

IPFS Network: Deployment and Performance



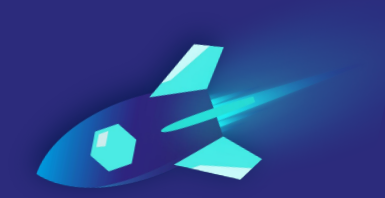
Cosener's House - July 2022

Yiannis Psaras
Research Scientist
Protocol Labs

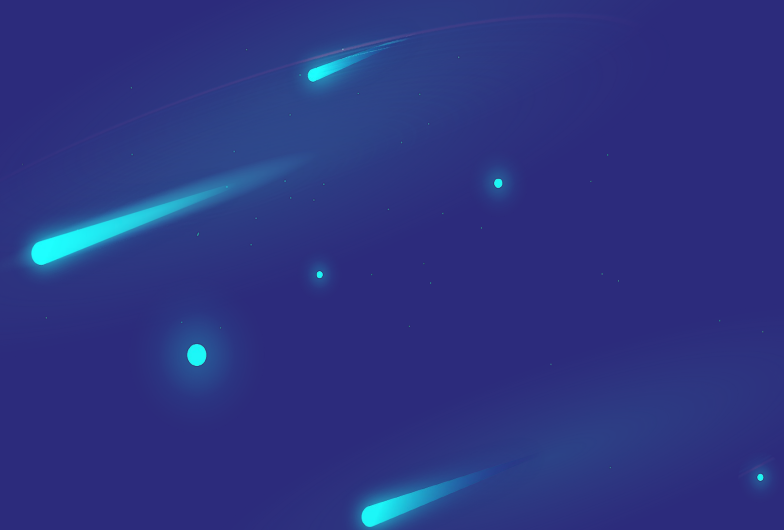


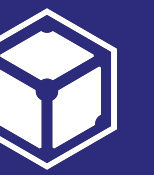


What is IPFS (in words)?



IPFS is a **decentralized storage and delivery network** which builds on fundamental principles of *P2P networking* and *content-based addressing*



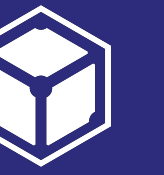


IPFS is a **decentralized storage and delivery network** which builds on fundamental principles of *P2P networking* and *content-based addressing*

- **decentralized:** open and permissionless
- **P2P network:** DHT-based (for the most part, but not exclusively)
- **content-addressed:** based on content hashes, or Content Identifiers (CIDs)

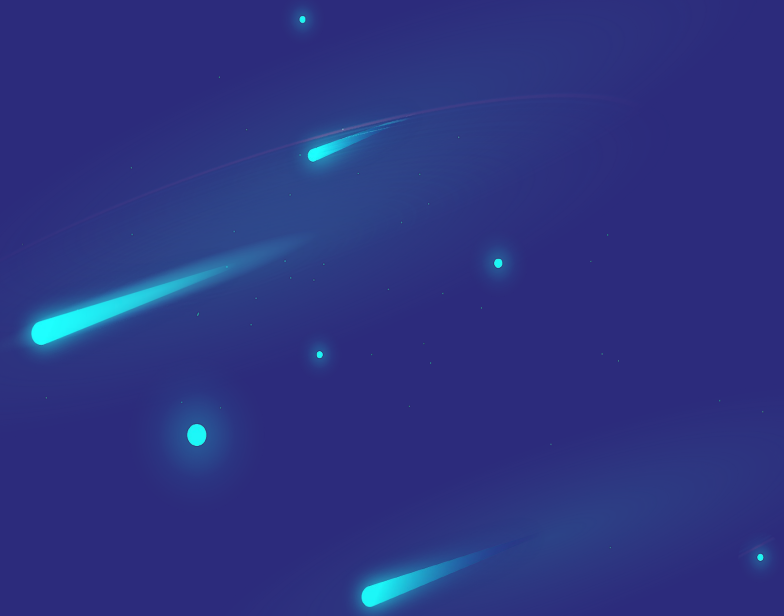
CID:

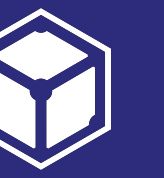
bafybeidbzzzvjuzuf7yjet27sftttod5fowge
3nzt3ybz5uxx1dsdonozq



decentralized: <https://doesipfshaveablockchain.com/>

no.

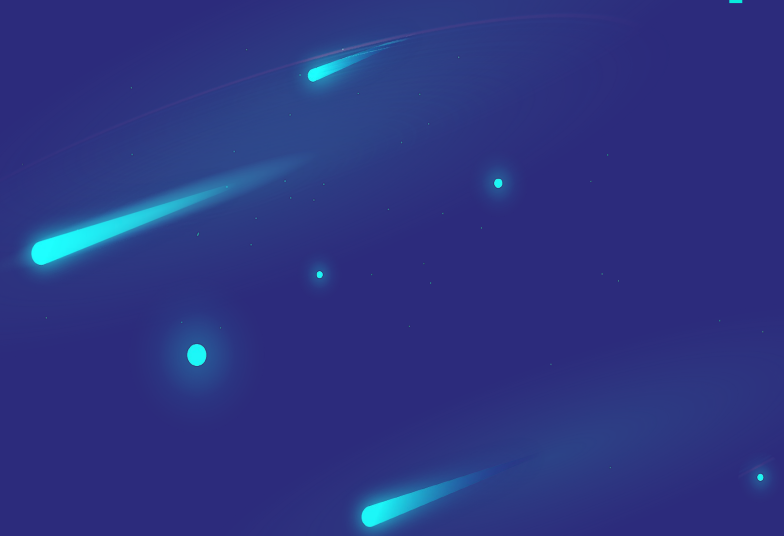




What is IPFS (in stats)?



- Traffic: >> 1B requests (weekly)
- Unique weekly users: several Ms
- DHT Servers: ~20k
- Unique (weekly) Server Peers: ~60k



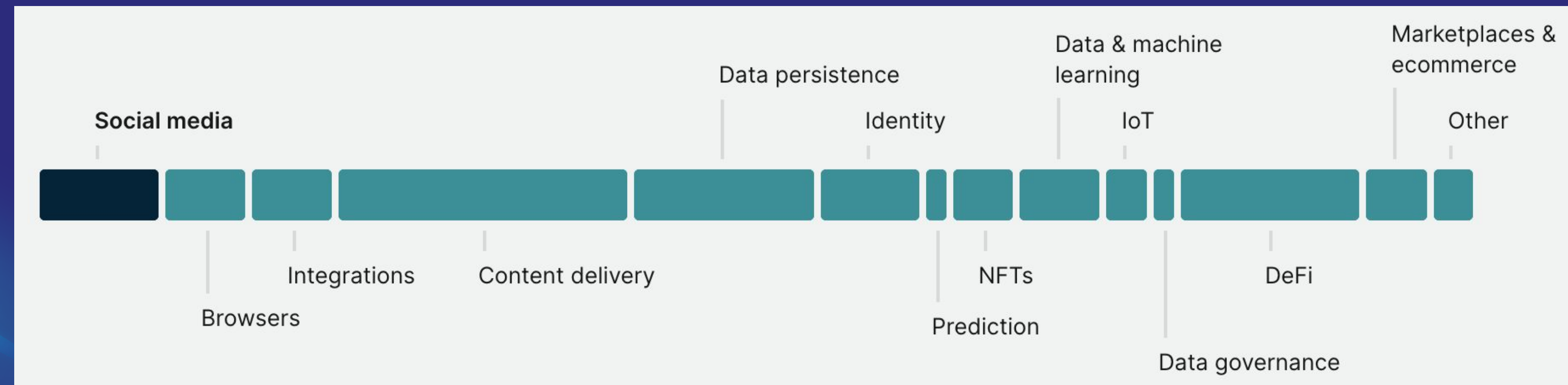
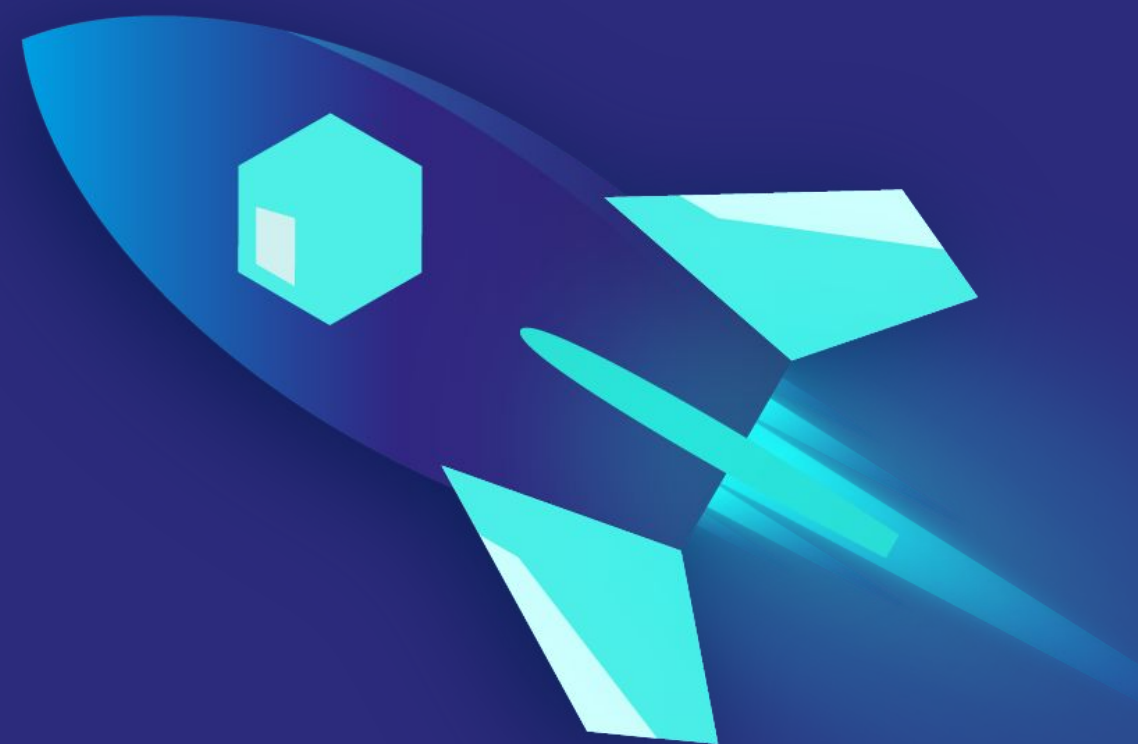


