughly 43.1% run on commercial hosting; the three crawlers agree closely on these figures, and they barely move between the two measurement days. Countries are more concentrated, with 30.7% of nodes in the US. At the AS level, concentration and the hosted share have both fallen since their 2020 peak. We also find that a single anomalous fleet can change which provider appears largest, and that 7.2% of reachable nodes follow a forked chain. In a topology model based on Bitcoin Core's connection rules, even the loss of all hosted nodes leaves the network connected. If current estimates of non-listening nodes are right, however, the remaining reachable nodes cannot serve part of the connection demand: about 14.0% with no failure, 21.1% without the largest provider and 32.3% without the five largest. Infrastructure concentration therefore does not amount to control over consensus. Its practical consequences are capacity pressure and exposure to network-level attacks.

1 Introduction

Bitcoin was designed so that no trusted intermediary is needed to agree on the state of the ledger [30]. Every full node checks every block against the consensus rules. Those nodes, however, are physical machines in data centres and homes, reached through autonomous systems (ASes) that belong to commercial firms in specific jurisdictions. A protocol can be permissionless and still run on concentrated infrastructure. Earlier work examined how reachable Bitcoin nodes are distributed across ASes and what that implies for centralisation [20, 39], and several papers have shown that routing-level attacks can exploit such concentration [3, 39, 42].

This paper asks to what extent the concentration of Bitcoin nodes in a limited number of hosting providers, ASes, ISPs or countries affects Bitcoin's operational independence and resilience. We split the working hypothesis into three parts so that each can fail on its own:

H1 The infrastructure on which observable Bitcoin nodes run is measurably concentrated, in the sense of non-trivial top-k shares, HHI and Gini at the provider, AS and country level.

H2 This concentration matters for operational resilience: plausible single-provider or single-country outages would substantially degrade node availability, connectivity or propagation.

H3 Infrastructure concentration amounts to control over consensus.

The null hypothesis for H1 is that observable nodes are spread across infrastructure without meaningful concentration. We expected H3 to be difficult to support even if H1 held, and treat it as a claim to examine rather than a premise.

The paper makes five contributions. First, it reports a measurement whose protocol was fixed in writing before the analysis and in which every metric is computed for independent observations from three crawlers, with bounds instead of point estimates wherever the classification is uncertain (§4, §5). Second, it documents two measurement artefacts that can dominate headline figures, an anomalous single-provider fleet and the dependence of crawl results on the vantage point (§6, §10). Third, it shows that the reachable node population of October 2026 is not a single consensus population (§6.2). Fourth, it separates the effects of

Table 1: Properties kept separate throughout the paper, and what an infrastructure provider can do to each.

Property	Meaning and possible influence of an infrastructure provider
Node availability	Whether a node runs and is reachable. A provider can switch off the nodes it hosts.
Peer connectivity	Whether nodes have peers and form one connected network. A provider can cut links to its nodes.
Transaction propagation	Speed and completeness of transaction relay over full-relay links. Affected through connectivity and latency.
Block propagation	Speed of block relay (compact blocks, block-relay-only links, miner relay networks). Affected through connectivity.
Eclipse resistance	How likely it is that one party controls all of a node's peers. A provider or AS can eclipse a node only if it hosts or intercepts all of that node's peers.
Consensus participation	Validation of blocks by full nodes, including non-listening ones. Not determined by hosting.
Mining hash power	Block production by pools and miners, a separate population that we do not measure.
Economic ownership	Who holds coins or runs economically important nodes such as exchanges. Not measured here.
Infrastructure ownership	Who owns the machines and networks. This is what we measure, via AS and organisation records.

infrastructure failures on availability, connectivity, propagation, capacity and eclipse exposure, using a model that follows Bitcoin Core's source code (§7). Fifth, it treats "independence" as several separate dimensions and explains why we do not combine them into one index (§8).

2 Background

Reachable and non-listening nodes. A node is reachable if it accepts inbound connections on a publicly routable address, and only reachable nodes can be enumerated by a crawler. Nodes behind NAT or a firewall, or configured not to listen, open outbound connections only and stay invisible. Their number is hard to pin down. Wang and Pustogarov put it at 10 to 100 times the number of reachable nodes [44], Grundmann et al. derived an estimate from peer announcements [22], and the estimate we use here (D4) puts non-listening nodes at 95.2% of the total [13].

How Bitcoin Core picks peers. The following details were checked in the Bitcoin Core source, both at master commit `66776840be` and at release v31.0 [7].

- A node keeps 8 outbound full-relay and 2 outbound block-relay-only connections (`MAX_OUTBOUND_FULL_RELAY_CONNECTIONS` and `MAX_BLOCK_RELAY_ONLY_CONNECTIONS` in `net.h`), plus short-lived feeler connections. Every release from v27.0 to v31.0 allows 125 connections by default (`DEFAULT_MAX_PEER_CONNECTIONS`); only the master branch uses 200. Subtracting the 11 automatic outbound slots leaves 114 inbound slots in released versions.
- Outbound IPv4 and IPv6 peers must come from different *netgroups* (`net.cpp`), which by default are /16 prefixes for IPv4 and /32 prefixes for IPv6 (`netgroup.cpp`). With `-asmap`, netgroups become ASNs. Default builds embed an asmap (`WITH_EMBEDDED_ASMAP=ON`), but it is used only if the operator sets `-asmap`.
- Once its full-relay slots are filled, a node opens extra connections to keep at least one outbound peer in every network it can reach, i.e. IPv4, IPv6, Tor, I2P and CJDNS (`MaybePickPreferredNetwork`).
- Outbound peers must announce `NODE_WITNESS`, otherwise they are disconnected (`GetDesirableServiceFlags` and `HasAllDesirableServiceFlags` in `net_processing.cpp`).

Because one provider typically owns many /16 prefixes, the default netgroup rule limits how many outbound peers share a prefix, not how many share a provider. This is why eclipse attacks [24] and routing attacks [3, 42] are the threat models in which infrastructure concentration is most relevant.

What we keep apart. Several properties that are often discussed together behave very differently under an infrastructure failure. Table 1 lists the ones we distinguish. A provider can directly affect the first two, it can affect the next three only through the shape of the topology, and it has no direct influence on the remaining four.

3 Related work

Measuring the Bitcoin network. Some of the earliest studies measured the reachable network and how quickly information spreads through it [15, 17]. Feld et al. looked at the network's deployment from an AS-level perspective [20], and Park et al. scanned it for 37 days in 2018 and compared their results with earlier measurements [36]. Neudecker and Hartenstein survey the network layer of permissionless blockchains, including measurement methods [32]. Bitcoin hides its topology on purpose, so inferring it requires indirect techniques such as timing analysis [33], orphan transactions [16] or other forms of active probing [18]; we therefore model the topology instead of measuring it. Estimates of peer degree and of the size of the reachable and unreachable network are given in [22, 23, 44].

Measuring decentralisation. Gencer et al. measured the network resources and interconnection of Bitcoin and Ethereum nodes as indicators of decentralisation [21]. Kwon et al. showed that without a positive Sybil cost, good decentralisation of a permissionless blockchain is almost impossible to reach [27]. Sai et al. proposed a taxonomy of centralisation across layers [40], and Lin et al. compared several metrics at different time granularities [28]. The "Nakamoto coefficient" was popularised by Srinivasan and Lee [41]. We rely mainly on the HHI, using the antitrust thresholds of [43] purely as a point of reference, and on the Gini coefficient [12]. Kim et al. carried out a comparable measurement of Ethereum peers [26].

Attacks at the network layer. An eclipse attack takes over all of a victim's connections [24]. ASes that hijack or intercept BGP routes can partition the network or delay blocks [3], a single transit AS can do so stealthily [42], and partitioning attacks can combine spatial, temporal and logical properties of the network [39]. Clients can be deanonymised at the P2P layer [6], and running Bitcoin over Tor brings attack surface of its own [5]. Bitcoin Core's asmap feature is meant to address the AS-level threat. Transaction relay efficiency is studied in [31], and mining incentives, a separate dimension of centralisation, in [19].

Concentration elsewhere on the Internet. Outside cryptocurrencies, researchers have measured how websites depend on a few third-party providers [25] and how much of the Internet the large cloud providers can reach directly [4]. Classifying who owns an AS is itself error-prone [45], and so is IP geolocation [38]; both kinds of error carry over into our results (§10).

4 Data

Table 2 summarises the four node datasets. We never rely on a single node-count service, and each result is reported for every dataset that can support it. Two days after the main measurement, on 6 October, we took a second Bitnodes snapshot and ran the AWS crawl again, to see how much the results depend on the day (§6.9). Raw files, URLs, retrieval times and SHA-256 hashes are listed in `data/raw/MANIFEST.csv`.

To attribute addresses we combine CAIDA's RouteViews prefix-to-AS tables for IPv4 and IPv6 (30 September 2026) [11], CAIDA's AS-to-organisation mapping (1 September 2026) [9], the NRO extended delegation statistics for registration country [35], DB-IP Lite for geolocated country (October 2026) [14], PeeringDB network types [37] and APNIC's per-AS user estimates [2]. For cloud addresses we use the ranges that AWS, Google Cloud, Oracle, DigitalOcean and Akamai/Linode publish themselves, and CAIDA's collection of cloud prefixes for Azure, IBM, Cloudflare and Fastly [10].

5 Method

We wrote the measurement protocol down on 4 October 2026, while our first crawl was running and before computing any concentration metric (`analysis/00_methodology_preregistration.md`). The protocol was not deposited with a third-party registry, so its timing rests on our own statement; the file also discloses the one look at the data that preceded it. Everything added later is marked as such below.

Unit of analysis. We count unique IP addresses that completed a version handshake. Tor, I2P and CJDNS endpoints cannot be tied to any infrastructure and are kept as a separate group. As a sensitivity check we also aggregate addresses to IPv4 /24 and IPv6 /48 subnets.