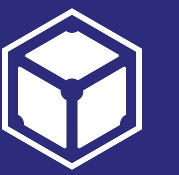
Coolness 1: Apps

Why is it cool?

You can build apps on top!

- Already a ton of apps on many verticals.
- <https://ecosystem.ipfs.io>





Coolness 1: Apps Why is it cool?



Data



Identity



Productivity



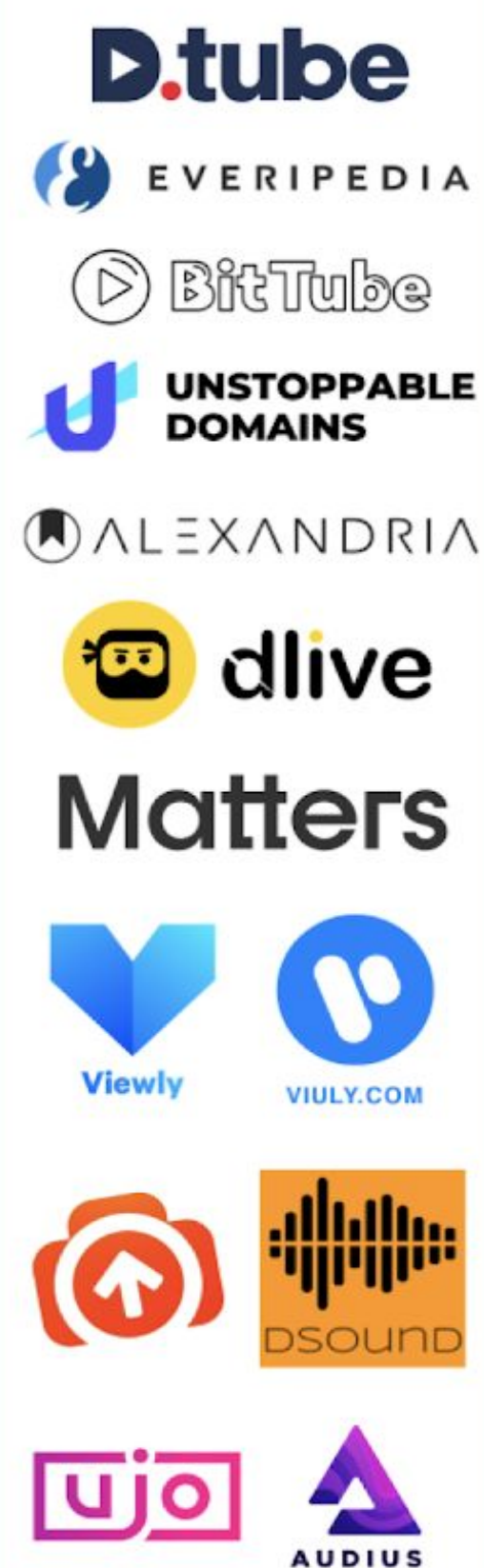
Marketplace



NFT



Content



Other



Social Media



Integrations & Collabs



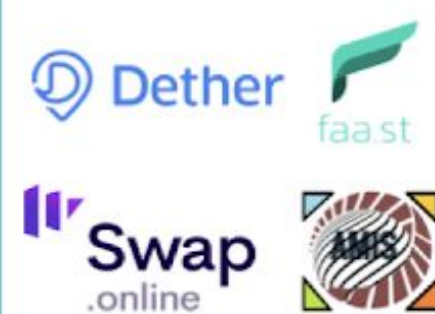
Prediction and betting



Governance

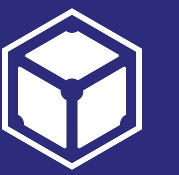


Exchange



Finance



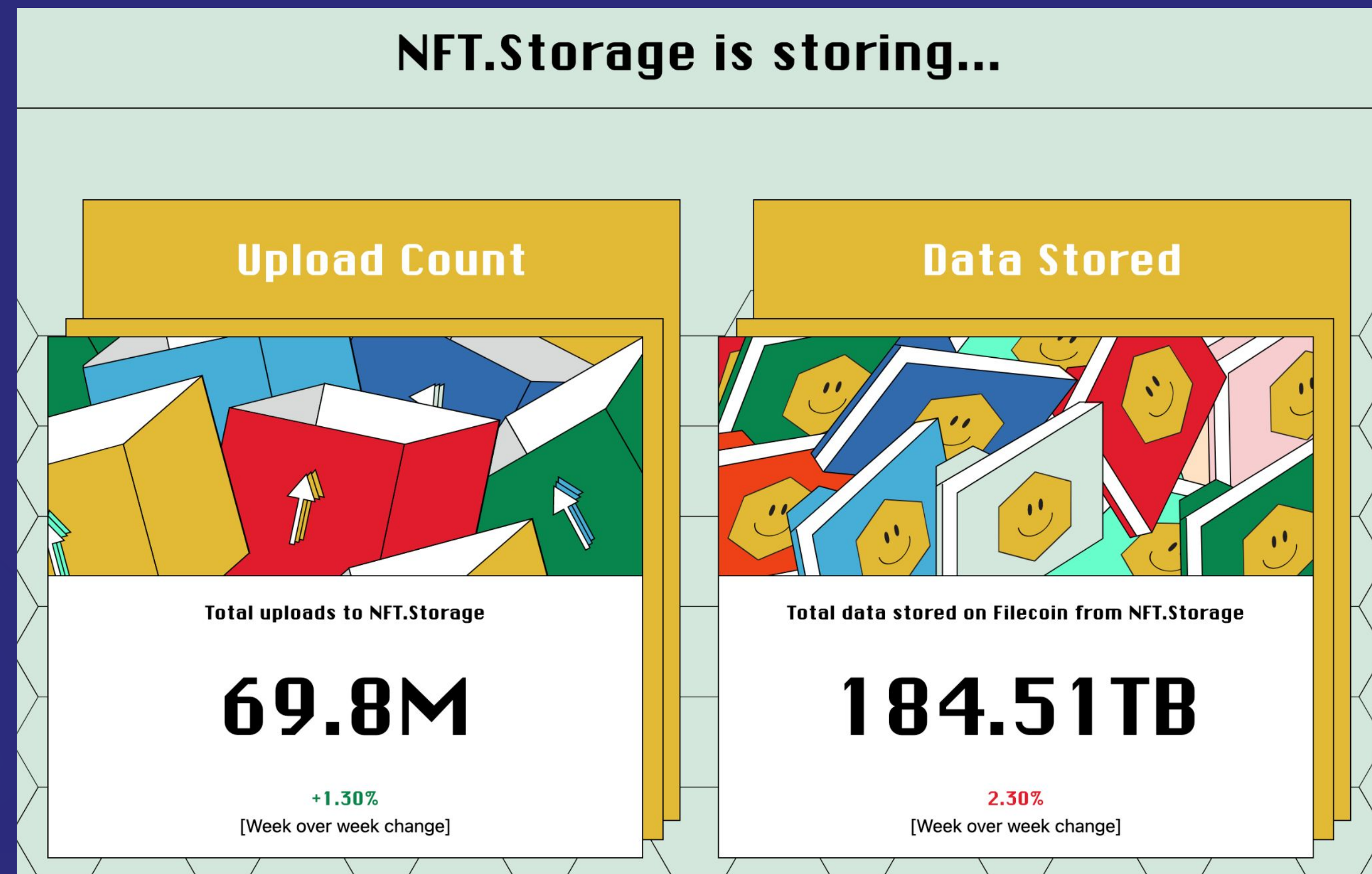


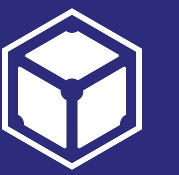
Coolness 2: NFTs

Why is it cool?

You can store your NFTs!

- Most used platform for NFT storage
- Content addressing comes handy
- <https://nft.storage>




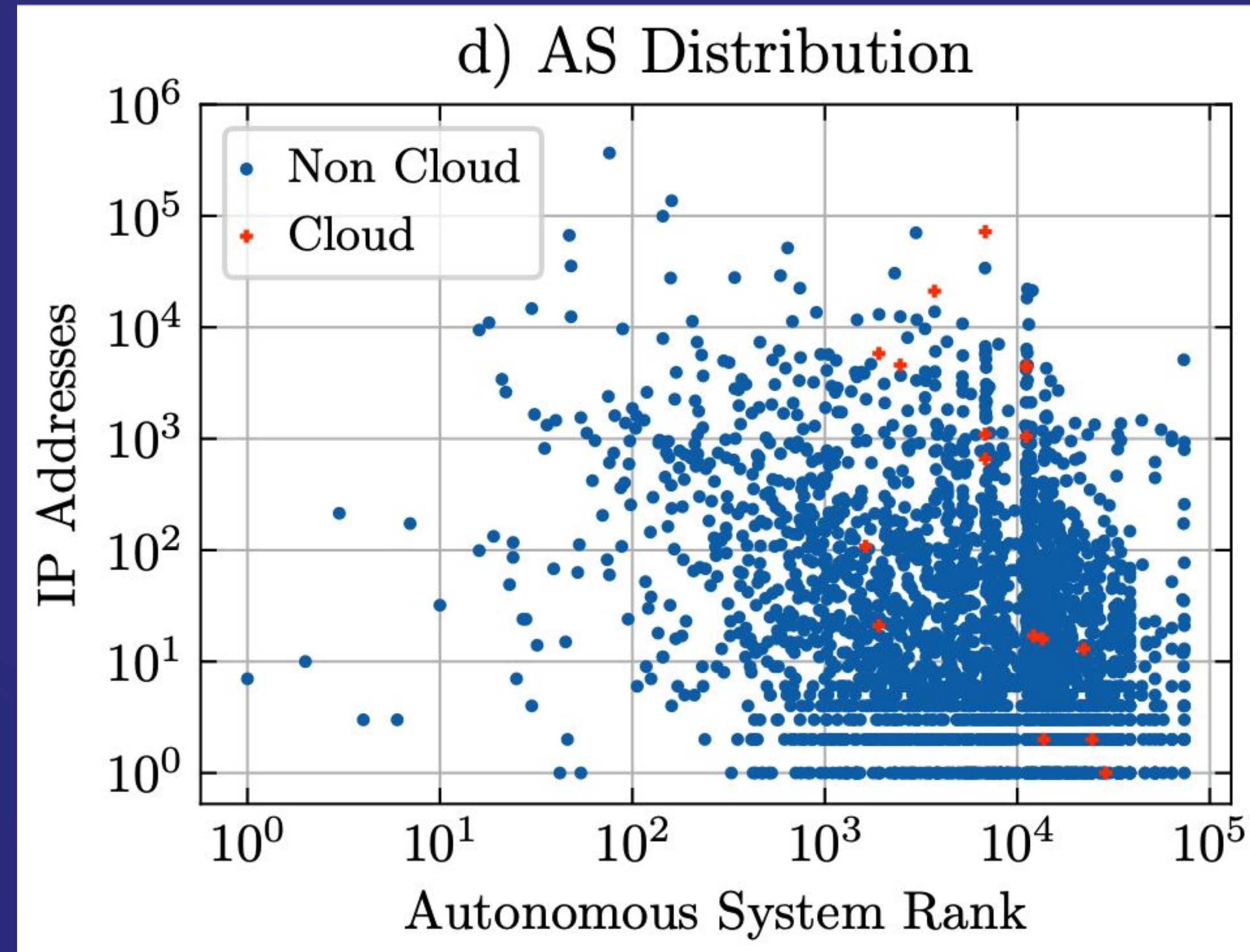
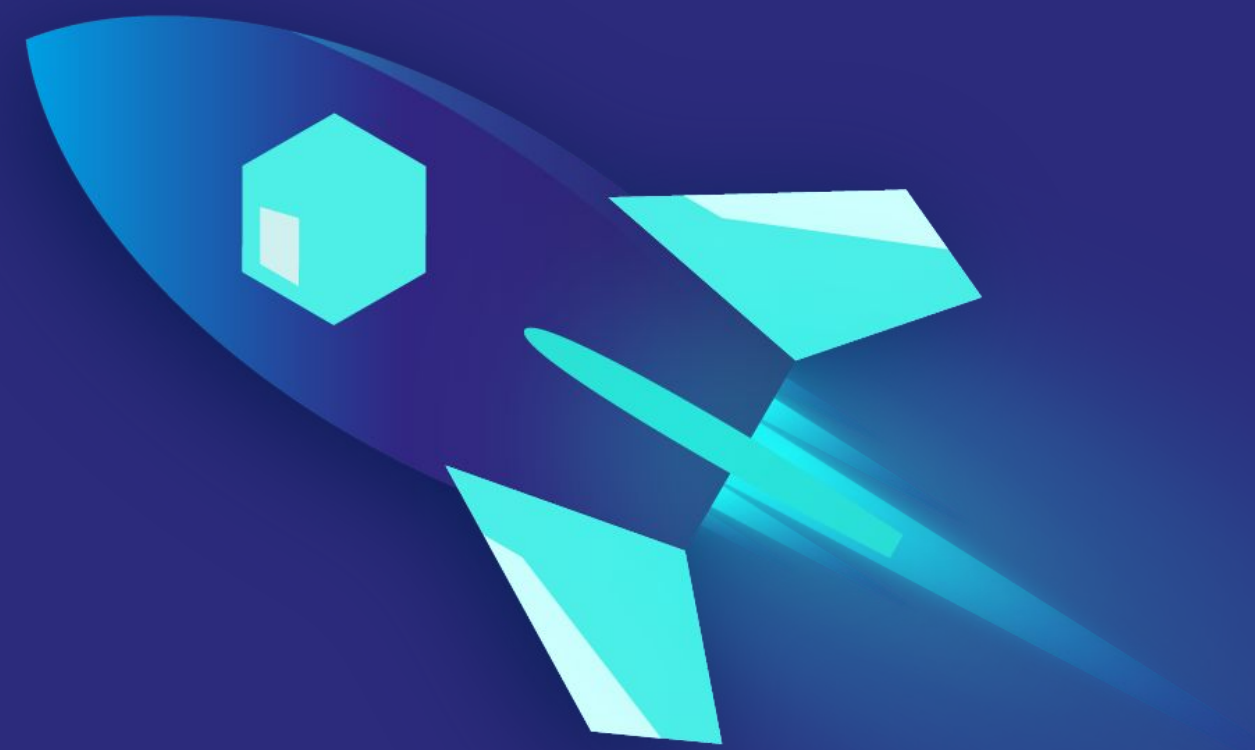


Coolness 3: Open

Why is it cool?

No one is in charge of the network!

- ~100 Gateways
 - <https://ipfs.github.io/public-gateway-checker/>
- Server Nodes in > 2700 ASes
- Think of IPFS as a decentralized CDN 





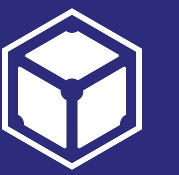
Coolness 4: It's Content Addressed

Why is it cool?