Table 2: Node datasets. "Reachable only" means that the dataset contains only nodes that accept inbound connections.

	D1 / D1b own crawls	D2 Bitnodes/btcnodes [29]	D3 KIT DSN [34]	D4 Luke Dashjr [13]
Collected	2026-10-04; D1 home vantage (Türkiye), 3.0 h; D1b AWS EC2, 2.9 h; AWS crawl repeated 2026-10-06 (IPv4)	2026-10-04 09:12 UTC; repeated 2026-10-06 14:25 UTC	2026-10-03 (daily dossier)	2026-10-04 (file date)
Method	BFS from 8 Core DNS seeds; version handshake + `getaddr`; gossip ≤7 days old; 1 retry; concurrency 300 (D1) / 250 (D1b)	Crawler snapshot (open-source Bitnodes deployment)	Long-running monitor connected to reachable peers	Not published; per-user-agent counts
Size	D1: 4,804 IPs (3,075 v4, 1,729 v6) of 291,336 probed; D1b: 8,685 IPs (6,956 v4 from AWS of 316,351 probed, 1,729 v6 from D1)	25,019 hosts (8,637 IP)	8,959 peers (7,142 v4, 1,817 v6)	5,297 listening + 105,153 est. non-listening
Reachable only	yes	yes	yes	no (estimate)
Tor / I2P	advertised addresses only (D1b: 20,787 / 5,384), not verified	verified (12,246 / 4,131)	no	unknown
IPv4 / IPv6	D1: both, IPv4 undercovered (§10); D1b: AWS IPv4 (instance had no IPv6) + D1 IPv6	both	both	unknown
Duplicates	IP:port deduplicated to IP	IP:port deduplicated to IP	anonymised IDs; no IPs	unknown
Known biases	D1: NAT saturation; D1b: cloud vantage point; IPv4 and IPv6 from different vantage points	crawler-specific discovery; bitnodes.io redirects to btcnodes.io, a deployment of the open-source crawler	monitor-specific peer set; ASN from KIT whois	opaque methodology

Attribution. Each IP is mapped to its origin ASN by longest-prefix match and to an organisation through CAIDA's AS2Org data. Because some companies appear under several AS2Org records (Amazon has two, for example), we also report a *provider* label that merges such records through an explicit table published with the code. Merging can only make concentration look higher, so results are reported both with and without it. Country is reported twice, once by geolocation and once by RIR registration.

Hosting or residential. Every node is put into one of four classes. H1 covers IPs inside a cloud provider's published ranges. H2 covers ASes that PeeringDB lists as *Content*, that appear on our list of hosting companies, or whose name contains a hosting keyword. R covers ASes with at least 10,000 estimated users according to APNIC, or listed as *Cable/DSL/ISP* in PeeringDB, which in practice means home and small-business access networks. Everything else falls into U. We report the hosted share as a range from H1 alone to H1+H2+U, with H1+H2 as the central estimate.

Metrics. For a dimension in which group g holds a share s_g of attributable nodes, we compute four measures. The *top-k share* is the share held by the k largest groups and says directly how much depends on them. The *HHI*, $10^4 \sum_g s_g^2$, gives large groups more weight; its reciprocal $N_{\text{eff}} = 10^4/\text{HHI}$ is the number of equally sized groups that would yield the same HHI. The *Gini coefficient* measures inequality among the groups that exist. It is strongly affected by the long tail of groups with a single node, so we always report it together with the number of groups. Finally, N_{50} (N_{33}) is the smallest number of groups that together hold at least half (a third) of the nodes. HHI and top-k shares tell us how much the failure of a few entities would remove, and the Gini coefficient tells us how unequal the distribution is. None of them measures power over consensus. The HHI threshold of 1,800 in the U.S. merger guidelines [43] was set for product markets, and we mention it only as a familiar reference point.

Failure scenarios. We consider the loss of the largest ASN, the largest provider, the largest country, the top-5 and top-10 providers, and all hosted (H1+H2) nodes. For each scenario we first count how many observed nodes would disappear, relative to three different totals. We then run a model. Each node draws 8+2 outbound peers at random from the reachable nodes, subject to the netgroup rule (/16 and /32, or ASN under asmap); 25 % of nodes are assumed to use an overlay network as well, and nodes keep one peer in every network they can reach, as Core does. The model is run with no non-listening nodes and with the 105,153 non-listening nodes estimated by D4. We record how many nodes are left without outbound peers, the size of the largest connected component (LCC), the mean number of hops before and after nodes reconnect as a

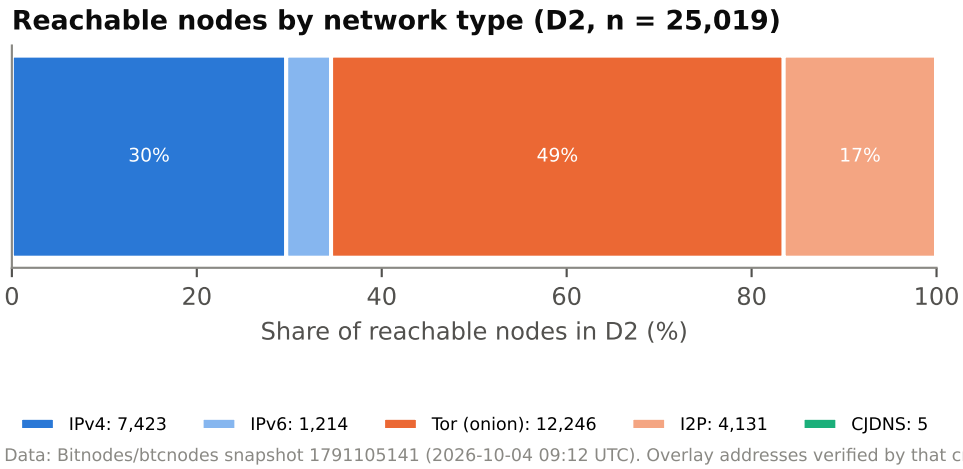


Figure 1: Reachable nodes by network type (D2). Only the IPv4 and IPv6 parts can be attributed to ASes, providers or countries.

rough proxy for propagation delay, how much inbound demand exceeds the default capacity (114 slots in released versions, with master's 189 as a second case), and how often a node's peers are concentrated in one provider. The model population consists of nodes that announce `NODE_WITNESS` and report a height on the SHA256d chain (§6.2). All model results are averages over three random seeds.

Changes to the protocol. The data contained several things the protocol had not anticipated, and we made four additions in response. We added a *functional* filter that excludes endpoints without the `NODE_WITNESS` bit, since current Core nodes will not keep them as outbound peers. We classified nodes by the start height they report, relative to the most common height in the same 15-minute window (the *chain view*). We re-probed, at low concurrency, IPv4 addresses that our home crawl had missed, to find out why it missed them. Once that showed the gap was an artefact, we ran the same crawler a second time from an AWS EC2 instance without NAT (D1b), starting about 1.5 hours after the first crawl had finished. Table 4 reports the unfiltered results specified in the protocol next to the filtered ones.

6 Results

6.1 How much can be observed

Bitnodes reports 25,139 reachable endpoints on 25,019 unique hosts, but only 8,637 of these hosts (34.5%) use IPv4 or IPv6 (Fig. 1). Tor alone accounts for 48.9%. If the D4 estimate is roughly right, all reachable hosts together make up 19.2% of the nodes in the network, and the IPv4/IPv6 hosts that we can attribute make up only 6.6%. Everything we say about providers and countries below refers to that fraction.

One group of nodes stands out. 849 IPv4 endpoints in D2, or 9.8% of its IP hosts, look exactly alike: user agent `/Satoshi:27.0.0/`, service bits 3077 (without `NODE_WITNESS`) and a reported height of 0. All 849 are on Akamai/Linode, spread over 200 /24 subnets and 23 /16 netgroups. KIT's dossier contains only 15 endpoints with this fingerprint. Our home crawl came across 91 of the addresses in gossip, and the AWS crawl 138, all of which completed a handshake. A genuine Bitcoin Core 27.0 node always announces `NODE_WITNESS`, so we think the fleet is most likely run by a single party for some purpose other than serving the network. It matters for the results: counted as ordinary nodes, it makes Akamai/Linode the largest provider in D2 (10.2%), whereas without it Hetzner is the largest (11.0%). A study based on one crawler alone would have no way of noticing this.

6.2 Reachable nodes follow more than one chain

Table 3 groups all reachable D2 hosts by the start height they report during the handshake. 22,168 hosts (88.6%) are within six blocks of the most common height, 969,818. Another 1,070 (4.3%) report heights far above it, by a median of 5,647 blocks; 1,068 of them run Bitcoin Knots, nearly all version 29.4.1 or 29.4.2. A further 720 (2.9%) report heights between 961,600 and 961,700 (median 961,639), and almost all of these

Table 3: Start height reported in the version handshake by all reachable nodes in D2 (IPv4, IPv6, Tor, I2P, CJDNS; unique hosts), relative to the modal height of the snapshot.

Height relative to SHA256d tip	Bitcoin Core	Bitcoin Knots	other	Total	Share (%)
at tip (±6 blocks)	19820	2284	64	22168	88.6
below tip (lagging/syncing)	188	17	5	210	0.8
above tip (>6 blocks)	1	1068	1	1070	4.3
961,600–961,700	4	716	0	720	2.9
zero	849	0	2	851	3.4

run Knots 29.3 or 29.4.0 or a BIP-110 build. Our own crawls show the same pattern: 198 nodes above the tip and 61 near 961,6xx from home, and 300 and 123 among the IPv4 nodes seen from AWS.

According to the Knots release notes, version 29.4.1 introduces a backward-incompatible switch to a BLAKE2b proof of work, and version 29.3 adds the BIP-110 "Reduced Data Temporary Softfork" [8]. Press reports describe a BIP-110 chain that stopped after splitting off in August 2026 [1]. Heights far ahead of the SHA256d tip and heights frozen near 961,6xx fit nodes following these two chains. We did not validate the chains ourselves, so this reading rests on the release notes and on the heights the nodes report.

Either way, several thousand reachable nodes enforce rules that the dominant chain does not, and every crawler counts them anyway. They also use the same infrastructure: 145 of the 194 providers that host the 479 forked IPv4/IPv6 nodes also host main-chain nodes, and 170 of the forked nodes sit at one of the ten largest main-chain providers, although they are more often on access networks (68.3% against 49.7%). Where a node is hosted tells us nothing about which chain it validates, which is also why a node count is not a measure of consensus.

6.3 Geographic distribution

Functional D2 nodes are located in 94 countries. The US has 30.7% of them, followed by DE with 18.1% and CA with 4.7%. The five largest countries hold 62.1% of nodes, the country HHI is 1,397 and $N_{50} = 3$ (Fig. 2). By registration rather than geolocation, the US share rises to 38.9% (HHI 1,997), and KIT, which records the country in which an AS is registered, gives 39.1% (HHI 1,997).

Country is the most concentrated dimension we measured: the two largest countries together hold 48.8% of attributable nodes, and 3 countries are enough to reach half. The country of a node is itself uncertain, though. Geolocation and registration agree for 79.2% of nodes overall but for only 38.4% of nodes in cloud ranges, compared with 97.0% for residential nodes, because cloud address space registered in one country is often used in another.

6.4 ASes and providers

Functional D2 nodes are spread over 1,261 ASNs and 1,135 providers (Table 4, Figs. 3–5). Hetzner hosts 11.0% of them, followed by OVHcloud (5.1%), Google (4.9%), Amazon (4.2%) and Charter (3.5%); Table 5 lists the ten largest, which together host 43.3%. The provider HHI is 267 ($N_{\text{eff}} = 37.5$), the ASN HHI is 226 ($N_{\text{eff}} = 44.3$, see also Fig. 4), and half of the attributable nodes are covered by $N_{50} = 15$ providers or 26 ASNs. KIT's data give almost the same picture, with Hetzner at 10.1%, a provider HHI of 264, an ASN HHI of 221 and $N_{50} = 14$ (Fig. 6). The Gini coefficient across providers is high (0.78) because of the long tail of providers that host only one node.

These numbers support H1, with a qualification. Concentration is clearly there, since 15 providers host half of the attributable nodes and the largest one hosts about one in ten. In absolute terms, however, it is moderate. The provider and AS HHIs are far below the 1,800 reference level, while the country HHI comes close to it under geolocation and exceeds it under KIT's registration-based country. The high Gini and the moderate HHI describe the same situation from two sides: many very small groups next to a handful of medium-sized ones, and no group that dominates.

6.5 Hosting and residential connections

Of the functional D2 nodes, 12.3% lie in published cloud ranges (H1), 30.8% in hosting ASes (H2) and 50.9% in access networks (R), while 6.0% could not be classified (Table 6, Fig. 7). The hosted share is therefore somewhere between 12.3% and 49.1%. KIT's data, which can only be classified at the AS level, give 42.1% hosted and 51.8% residential.

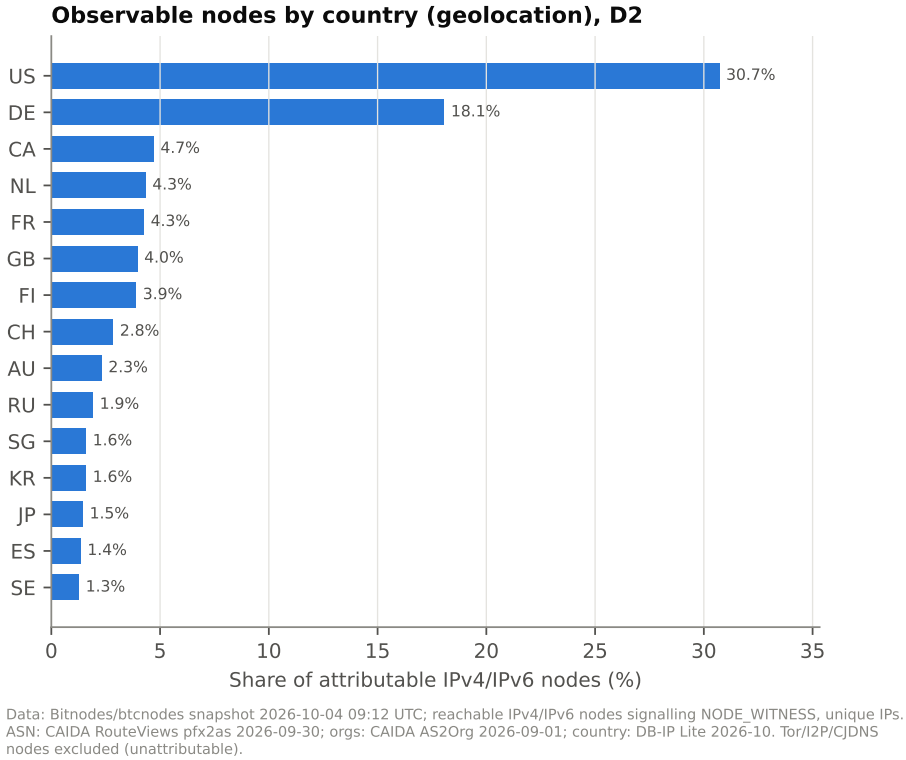


Figure 2: Functional IPv4/IPv6 nodes by geolocated country (D2).

So a large minority of attributable nodes, probably a little under half, runs on commercial hosting, and roughly half runs on home or small-business connections. The R class is an inference made at the AS level; a node in an access network can still sit in a company's server room, and a node at a small VPS provider that is missing from our lists ends up in U. Our name-based rule matched 134 nodes (1.7%), including some false positives such as ISP names that contain "colo", so its effect on the central estimate is at most that large.

6.6 Client software

Bitcoin Core runs on 87.8% of functional IP nodes. Bitcoin Knots accounts for 16.3% of all reachable hosts and 19.5% of the Tor and I2P hosts, and it is far more common on residential addresses (15.9%) than in cloud ranges (2.8%). By the HHI, client diversity is low (7,836). The minority client has grown quickly, though: in KIT's data it stayed below one percent until 2024 and reached 11.4% in 2026. Since that client is now also the one behind the chain splits described in §6.2, more software diversity has in this case come with less agreement on the rules.

6.7 Change over time

In KIT's annual snapshots (Table 7, Fig. 8), the ASN HHI went from 149 in 2015 up to 408 in 2020 and back down to 219 in 2026. The share of peers in hosting ASes peaked at 61.3% in 2020 and is now 42.4%. Hetzner has been the largest provider in every year since 2016. Over the seven days before our crawl, none of the KIT metrics moved by more than 3.5% relative to its mean, so our snapshot is not an unusual day. Over the longer run, concentration of the reachable network has not simply kept rising; by these measures it is lower today than in 2020–2022.

6.8 Agreement between datasets

The AWS crawl reached 85.9% of the IPv4 hosts listed by D2, against 38.5% for the home crawl, and it found 581 IPv4 hosts that D2 does not list. D1b and D2 share 7,493 IPs, with 1,192 seen only by D1b and 1,144 only by D2, a Jaccard similarity of 0.76 (0.42 for the home crawl alone). For functional nodes, D1b identifies the same largest provider (Hetzner at 10.6%) and gives a provider HHI of 293, an ASN HHI of 250, $N_{50} = 13$ providers, a country HHI of 1,355 and a central hosted share of 46.1%, with 48.2% residential. The home crawl alone, affected by NAT, gives a higher provider HHI (379) and a higher hosted share (51.7%),

Table 4: Concentration of observable IPv4/IPv6 nodes (suffix f: NODE_WITNESS nodes only). Top-k shares in % of attributable nodes; N_eff = 10,000/HHI; N50 = smallest number of groups jointly holding at least 50% of attributable nodes.