Content addressing has always been cool! 😁

- Requests are for the actual content, not for the address where the content is supposed to be.
- Well-known advantages
 - bandwidth savings
 - faster retrieval latency
 - both through caching and serving local copies of content
 - self-verification of content out of the box - no need to place trust on the DNS system, or CAs.
 - can work offline

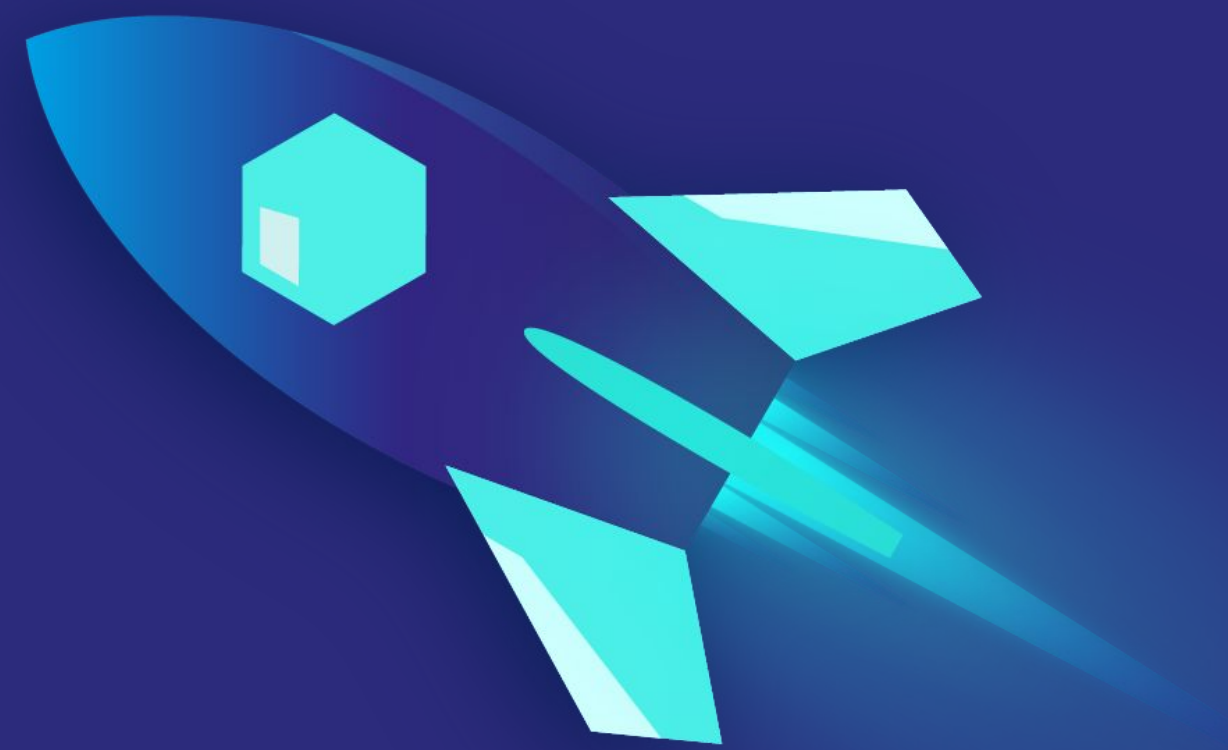
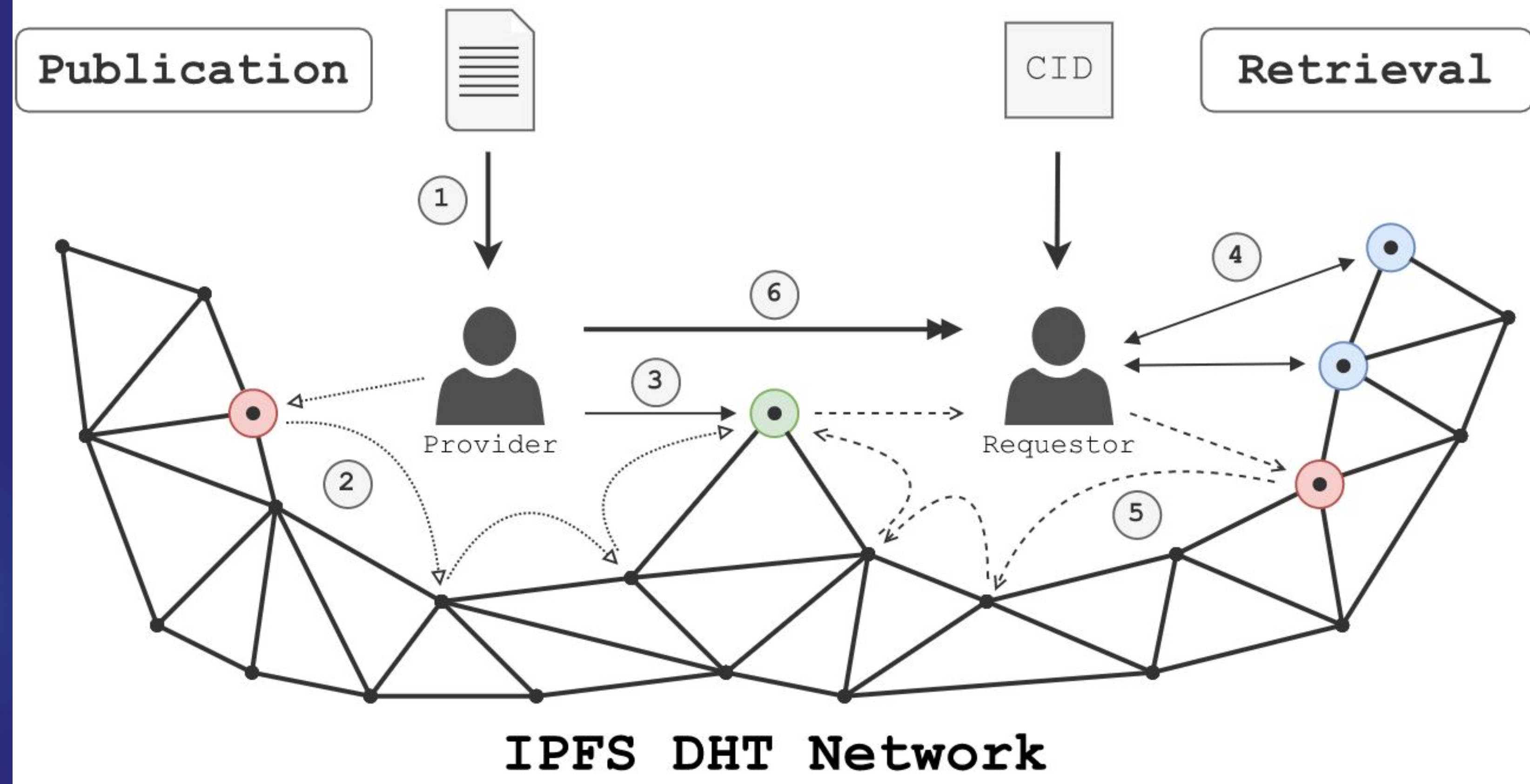
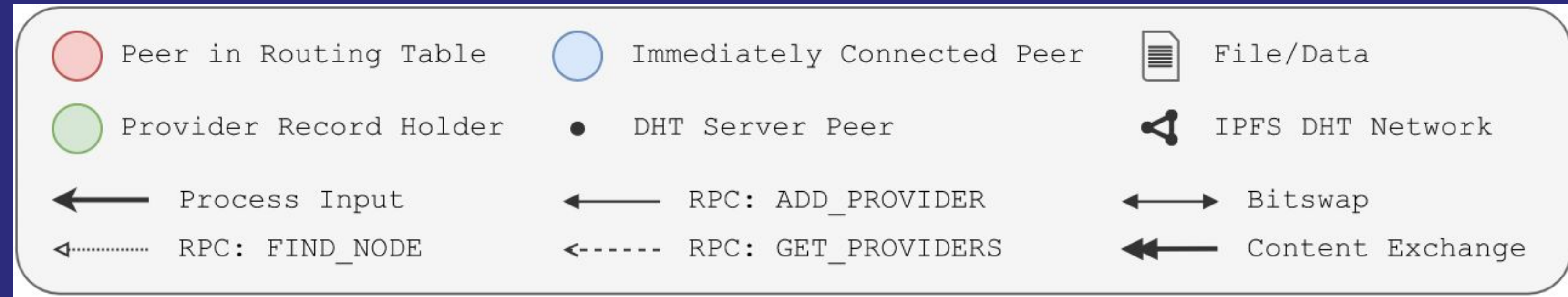




Architecture Overview (partial)

Nutshell

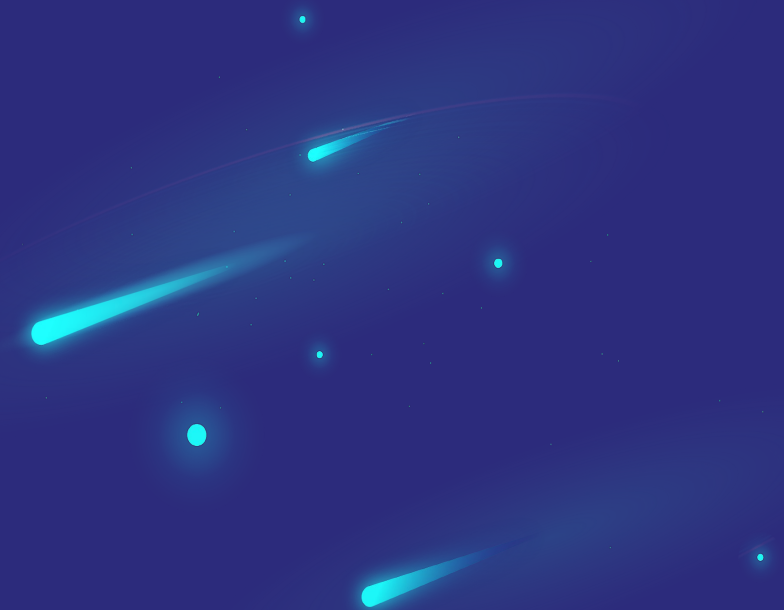
How does IPFS work?