Data	Dimension	n	Groups	Top-1	Top-5	Top-10	HHI	N_eff	Gini	N50
D1	Country (geolocation, DB-IP)	4804	82	27.5	60.8	76.7	1268	7.9	0.8	3
D1	ASN	4804	864	12.4	35.0	44.8	336	29.8	0.8	15
D1	Organisation (CAIDA AS2Org)	4804	803	12.7	35.9	46.6	356	28.1	0.8	13
D1	Provider (sibling-merged)	4804	780	12.7	36.5	48.7	373	26.8	0.8	11
D1b	Country (geolocation, DB-IP)	8685	92	30.6	61.7	76.4	1373	7.3	0.8	3
D1b	ASN	8685	1285	10.2	29.4	40.8	247	40.6	0.8	21
D1b	Organisation (CAIDA AS2Org)	8685	1195	10.4	30.5	42.7	267	37.5	0.8	16
D1b	Provider (sibling-merged)	8685	1166	10.4	31.3	45.2	288	34.7	0.8	13
D2	Country (geolocation, DB-IP)	8637	94	31.6	64.8	78.8	1491	6.7	0.8	2
D2	ASN	8637	1264	10.2	31.4	42.9	287	34.8	0.8	18
D2	Organisation (CAIDA AS2Org)	8637	1169	10.2	32.3	44.8	304	32.9	0.8	14
D2	Provider (sibling-merged)	8637	1138	10.2	32.9	47.0	320	31.2	0.8	12
D3	Country (AS registration, KIT whois)	8946	96	39.2	71.2	80.6	2002	5.0	0.9	2
D3	ASN	8924	1300	9.8	27.3	38.6	219	45.6	0.8	25
D3	Organisation (CAIDA AS2Org)	8924	1197	10.1	28.3	40.8	239	41.8	0.8	19
D3	Provider (sibling-merged)	8924	1168	10.1	29.3	43.6	262	38.2	0.8	14
D1f	Country (geolocation, DB-IP)	4743	82	27.4	60.5	76.4	1259	7.9	0.8	3
D1f	ASN	4743	863	12.6	35.4	45.2	341	29.3	0.7	16
D1f	Organisation (CAIDA AS2Org)	4743	802	12.8	36.3	47.2	361	27.7	0.8	12
D1f	Provider (sibling-merged)	4743	779	12.8	36.9	49.2	379	26.4	0.8	11
D1bf	Country (geolocation, DB-IP)	8523	92	30.4	61.2	76.1	1355	7.4	0.8	3
D1bf	ASN	8523	1283	10.3	29.9	40.7	250	39.9	0.8	22
D1bf	Organisation (CAIDA AS2Org)	8523	1193	10.6	31.0	43.3	272	36.8	0.8	17
D1bf	Provider (sibling-merged)	8523	1164	10.6	31.8	45.9	293	34.1	0.8	13
D2f	Country (geolocation, DB-IP)	7771	94	30.7	62.1	77.0	1397	7.2	0.8	3
D2f	ASN	7771	1261	10.7	26.8	38.1	226	44.3	0.8	26
D2f	Organisation (CAIDA AS2Org)	7771	1166	11.0	27.8	40.5	247	40.6	0.8	20
D2f	Provider (sibling-merged)	7771	1135	11.0	28.8	43.3	267	37.5	0.8	15
D3f	Country (AS registration, KIT whois)	8861	95	39.1	71.2	80.6	1997	5.0	0.9	2
D3f	ASN	8839	1294	9.8	27.5	38.7	221	45.3	0.8	25
D3f	Organisation (CAIDA AS2Org)	8839	1191	10.1	28.4	41.0	241	41.5	0.8	19
D3f	Provider (sibling-merged)	8839	1162	10.1	29.5	43.7	264	37.9	0.8	14

as expected from its bias (§10). Three crawlers with different code, vantage points and discovery methods thus agree closely on who the largest provider is, on its share and on the provider and ASN HHIs, so the concentration figures do not depend on one particular crawler.

6.9 A second measurement day

On 6 October we repeated the Bitnodes snapshot and the AWS crawl; the AWS machine again had no IPv6, so that comparison covers IPv4 only (Table 8). The concentration figures hardly changed. In Bitnodes, Hetzner went from 11.0% to 10.8% and the provider HHI from 267 to 261, and the hosted share moved from 43.1% to 42.6%. In the AWS crawl the provider HHI was 221 on both days. Across both datasets the provider HHI changed by at most 6 points.

The node population itself turned over more. Of the functional IP nodes in the first AWS crawl, 93.0% were reachable again two days later, while 525 nodes appeared that had not been seen before. Hosted nodes were more stable (96.7% seen again) than nodes on access networks (90.1%). Bitnodes shows a lower overall retention of 88.7% and a less clear pattern by class (82.2% for H1, 92.0% for H2, 87.9% for R), which may partly reflect day-to-day differences in what its crawler reaches. Either way, individual nodes come and go noticeably faster than the overall distribution changes.

The Linode fleet shrank from 849 to 617 endpoints in Bitnodes, 610 of which had the same addresses on both days. The number of nodes reporting heights above the SHA256d tip rose slightly, from 1,070 to 1,103 (AWS crawl: 300 to 313), while the number stuck near 961,6xx stayed flat (720 and 721).

7 Failure scenarios

Counting the losses. Table 9 and Fig. 9 start from the 7,291 functional IPv4/IPv6 nodes in D2 that are on the SHA256d chain. If Hetzner went offline, 842 of them would disappear. That is 11.5% of the

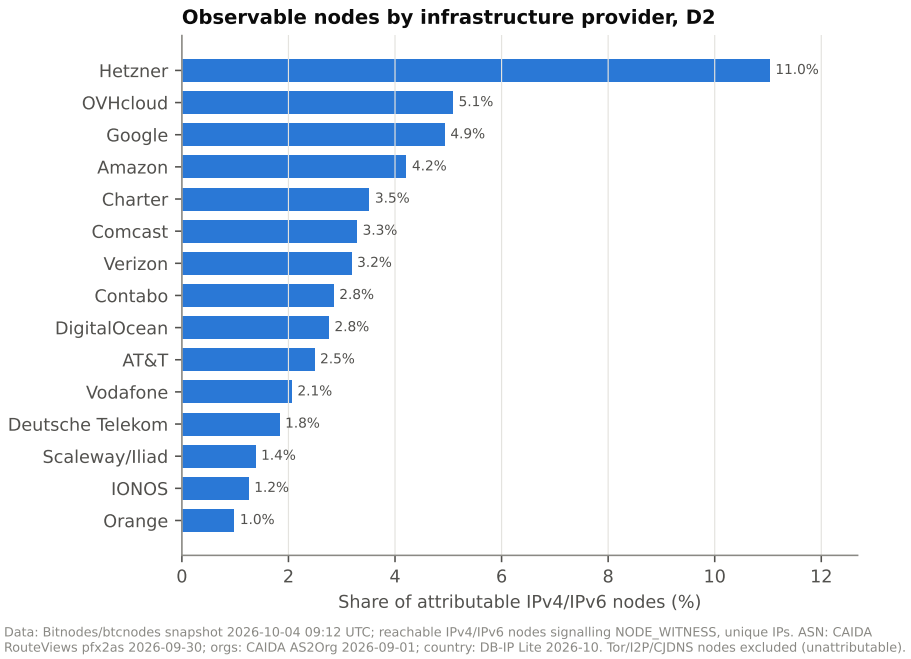


Figure 3: Functional IPv4/IPv6 nodes by provider (sibling-merged CAIDA organisations), D2.

Table 5: Ten largest infrastructure providers (sibling-merged CAIDA organisations) among functional (NODE_WITNESS) IPv4/IPv6 nodes in D2.

Group	Nodes	Share (%)	Cum. (%)
Hetzner	857	11.0	11.0
OVHcloud	395	5.1	16.1
Google	383	4.9	21.0
Amazon	327	4.2	25.2
Charter	273	3.5	28.8
Comcast	255	3.3	32.0
Verizon	247	3.2	35.2
Contabo	221	2.8	38.1
DigitalOcean	214	2.8	40.8
AT&T	193	2.5	43.3

IP-reachable nodes and 3.8% of all reachable nodes; relative to the whole network including the D4 estimate, it is between 0.66% (if non-listening nodes are unaffected) and 10.2% (if they are affected in the same proportion). Losing the top-5 providers would remove 29.5% of IP-reachable nodes, losing the US 29.6%, and losing every hosted node 44.2%. Even then, at least 954 ASNs and 92 countries would still have nodes.

Connectivity in the model. With the D4 estimate included, a message needs 3.74 hops on average to reach a node in the model network (3.88 over full-relay links only). Each reachable node receives 57.0 inbound connections on average and at most 187, against a default capacity of about 114. Removing every hosted node at once leaves 0.01% of the remaining nodes without any outbound peer and 0.08% without a full-relay peer, and the largest connected component still contains 99.99% of them. The average path grows to 3.95 hops before nodes reconnect and falls to 3.65 afterwards. In no scenario do more than 0.01% of nodes lose all outbound peers, and the LCC never drops below 99.99%. If we assume that no node uses both clearnet and an overlay network, the LCC after losing all hosted nodes is 87.9%, but it is already 88.2% without any failure, because under that assumption the Tor- and I2P-only nodes form a separate network from the start (Fig. 10).

Inbound capacity. Since only reachable nodes accept connections, they have to absorb the outbound connections of all other nodes. Table 10 shows what share of the outbound demand would exceed the inbound slots of the nodes it is aimed at. Without non-listening nodes this never happens (0%). With the D4 estimate and the 114 slots of released versions, 14.0% of the demand cannot be served even when nothing

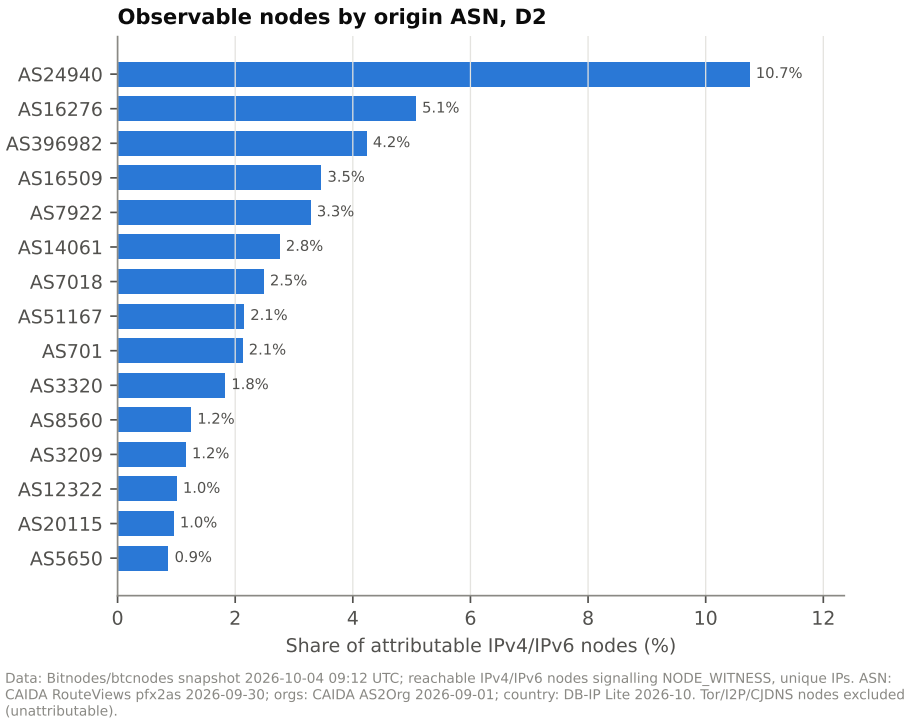


Figure 4: Functional IPv4/IPv6 nodes by origin ASN, D2.