Measuring IPFS



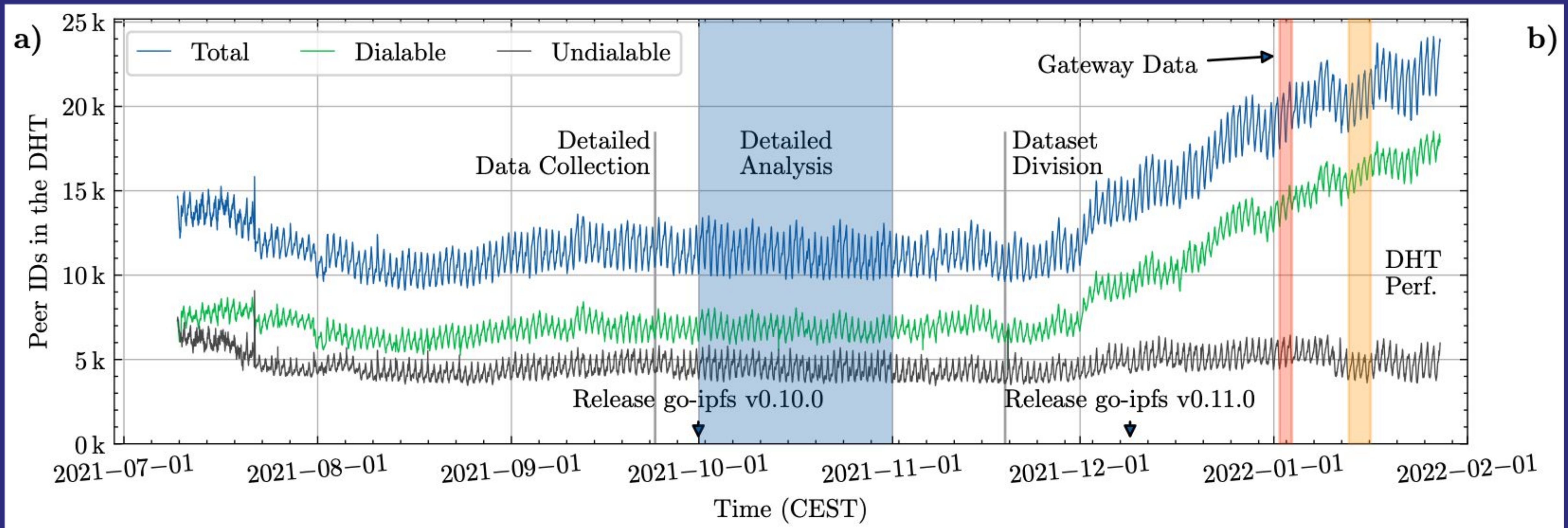
You can't improve what you don't measure and you should measure what (you *think*) you've just improved.

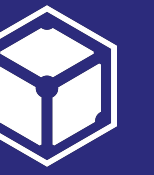




The Nebula Crawler Measurement Campaign

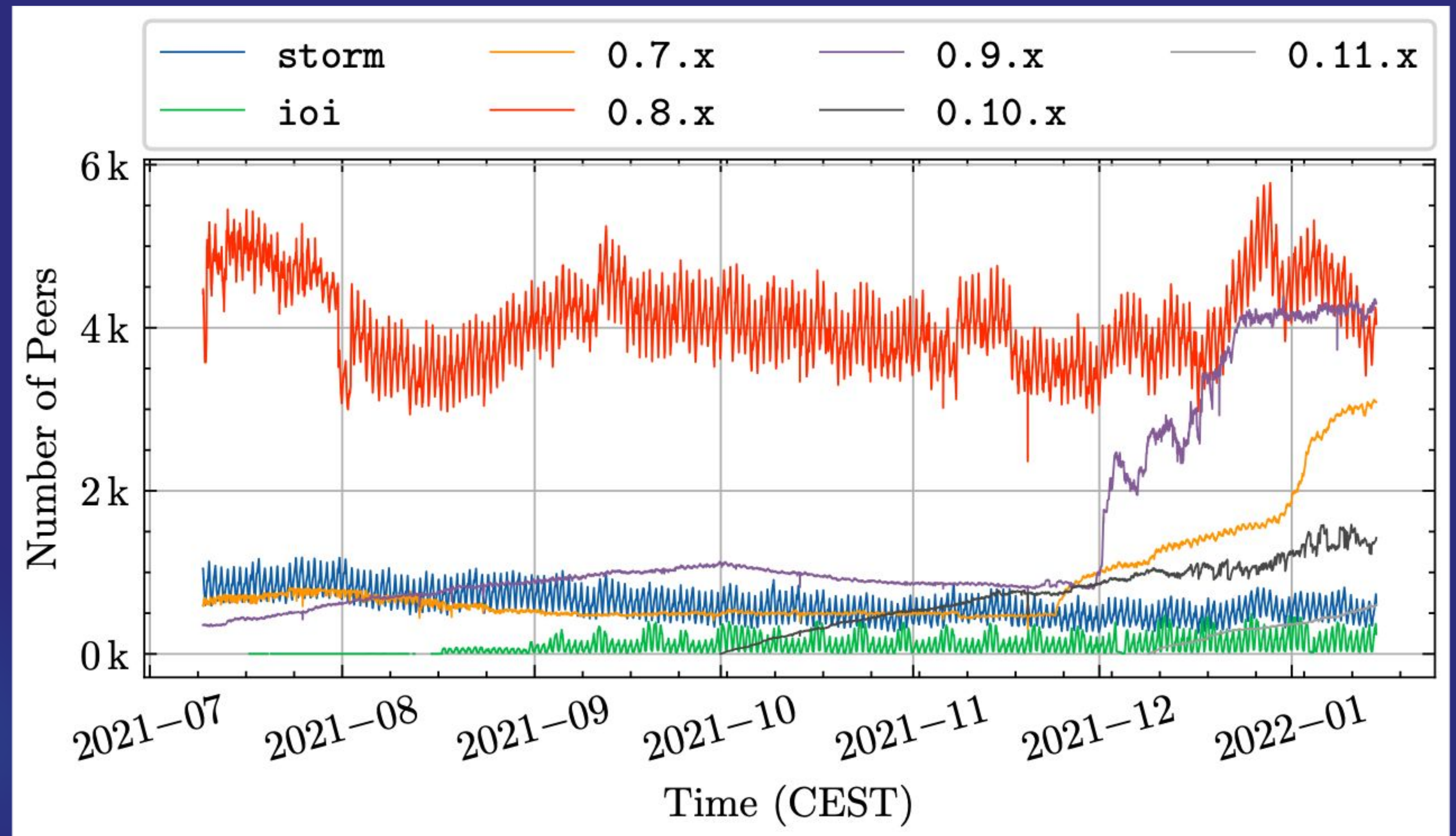
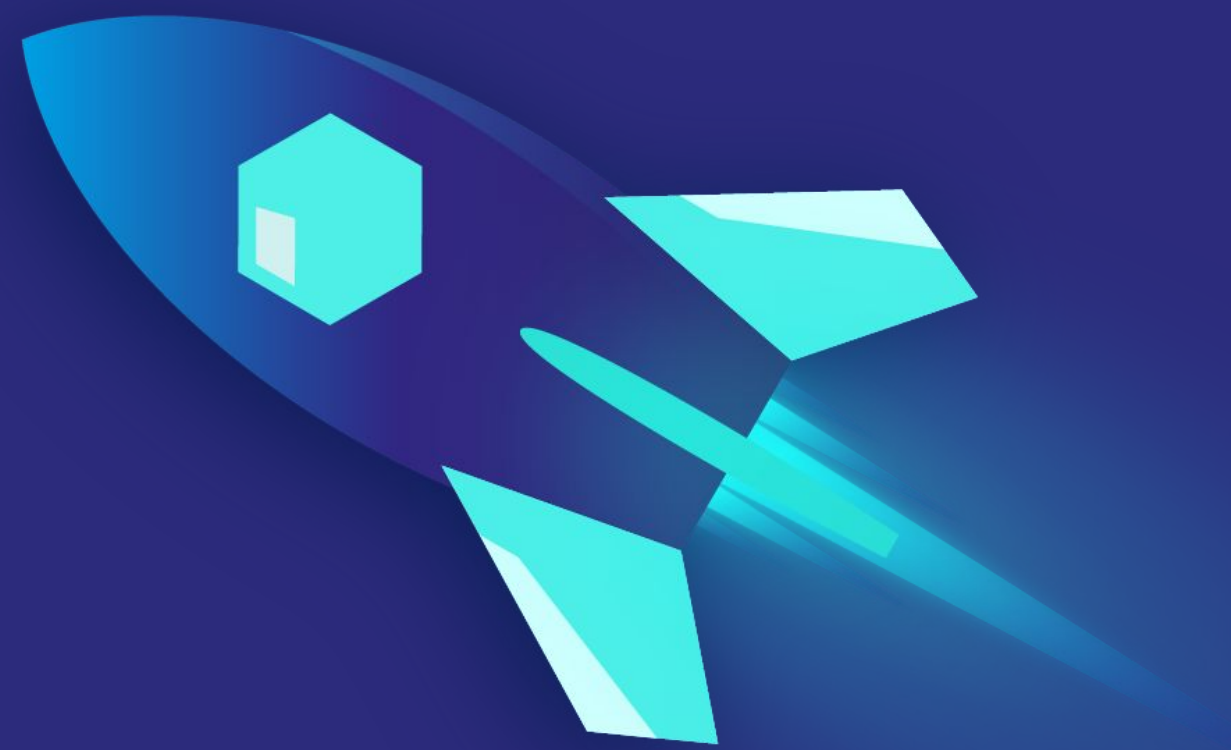
198,964 Peers | 1,998,825 Multi-Addresses | 464,303 IP Addresses | 152 countries

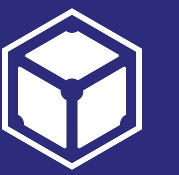




Metrics & Statistics

Agent Version Uptake

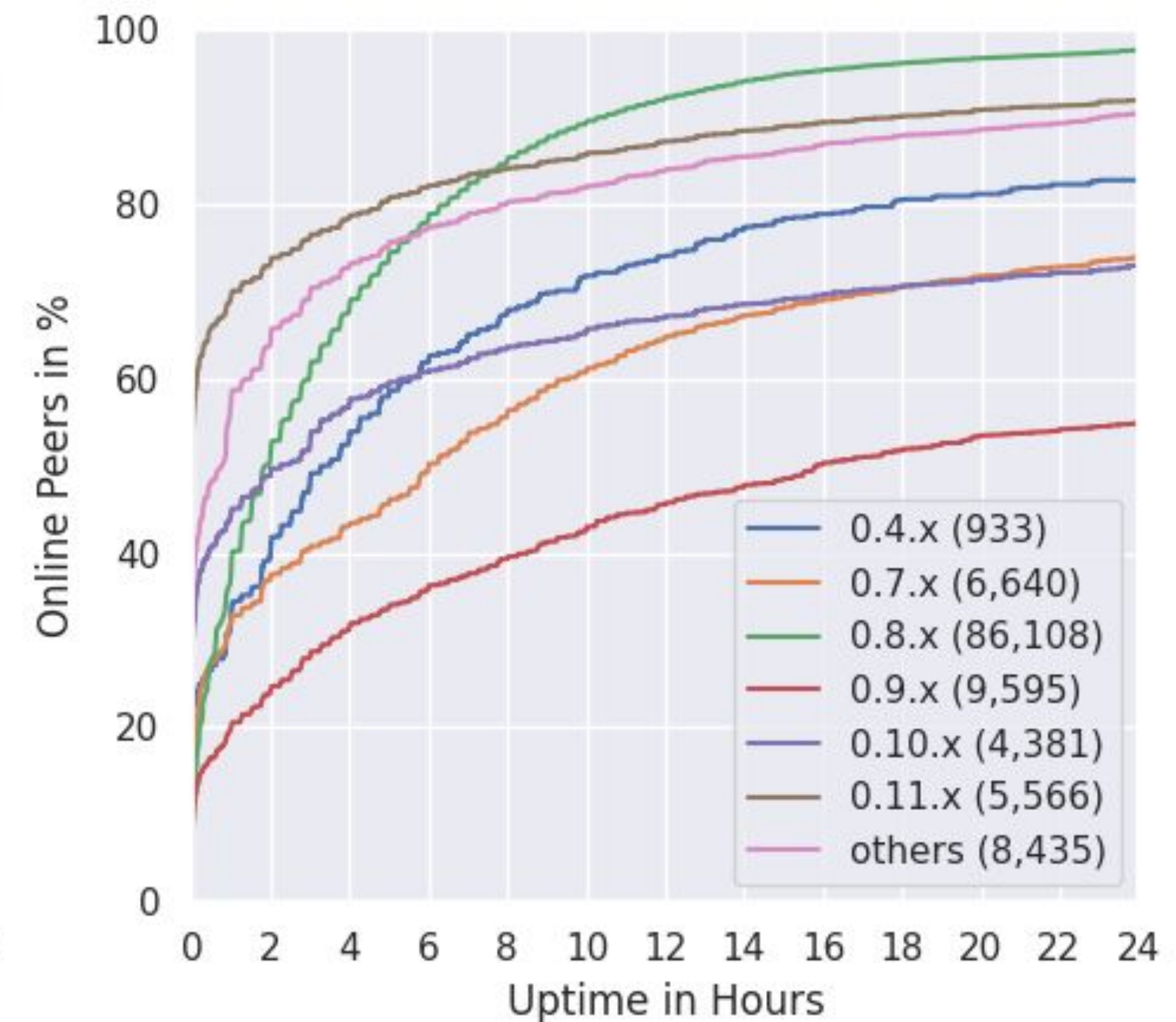
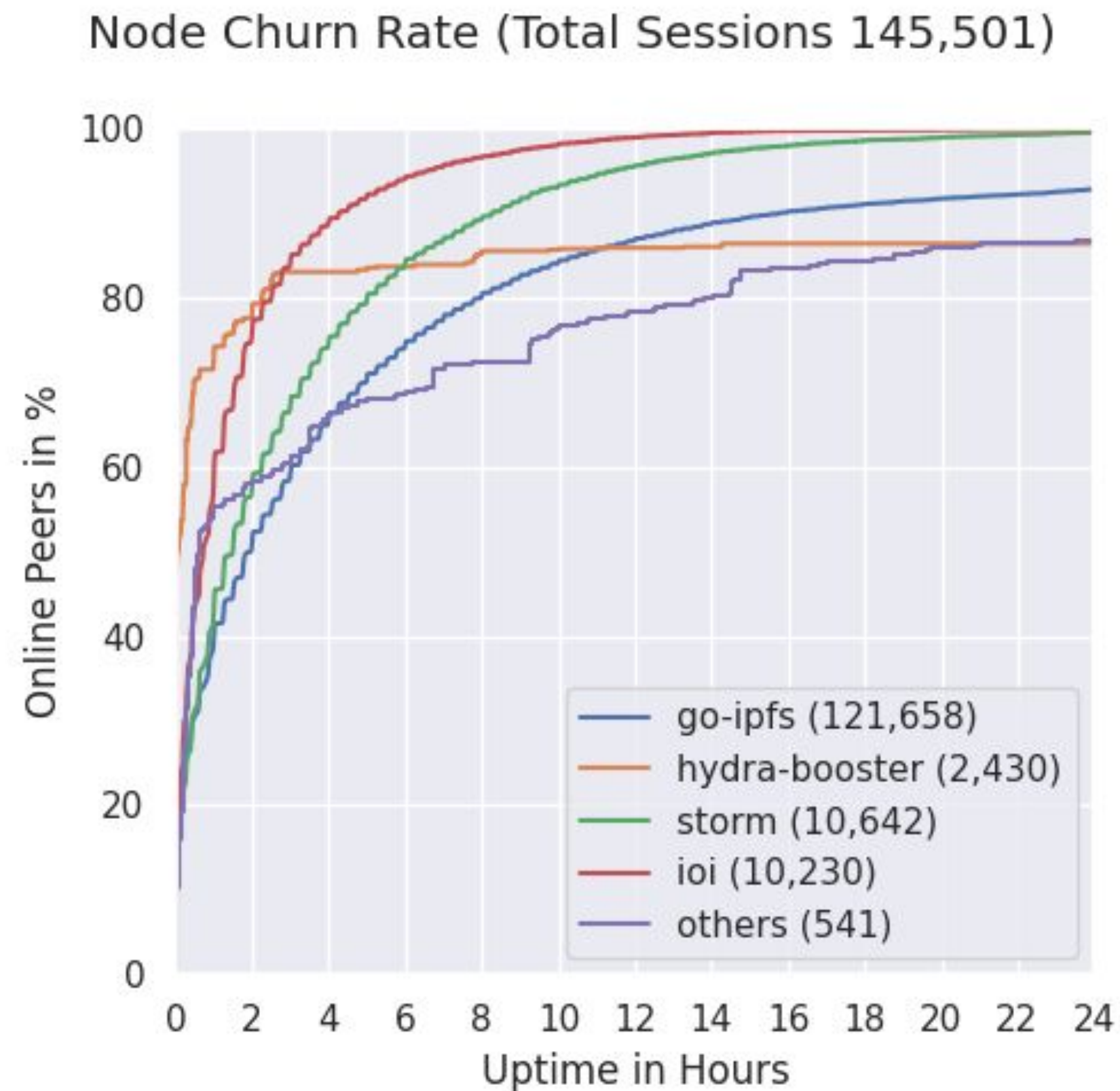
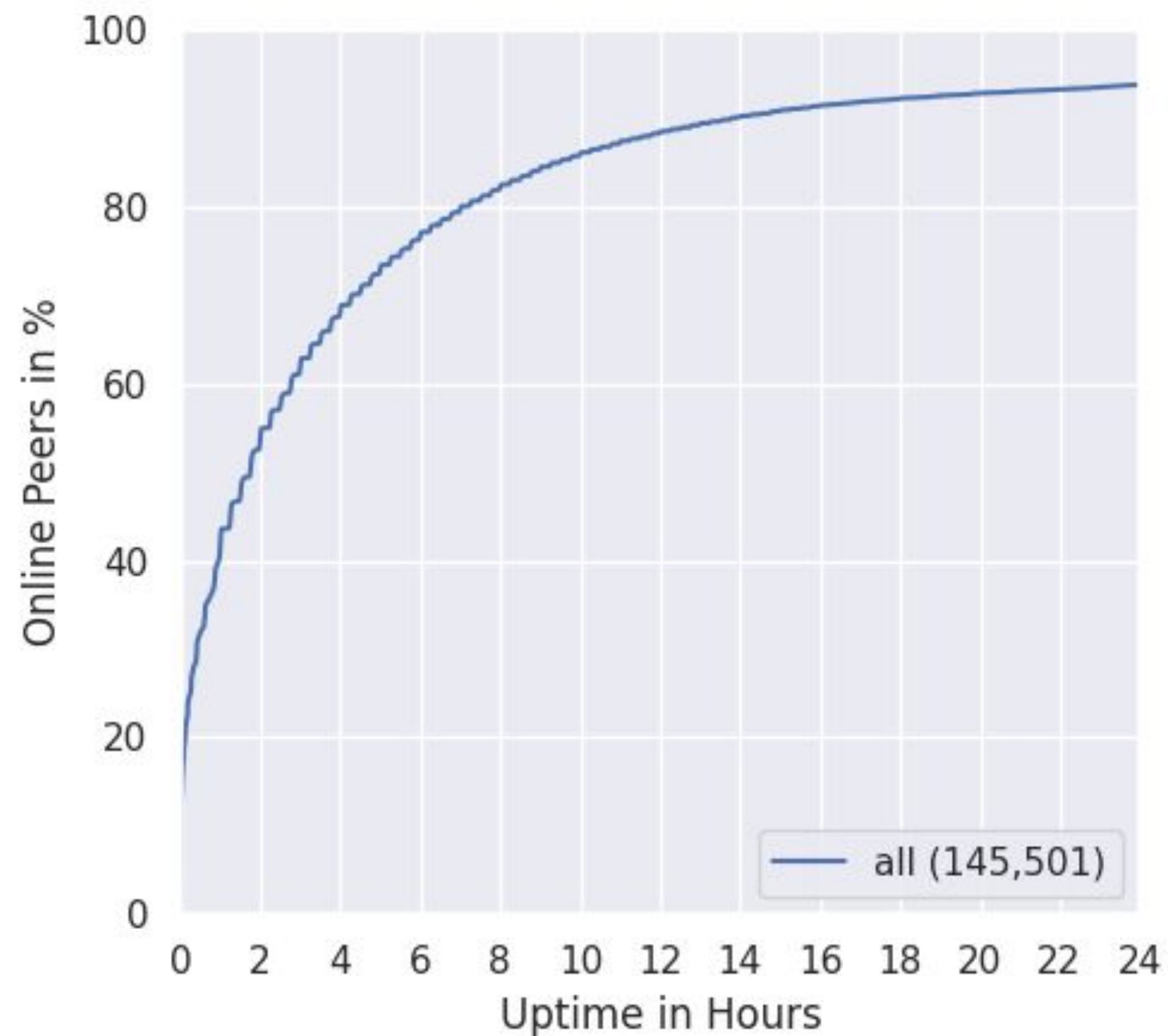