Table 6: Infrastructure class of attributable IPv4/IPv6 nodes (% of nodes). D3 has no IP addresses, so H1 (published cloud ranges) cannot be applied there; its cloud nodes fall into H2.

Data	n	H1	H2	R	U	Hosted low	Hosted central	Hosted high
D1	4804	19.5	32.7	42.0	5.7	19.5	52.2	58.0
D1b	8685	17.8	29.2	47.3	5.6	17.8	47.0	52.7
D2	8637	20.9	27.8	45.8	5.5	20.9	48.7	54.2
D3	8959	0.0	42.1	51.7	5.7	0.0	42.1	47.9
D1f	4743	18.6	33.0	42.6	5.8	18.6	51.7	57.4
D1bf	8523	16.5	29.6	48.2	5.7	16.5	46.1	51.8
D2f	7771	12.3	30.8	50.9	6.0	12.3	43.1	49.1
D3f	8874	0.0	42.1	51.8	5.7	0.0	42.1	47.8

has failed. Losing the largest provider raises this to 21.1%, losing the top-5 providers or the largest country to about 32.3%, and losing all hosted nodes to 41.6%. With master's 189 slots the corresponding figures are 0%, 0.01%, 3.1% and 17.1%. A node that is turned away would try other peers, so these numbers describe how much pressure the remaining nodes are under rather than how many connections are lost for good.

Eclipse exposure. With the default /16 netgroups, 0.20% of modelled nodes get most of their clearnet outbound peers from a single provider, and 1.5% get at least half of their full-relay peers from Hetzner. With ASN-based netgroups (`-asmap`) these shares drop to <0.01% and 0.21%. Nodes whose clearnet peers all belong to one provider are rare either way, <0.01% with /16 netgroups (a handful in a population of over 100,000) and 0% with asmap.

What the scenarios show. The model gives no support to H2 as far as connectivity is concerned. Even the extreme and physically implausible loss of every hosted node at once leaves the rest of the network connected, because outbound connections are spread at random over thousands of netgroups and non-listening nodes connect to reachable nodes of every kind; paths become only slightly longer. Capacity is another matter. With the D4 estimate the model network is saturated before anything fails, which means that either D4 overstates the number of non-listening nodes or reachable nodes are in fact often full and the

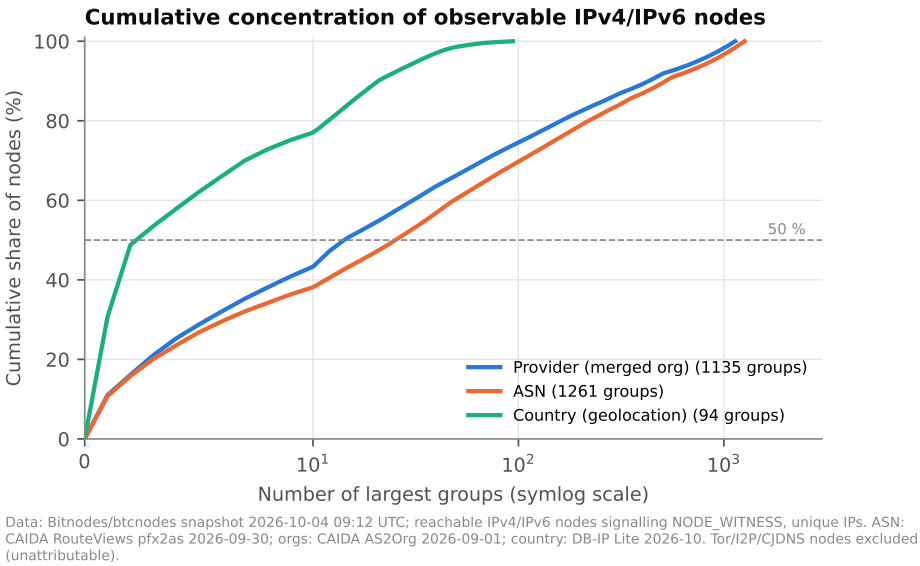


Figure 5: Cumulative share of functional IPv4/IPv6 nodes held by the k largest providers, ASNs and countries (D2).

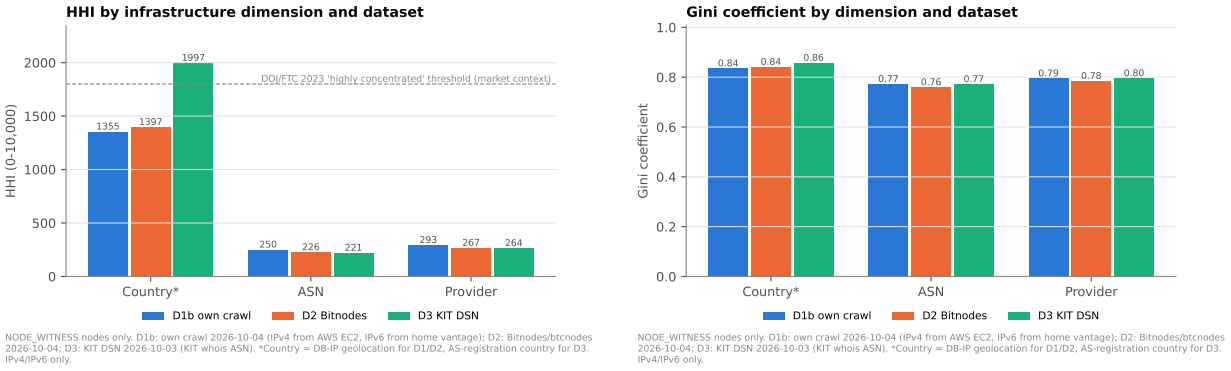


Figure 6: HHI and Gini coefficient per dimension for the three datasets (functional nodes). The dashed line marks the 1,800 HHI reference of [43].

rest get by through retries, overlay networks or higher limits set by operators. Our data cannot tell these apart. Grundmann et al., however, measured that about half of reachable peers run with the default limit of 125 connections and that nearly all connection slots are in use [23], which fits a network running close to saturation. What the model shows more robustly is the change. Losing a large share of reachable IPv4/IPv6 capacity, such as the top-5 providers or the largest country, adds about 18 percentage points of unmet demand under the defaults of released versions. Hosted reachable nodes therefore supply a good part of the capacity that the non-listening majority relies on, and in that respect H2 is partly supported. Without non-listening nodes the effect disappears. The model leaves out miner relay networks, latency and the non-uniform way in which addrman really selects peers. None of the scenarios changes the rules that the remaining nodes enforce, so losing nodes is not the same as losing control. Concentration matters most where someone acts on purpose: a provider or AS that is malicious, or forced to cooperate, could tilt the peer sets of some nodes towards itself, and in the model asmap makes this considerably harder.

8 Dimensions of independence

Table 11 brings the measured values together, one row per dimension. Combining them into a single "Infrastructure Independence Index" would require a validated model of how each dimension affects some resilience outcome. We did not find such a model in the literature, and our data do not provide one, so any set of weights would simply reflect our own judgement. In line with the rule we fixed beforehand, we report the dimensions separately and do not compute a combined score.

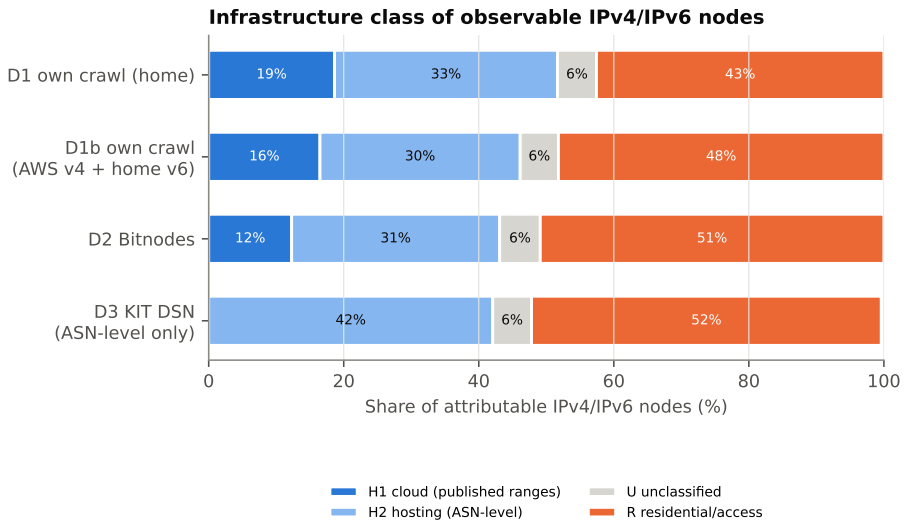


Figure 7: Infrastructure class of functional IPv4/IPv6 nodes.

Table 7: Infrastructure diversity of peers in KIT DSN dossiers (1 October of each year, IPv4/IPv6). Hosted % = share of peers in ASNs classified as hosting (H2) with 2026 classification sources.

Date	Peers	ASNs	ASN HHI	Top-10 ASN %	Top provider	%	Hosted %	Top AS-cc	%	Knots %
2015-10-01	4934	1032	149	29.7	Comcast	7.7	29.4	US	39.1	0.0
2016-10-01	5419	941	203	34.8	Hetzner	8.3	43.5	US	33.8	0.5
2017-10-02	8868	1224	249	39.4	Hetzner	10.5	51.6	US	33.9	0.5
2018-10-01	10158	1312	263	42.1	Hetzner	9.8	52.7	US	33.0	0.4
2019-10-01	10021	1192	333	46.0	Hetzner	12.0	57.6	US	38.6	0.2
2020-10-01	8843	1083	408	50.5	Hetzner	13.4	61.3	US	36.5	0.4
2021-10-01	9858	1223	358	46.2	Hetzner	13.3	53.9	US	39.1	0.3
2022-10-01	7679	1001	390	47.7	Hetzner	12.5	53.5	US	40.2	0.5
2023-10-01	6895	985	293	42.1	Hetzner	12.8	50.2	US	36.7	0.3
2024-10-01	7804	1104	265	41.0	Hetzner	12.3	50.1	US	37.3	0.8
2025-10-01	9886	1324	228	40.7	Hetzner	9.7	43.9	US	39.3	9.9
2026-10-01	9011	1309	219	38.5	Hetzner	10.0	42.4	US	39.1	11.4

9 Counter-arguments

We set out to look for evidence against our own hypothesis, so we went through the usual objections one by one and asked whether each would change our conclusions.

1. *Nodes at the same provider may belong to different people.* We agree, and our data cannot test it. AS2Org tells us who owns the network, not who runs the node. This weakens any reading of provider shares as ownership, but it does not affect the availability analysis, so it changes how the numbers should be read rather than the numbers themselves.
2. *Concentrated hosting is not concentrated ownership.* Again we agree, and we only measure hosting (Table 1). The Linode fleet (§6.1) shows that the reverse can also happen, with one operator appearing as hundreds of nodes.
3. *Consensus does not require a majority of public nodes.* This is correct. Every full node validates on its own, including non-listening ones, and on its own this argument is enough to reject H3.
4. *Reachable nodes are not all nodes.* D4 suggests that they are a minority. If the non-listening majority is mostly residential, our hosted shares overstate hosting in the network as a whole. We cannot check this, which is why we report two bounds (assumptions A and B).
5. *Tor nodes may be missing.* We do count Tor and I2P nodes, but we cannot say where they are hosted, and they make up a large share of reachable hosts. Tor relays themselves run somewhere, and we did not measure how concentrated they are.

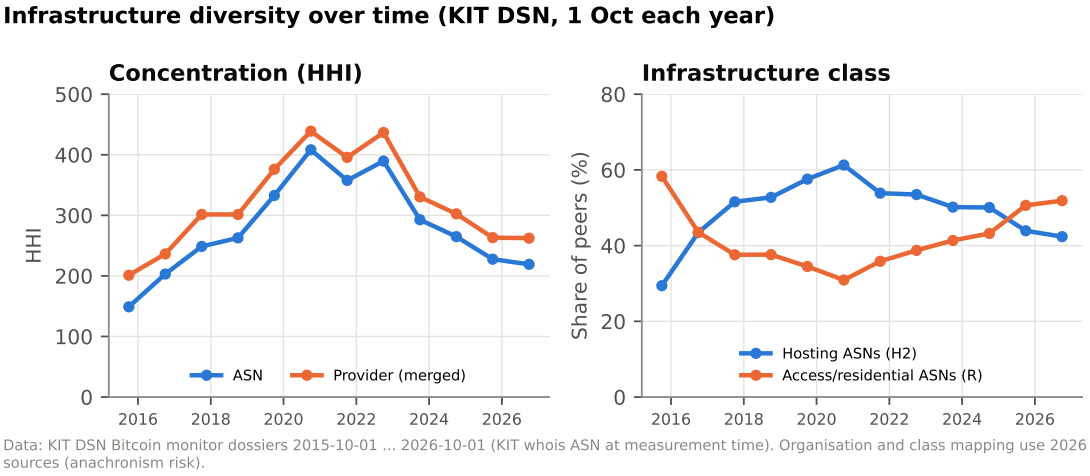


Figure 8: Infrastructure diversity over time (KIT DSN). Organisation and class mapping use 2026 sources.

Table 8: The same measurements on 4 and 6 October 2026 (functional IPv4/IPv6 nodes for D2, functional IPv4 nodes for the AWS crawl). Hosted % is the central estimate H1+H2.

Dataset	Day	IP nodes	Top provider	Top %	Top-5 %	Prov. HHI	ASN HHI	Country HHI	Hosted %
Bitnodes (D2)	4 Oct	7771	Hetzner	11.0	28.8	267	226	1397	43.1
Bitnodes (D2)	6 Oct	7792	Hetzner	10.8	28.4	261	220	1387	42.6
AWS crawl, IPv4 (D1b)	4 Oct	6796	Hetzner	8.3	26.2	221	177	1474	40.1
AWS crawl, IPv4 (D1b)	6 Oct	6842	Hetzner	8.3	26.2	221	176	1478	39.8

6. *Private nodes cannot be observed.* They enter our analysis only through the D4 estimate, whose method has not been published.
7. *Running a node is cheap and operators can move.* We did not measure costs. The time series does show a lot of movement, with hosted shares rising and falling by more than ten percentage points within a decade. This limits how long any dependence on a provider lasts, but it does not help during a sudden outage.
8. *Cloud providers have no direct control over consensus.* This matches our analysis. A provider can switch off nodes or interfere with traffic, but it cannot change the rules that other nodes enforce.
9. *Mining decentralisation and node decentralisation are different things.* We did not measure hash power. Block production can be concentrated while nodes are not, and the other way round.
10. *A provider that shuts down all its nodes still does not control consensus.* The model agrees: the rest of the network stays connected, and for the largest single provider unmet connection demand rises only moderately, from 14.0% to 21.1% under the defaults of released versions.

Objections 3, 8 and 10 lead us to reject H3. Objections 4 to 6 make the H1 estimates less certain without overturning them. None of the objections touches the finding that the infrastructure we can observe is measurably concentrated, at the provider level and even more so at the country level.

10 Threats to validity

Crawler and vantage point. From home, our crawler reached only 3,075 IPv4 nodes. When we re-probed the 3,414 IPv4 addresses that D2 lists as reachable but that had timed out, using 30 instead of 300 parallel connections, 2,841 of them (83.2%) answered, which points to a saturated home NAT rather than to missing nodes. Residential nodes were lost more often than hosted ones, so D1 is biased towards hosting. We therefore repeated the crawl from an AWS EC2 instance without NAT (D1b), which reached 85.9% of D2's IPv4 hosts. D1b has limitations of its own: it observes from inside a cloud provider, it started about five hours after the D2 snapshot and ran for three hours, and its IPv6 part comes from the home crawl. We keep D2 as the primary dataset because it verifies overlay endpoints and is a single consistent snapshot, and use D1b and D3 as replications. The Linode fleet also shows that crawlers differ in what they discover in the first place.

Table 9: Static accounting of hypothetical failures (D2, SHA256d-chain functional IPv4/IPv6 nodes). % all (A): non-listening nodes (Luke Dashjr estimate) assumed unaffected; (B): affected in the same proportion as reachable IP nodes.

Scenario	Lost	% IP	% reach.	% all (A)	% all (B)	ASNs left	Countries left
Baseline (no failure)	0	0.0	0.0	0.0	0.0	1209	94
Largest ASN (AS24940)	821	11.3	3.7	0.6	9.9	1208	94
Largest provider (Hetzner)	842	11.5	3.8	0.7	10.2	1205	94
Largest country (US)	2159	29.6	9.7	1.7	26.1	954	93
Top-5 providers	2154	29.5	9.6	1.7	26.1	1186	94
Top-10 providers	3194	43.8	14.3	2.5	38.6	1175	94
All hosted (H1+H2)	3220	44.2	14.4	2.5	38.9	1002	92

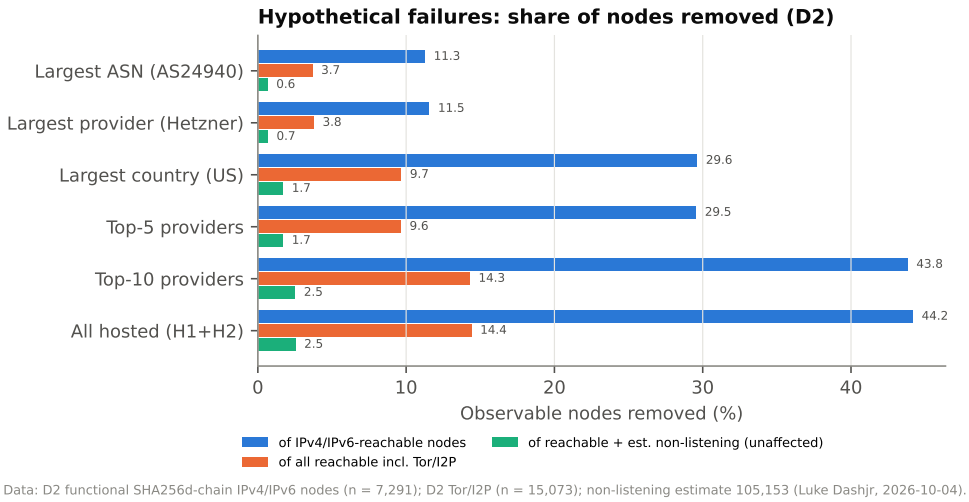


Figure 9: Share of nodes removed in each hypothetical scenario, under three denominators.

Survivorship. A snapshot only contains nodes that were online at that moment, so short-lived and intermittently reachable nodes are under-represented.

Geolocation. Geolocation and registration country agree for only 79.2% of nodes, and for 38.4% of nodes in cloud ranges [38].

ASN attribution. We use a single BGP view (RouteViews). Prefixes with several origin ASes affect 0.35% of nodes and are assigned to the first origin. KIT's ASNs come from its own whois lookups.

Classification. Classes assigned per AS do not necessarily describe individual hosts [45], which is why we report bounds. The name-based rule produces some false positives and affects 1.7% of nodes.