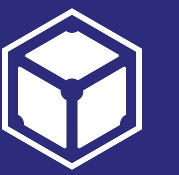


Metrics & Statistics

Peer Churn

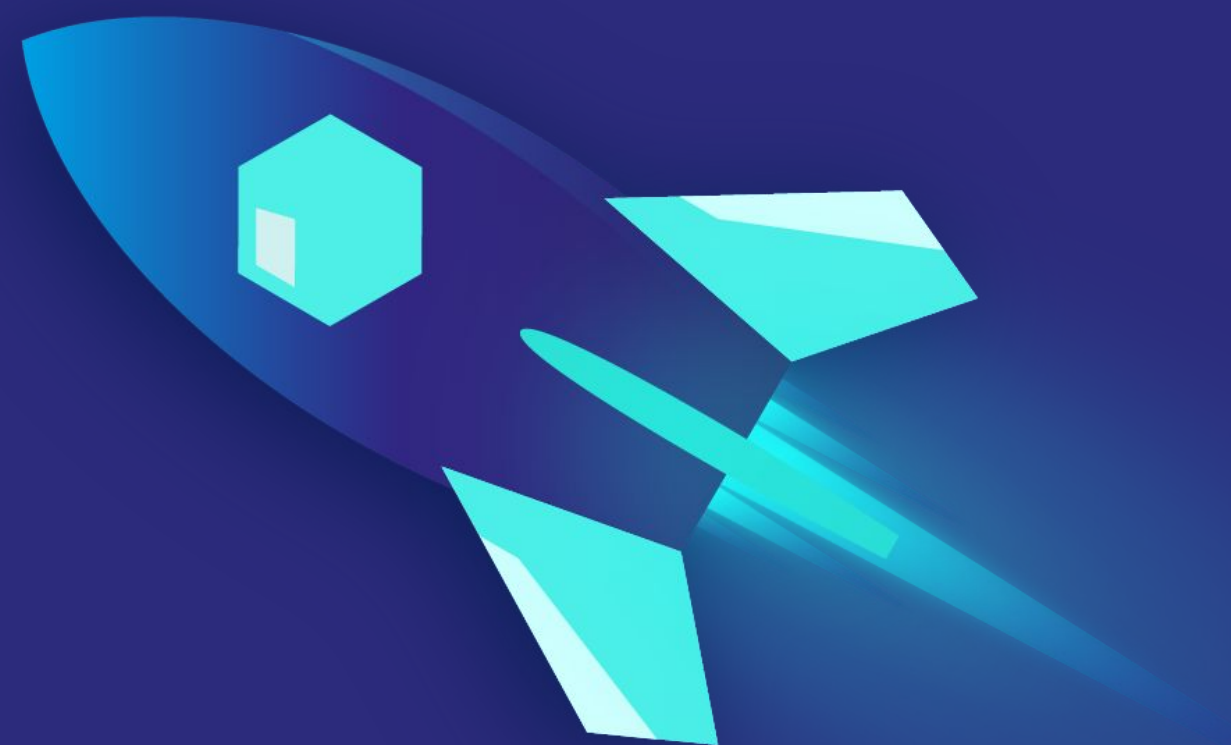
- Easy to spot inconsistencies, misbehaviour, as well as potential bugs in new releases.
- Churn rate influences settings such as the provider record replication.





Metrics & Statistics

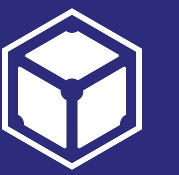
Cloud Provider Dependency



- Very small minority of nodes hosted on centralised cloud infrastructure!
 - At least on providers whose IP addresses are public.

Table 3: Percentage of nodes hosted on cloud providers. The table shows the top ten and selected cloud providers.