Duplicates, NAT and dynamic addresses. Counting by IP merges several nodes behind one NAT and splits a node whose address changes. At the /24 level there are 7,121 distinct subnets and the largest holds 0.30% of nodes, so clustering within subnets is small.

Tor, I2P and IPv6. Overlay nodes cannot be attributed, and IPv6 attribution depends on a smaller BGP table.

Private nodes and stale data. Non-listening nodes enter only through D4, and the attribution sources are up to 33 days older than the node data.

Change over time. The KIT series shows that these distributions change over the years, and our two measurement days show that individual nodes turn over within days even when the distribution stays put. Two days are still a short window, and our results describe October 2026.

Time series. Organisation names and classes for 2015–2025 are taken from 2026 sources, so an AS that changed owner or business is classified by its current state.

Model. Random peer choice, a fixed share of dual-stack nodes and the absence of miner relay networks are simplifications, and the model's outputs are not measurements.

Table 10: Share of outbound connection demand exceeding inbound capacity (model, including the D4 non-listening estimate, mean of three seeds).

Scenario	Released v27–v31 (114 slots)	Master (189 slots)
Baseline (no failure)	14.0%	0%
Largest ASN	20.9%	<0.01%
Largest provider	21.1%	0.01%
Largest country	32.3%	3.3%
Top-5 providers	32.3%	3.1%
Top-10 providers	41.4%	16.7%
All hosted (H1+H2)	41.6%	17.1%

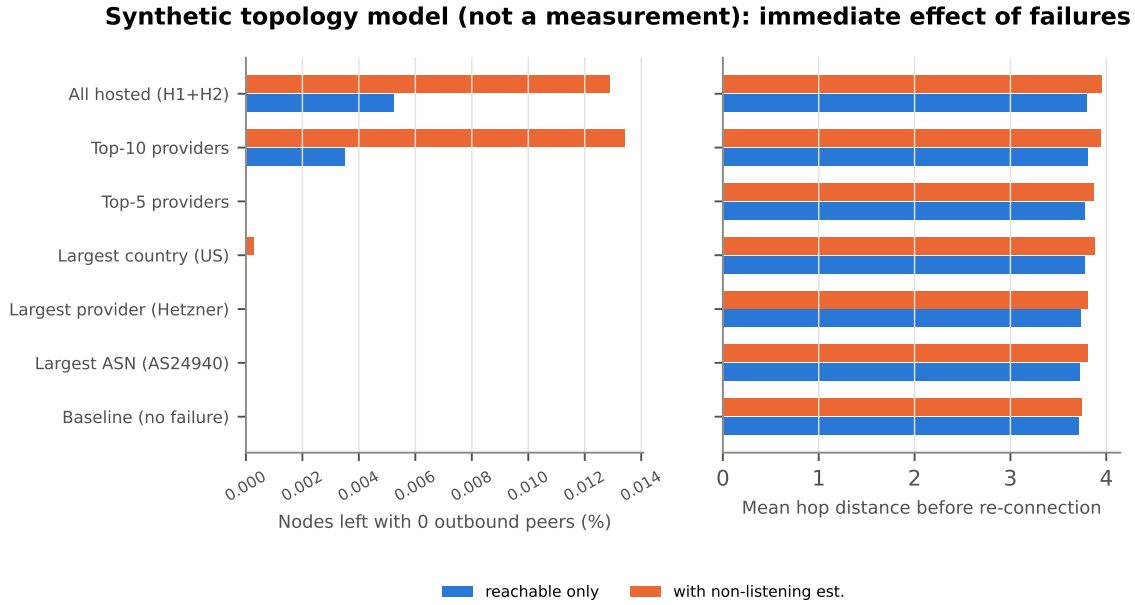


Figure 10: Model results (not measurements): share of nodes left without outbound peers and mean hop distance immediately after each failure.

11 Discussion

What we observed. Among nodes that can be attributed, the provider HHI is about 267, the largest provider hosts about 11.0%, half the nodes are covered by about 15 providers, the country HHI is about 1,397, and between 12.3% and 49.1% of nodes are hosted. Bitnodes, KIT and our own AWS crawl all produce very similar figures, and repeating two of them two days later changed the provider HHI by at most 6 points.

How we interpret it. Concentration exists and can be measured (H1), but at the provider and AS level it is moderate and has decreased since 2020. Losing several large providers at once would remove a substantial part of the reachable IP nodes without, according to the model, splitting the network (H2 not supported for connectivity). It could, however, push the remaining reachable nodes beyond their inbound capacity, to an extent that depends on how many non-listening nodes there really are (H2 partly supported for capacity). Hosting concentration does not give anyone control over consensus (H3 rejected). The more relevant risk is deliberate interference by an AS or provider that biases peer selection or intercepts traffic [3, 42]. In that context, the fact that ASN-based bucketing exists but is not enabled by default deserves attention.

Limitations. The nodes we can attribute are a minority of all nodes, we cannot see who operates them, and the topology is modelled rather than measured.

Open questions. Two hypotheses remain untested: that most non-listening nodes are residential, and that the infrastructure behind Tor relays is concentrated in the same way as node hosting. Testing either

Table 11: Dimensions of independence (functional D2 nodes unless stated). "n/m" means not measurable with the available sources.

Dimension	Measure	Value
Geographic	Country HHI (geoloc.) / top-1 / N_{50}	1,397 / 30.7% / 3
ASN	ASN HHI / top-1 / N_{50}	226 / 10.7% / 26
ISP (access networks)	Share of nodes in R-class ASes	50.9%
Hosting provider	Provider HHI / top-5 / N_{50}	267 / 28.8% / 15
Cloud vs. residential	Hosted share bounds	[12.3%, 49.1%]
Software/client	Client HHI / Core share	7,836 / 87.8%
Network layer	Share of reachable hosts on overlay networks	65.5%
Ownership	Operators behind IPs	n/m (AS ownership $\neq$ node ownership)
Self-hosting	Share of reachable+est. non-listening nodes that are reachable	19.2% (D4, uncertain)
Outage resilience	Model, top-5 providers removed: zero-outbound share / unmet connection demand (baseline 14.0%)	0% / 32.3%

would require data we do not have.

12 Conclusion

The Bitcoin nodes we can observe are measurably concentrated on several infrastructure dimensions, most strongly by country and less so by AS and hosting provider, and the concentration is lower in 2026 than it was in 2020–2022. The evidence does not support the stronger claims that this concentration threatens the network's connectivity or hands control over consensus to infrastructure owners: in our model the network stays connected even without any hosted nodes, and nodes on different chains share the same providers. The weaker point is capacity. Reachable nodes, a small minority of all nodes if the D4 estimate holds, carry the connections of everyone else, and losing many of them at once would shift that load onto the rest.

Several of our findings concern measurement as much as the network itself. Most reachable nodes cannot be attributed to any infrastructure, a single fleet or a single vantage point can change the headline numbers, and node counts include nodes that follow different consensus rules. We think infrastructure concentration is best understood as part of Bitcoin's attack surface, to be addressed through peer diversity mechanisms such as ASN-based bucketing, and not as a measure of who controls Bitcoin.

Use of AI tools. The measurement and analysis code (crawler, data download and attribution pipeline, metric computation, topology model, figure and table generation) was developed with the assistance of the large language model Claude Opus 5.5 (Anthropic), used through Claude Code. The model also assisted with drafting the manuscript text and with automated checks of bibliographic metadata against Crossref. All outputs were reviewed by the author, who takes full responsibility for the methodology, the data, the results and their interpretation. The model is not an author of this work.

Data availability and ethics. Many node operators run nodes from home, which makes node addresses (IPv4, IPv6, onion and I2P) personal data, so we do not publish them. In the public release every address, subnet and netgroup is replaced by a pseudonym computed with a keyed HMAC-SHA256, and the raw crawl and Bitnodes files are left out. We checked that running analysis steps 05 to 08 on the pseudonymised data reproduces every number in this paper. Our crawler only performed the standard handshake and address exchange and identified itself with an honest user agent.

Reproducibility. The code, the hashes of all raw files and the derived tables are in the accompanying repository. The numbered scripts in `analysis/` reproduce every number in the paper, and the numbers in the text are generated macros (`paper/numbers.tex`).

References

[1] Oluwapelumi Adejumo. Bitcoin Knots is trying to fork Bitcoin again after its last chain died in two blocks. CryptoSlate, 29 August 2026, https://cryptoslate.com/bitcoin-knots-is-trying-to-fork-bitcoin-again-after-its-last-chain-died-in-two-blocks/, 2026.

[2] APNIC Labs. Estimates of user population per AS (60-day window ending 2026-10-01). https://stats.labs.apnic.net/aspop/, 2026.

[3] Maria Apostolaki, Aviv Zohar, and Laurent Vanbever. Hijacking Bitcoin: Routing attacks on cryptocurrencies. In *2017 IEEE Symposium on Security and Privacy (SP)*, 2017. doi: 10.1109/SP.2017.29.

[4] Todd Arnold, Jia He, Weifan Jiang, Matt Calder, Italo Cunha, Vasileios Giotsas, and Ethan Katz-Bassett. Cloud provider connectivity in the flat Internet. In *Proceedings of the ACM Internet Measurement Conference*, 2020. doi: 10.1145/3419394.3423613.

[5] Alex Biryukov and Ivan Pustogarov. Bitcoin over Tor isn't a good idea. In *2015 IEEE Symposium on Security and Privacy*, 2015. doi: 10.1109/SP.2015.15.

[6] Alex Biryukov, Dmitry Khovratovich, and Ivan Pustogarov. Deanonymisation of clients in Bitcoin P2P network. In *Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security*, 2014. doi: 10.1145/2660267.2660379.

[7] Bitcoin Core contributors. Bitcoin Core source code, master branch. https://github.com/bitcoin/bitcoin, commit 66776840beb558f7e84451c2c55457f0e06242f0, retrieved 2026-10-04, 2026.

[8] Bitcoin Knots contributors. Release notes, v29.3.knots20260508, v29.4.1.knots20260508 and v29.4.2.knots20260508. https://github.com/bitcoinknots/bitcoin, `doc/release-notes.md` at the respective tags, retrieved 2026-10-04, 2026.

[9] CAIDA. AS to organization mapping dataset, 2026-09-01. https://publicdata.caida.org/datasets/as-organizations/, 2026.

[10] CAIDA. Cloud prefixes dataset, 2026-09-05. https://catalog.caida.org/dataset/cloud_prefixes, 2026.

[11] CAIDA. Routeviews Prefix-to-AS mappings (IPv4 and IPv6), 2026-09-30. https://publicdata.caida.org/datasets/routing/routeviews-prefix2as/, 2026.

[12] Lidia Ceriani and Paolo Verme. The origins of the Gini index: Extracts from Variabilità e Mutabilità (1912) by Corrado Gini. *The Journal of Economic Inequality*, 2011. doi: 10.1007/s10888-011-9188-x.

[13] Luke Dashjr. Bitcoin node software (node-count estimates, `uainfo.json`). https://luke.dashjr.org/programs/bitcoin/files/charts/software.html, retrieved 2026-10-04, 2026.

[14] DB-IP. IP to Country Lite database, 2026-10 (CC BY 4.0). https://db-ip.com/db/download/ip-to-country-lite, 2026.

[15] Christian Decker and Roger Wattenhofer. Information propagation in the Bitcoin network. In *IEEE P2P 2013 Proceedings*, 2013. doi: 10.1109/P2P.2013.6688704.

[16] Sergi Delgado-Segura, Surya Bakshi, Cristina Pérez-Solà, James Litton, Andrew Pachulski, Andrew Miller, and Bobby Bhattacharjee. TxProbe: Discovering Bitcoin's network topology using orphan transactions. In *Financial Cryptography and Data Security (FC 2019)*, Lecture Notes in Computer Science, 2019. doi: 10.1007/978-3-030-32101-7_32.

[17] Joan Antoni Donet Donet, Cristina Pérez-Solà, and Jordi Herrera-Joancomartí. The Bitcoin P2P network. In *Financial Cryptography and Data Security (FC 2014 Workshops)*, Lecture Notes in Computer Science, 2014. doi: 10.1007/978-3-662-44774-1_7.

[18] Meryam Essaid, Changhyun Lee, and Hongteak Ju. Characterizing the Bitcoin network topology with Node-Probe. *International Journal of Network Management*, 2023. doi: 10.1002/nem.2230.

[19] Ittay Eyal and Emin Gün Sirer. Majority is not enough: Bitcoin mining is vulnerable. *Communications of the ACM*, 2018. doi: 10.1145/3212998.

[20] Sebastian Feld, Mirco Schönfeld, and Martin Werner. Analyzing the deployment of Bitcoin's P2P network under an AS-level perspective. *Procedia Computer Science*, 2014. doi: 10.1016/j.procs.2014.05.542.

[21] Adem Efe Gencer, Soumya Basu, Ittay Eyal, Robbert van Renesse, and Emin Gün Sirer. Decentralization in Bitcoin and Ethereum networks. In *Financial Cryptography and Data Security (FC 2018)*, Lecture Notes in Computer Science, 2018. doi: 10.1007/978-3-662-58387-6_24.

[22] Matthias Grundmann, Hedwig Amberg, Max Baumstark, and Hannes Hartenstein. Short paper: What peer announcements tell us about the size of the Bitcoin P2P network. In *Financial Cryptography and Data Security (FC 2022)*, Lecture Notes in Computer Science, 2022. doi: 10.1007/978-3-031-18283-9_35.

[23] Matthias Grundmann, Max Baumstark, and Hannes Hartenstein. On the peer degree distribution of the Bitcoin P2P network. In *2022 IEEE International Conference on Blockchain and Cryptocurrency (ICBC)*, 2022. doi: 10.1109/ICBC54727.2022.9805511.

[24] Ethan Heilman, Alison Kendler, Aviv Zohar, and Sharon Goldberg. Eclipse attacks on Bitcoin's peer-to-peer network. In *24th USENIX Security Symposium*, pages 129–144, 2015.

[25] Aqsa Kashaf, Vyas Sekar, and Yuvraj Agarwal. Analyzing third party service dependencies in modern web services: Have we learned from the Mirai-Dyn incident? In *Proceedings of the ACM Internet Measurement Conference*, 2020. doi: 10.1145/3419394.3423664.

[26] Seoung Kyun Kim, Zane Ma, Siddharth Murali, Joshua Mason, Andrew Miller, and Michael Bailey. Measuring Ethereum network peers. In *Proceedings of the Internet Measurement Conference 2018*, 2018. doi: 10.1145/3278532.3278542.

[27] Yujin Kwon, Jian Liu, Minjeong Kim, Dawn Song, and Yongdae Kim. Impossibility of full decentralization in permissionless blockchains. In *Proceedings of the 1st ACM Conference on Advances in Financial Technologies*, 2019. doi: 10.1145/3318041.3355463.

[28] Qinwei Lin, Chao Li, Xifeng Zhao, and Xianhai Chen. Measuring decentralization in Bitcoin and Ethereum using multiple metrics and granularities. In *2021 IEEE 37th International Conference on Data Engineering Workshops (ICDEW)*, 2021. doi: 10.1109/ICDEW53142.2021.00022.

[29] Rodrigo Martínez. btcnodes.io project (deployment of the Bitnodes crawler). https://github.com/brunneis/bitnodes; snapshot 1791105141 via https://bitnodes.io/api/v1/snapshots/latest/ (redirects to btcnodes.io), retrieved 2026-10-04, 2026.

[30] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. https://bitcoin.org/bitcoin.pdf, 2008.

[31] Gleb Naumenko, Gregory Maxwell, Pieter Wuille, Alexandra Fedorova, and Ivan Beschastnikh. Erlay: Efficient transaction relay for Bitcoin. In *Proceedings of the 2019 ACM SIGSAC Conference on Computer and Communications Security*, 2019. doi: 10.1145/3319535.3354237.

[32] Till Neudecker and Hannes Hartenstein. Network layer aspects of permissionless blockchains. *IEEE Communications Surveys & Tutorials*, 2019. doi: 10.1109/COMST.2018.2852480.

[33] Till Neudecker, Philipp Andelfinger, and Hannes Hartenstein. Timing analysis for inferring the topology of the Bitcoin peer-to-peer network. In *2016 Intl IEEE Conferences UIC/ATC/ScalCom/CBDCom/IoP/SmartWorld*, 2016. doi: 10.1109/UIC-ATC-ScalCom-CBDCom-IoP-SmartWorld.2016.0070.

[34] Till Neudecker, Matthias Grundmann, and Hannes Hartenstein. Bitcoin peer-to-peer network monitoring. KIT DSN, https://dsn.tm.kit.edu/bitcoin/; snapshots retrieved 2026-10-04, 2025.

[35] Number Resource Organization. NRO extended allocation and assignment statistics. https://ftp.ripe.net/pub/stats/ripencc/nro-stats/latest/, retrieved 2026-10-04, 2026.

[36] Sehyun Park, Seongwon Im, Youhwan Seol, and Jeongyeup Paek. Nodes in the Bitcoin network: Comparative measurement study and survey. *IEEE Access*, 2019. doi: 10.1109/ACCESS.2019.2914098.

[37] PeeringDB. PeeringDB API, network records. https://www.peeringdb.com/api/net, retrieved 2026-10-04, 2026.

[38] Ingmar Poese, Steve Uhlig, Mohamed Ali Kaafar, Benoit Donnet, and Bamba Gueye. IP geolocation databases: Unreliable? *ACM SIGCOMM Computer Communication Review*, 2011. doi: 10.1145/1971162.1971171.

[39] Muhammad Saad, Victor Cook, Lan Nguyen, My T. Thai, and Aziz Mohaisen. Partitioning attacks on Bitcoin: Colliding space, time, and logic. In *2019 IEEE 39th International Conference on Distributed Computing Systems (ICDCS)*, 2019. doi: 10.1109/ICDCS.2019.00119.

[40] Ashish Rajendra Sai, Jim Buckley, Brian Fitzgerald, and Andrew Le Gear. Taxonomy of centralization in public blockchain systems: A systematic literature review. *Information Processing & Management*, 2021. doi: 10.1016/j.ipm.2021.102584.

[41] Balaji S. Srinivasan and Leland Lee. Quantifying decentralization. Blog post, 21 July 2017. Summary verified via https://messari.io/report/analysis-quantifying-decentralization-balaji-srinivasan-and-leland-lee, 2017.

[42] Muoi Tran, Inho Choi, Gi Jun Moon, Anh V. Vu, and Min Suk Kang. A stealthier partitioning attack against Bitcoin peer-to-peer network. In *2020 IEEE Symposium on Security and Privacy (SP)*, 2020. doi: 10.1109/SP40000.2020.00027.

[43] U.S. Department of Justice and Federal Trade Commission. Merger guidelines. https://www.justice.gov/atr/2023-merger-guidelines, 2023.

[44] Liang Wang and Ivan Pustogarov. Towards better understanding of Bitcoin unreachable peers. https://arxiv.org/abs/1709.06837, 2017.

[45] Maya Ziv, Liz Izhikevich, Kimberly Ruth, Katherine Izhikevich, and Zakir Durumeric. ASdb: A system for classifying owners of autonomous systems. In *Proceedings of the 21st ACM Internet Measurement Conference*, 2021. doi: 10.1145/3487552.3487853.