Rank	Provider	IP Addresses	IP Address Share
1	Contabo GmbH	2038	0.44 %
2	Amazon AWS	1792	0.39 %
3	Microsoft Azure/Coporation	1536	0.33 %
4	Digital Ocean	836	0.18 %
5	Hetzner Online	592	0.13 %
6	GZ Systems	346	<0.10 %
7	OVH	341	<0.10 %
8	Google Cloud	286	<0.10 %
9	Tencent Cloud	258	<0.10 %
10	Choopa, LLC. Cloud	244	<0.10 %
12	Alibaba Cloud	180	<0.10 %
13	CloudFlare Inc	140	<0.10 %
27	Oracle Cloud	27	<0.10 %
54	IBM Cloud	9	<0.10 %
	235 Other Cloud Providers	2017	0.43 %
	Non-Cloud	453,661	97.71 %



Content Routing

DHT Lookup Process

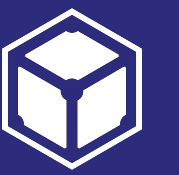
What we wanted to find out:

- Break down the content routing process in steps
- Identify what works well and what doesn't

What we did: a controlled experiment

- Deployed nodes in several geographic areas
- Published CIDs from one of them
- Requested those CIDs from other nodes
- Measured latencies from start to finish
- Repeated from all nodes

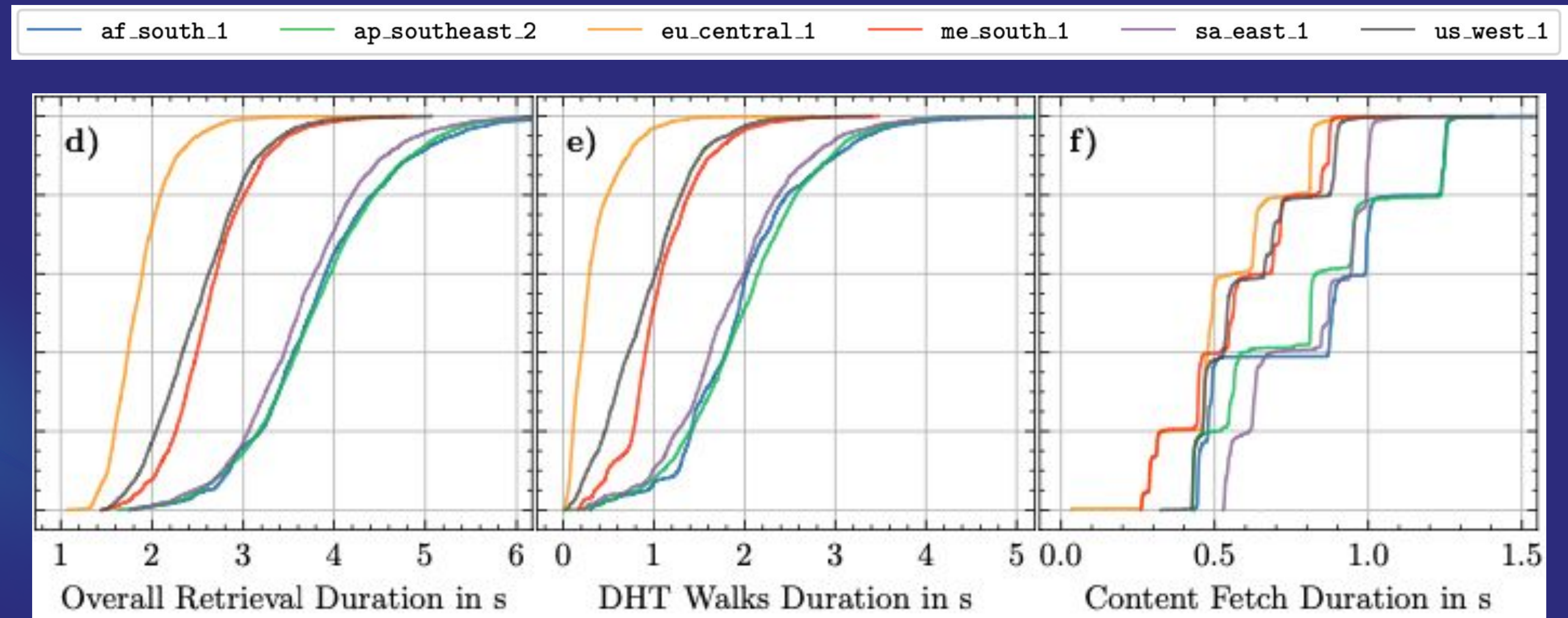


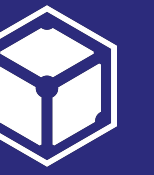


Opportunity 2: DHT Lookup Process

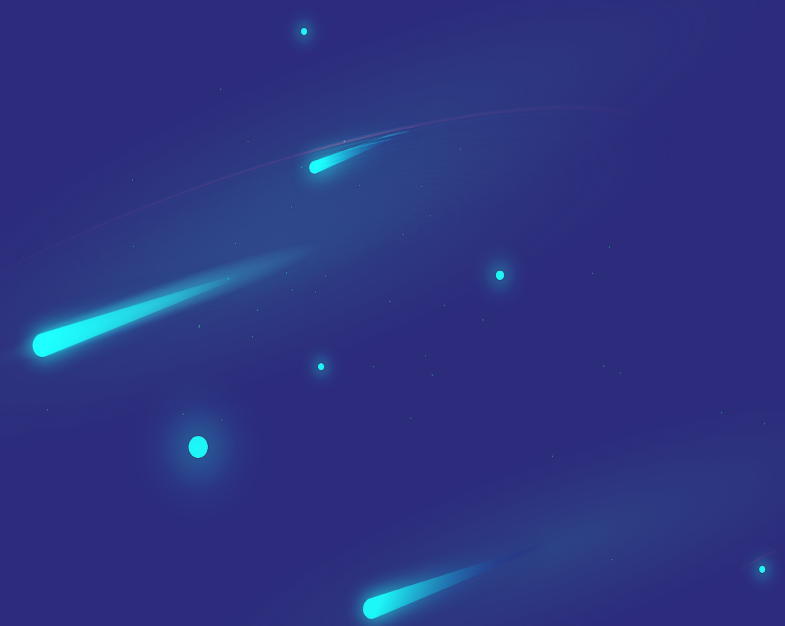
What we found out:

- Content fetching is reasonably fast (as expected).
- The DHT lookup is *very* fast - *very* often subsecond!





**We are producing weekly reports about
all this and a lot more at:
<https://stats.ipfs.network>**





You can read about all that and more in
our ACM Sigcomm 2022 paper.

CID: bafybeidbzzvjuzuf7yjet27sftttod5fowge3nzt3ybz5uxx1dsdonozq

Design and Evaluation of IPFS: A Storage Layer for the Decentralized Web

Dennis Trautwein
Protocol Labs &
University of Göttingen
dennis.trautwein@protocol.ai

Aravindh Raman
Telefonica Research
aravindh.raman@telefonica.com

Gareth Tyson
Hong Kong University of Science &
Technology (GZ)
gtyson@ust.hk

Ignacio Castro
Queen Mary University of London
i.castro@qmul.ac.uk

Will Scott
Protocol Labs
will@protocol.ai

Moritz Schubotz
FIZ Karlsruhe – Leibniz Institute for
Information Infrastructure
moritz.schubotz@fiz-karlsruhe.de

Bela Gipp
University of Göttingen
gipp@uni-goettingen.de

Yiannis Psaras
Protocol Labs
yiannis@protocol.ai

ABSTRACT

Recent years have witnessed growing consolidation of web operations. For example, the majority of web traffic now originates from a few organizations, and even micro-websites often choose to host on large pre-existing cloud infrastructures. In response to this, the “Decentralized Web” attempts to distribute ownership and operation of web services more evenly. This paper describes the design and implementation of the largest and most widely used Decentralized Web platform – the InterPlanetary File System (IPFS) – an open-source, content-addressable peer-to-peer network that

KEYWORDS

Interplanetary file system, content addressing, decentralized web, libp2p, content addressable storage

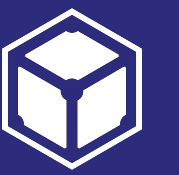
ACM Reference Format:

Dennis Trautwein, Aravindh Raman, Gareth Tyson, Ignacio Castro, Will Scott, Moritz Schubotz, Bela Gipp, and Yiannis Psaras. 2022. Design and Evaluation of IPFS: A Storage Layer for the Decentralized Web. In *ACM SIGCOMM 2022 Conference (SIGCOMM '22)*, August 22–26, 2022, Amsterdam, Netherlands. ACM, New York, NY, USA, 14 pages. <https://doi.org/10.1145/3544216.3544232>



CALL OUTS





Future/Open Projects

Get Involved!

Status: **Open**

Amount: 500 FIL

Funder: Protocol Labs

Next Deadline: 5/30/2022 - 10:59 PM

Ecosystem: IPFS

Discussion Link: [Link](#)

Application Required: Yes

App Deadline: 5/30/2022 - 10:59 PM

Connect to apply!

[Connect Wallet](#)

Grant

Effectiveness of Bitswap Discovery Process

Bitswap is involved in IPFS's content discovery and precedes the DHT walk. This step adds a delay of 1sec before content is resolved through the DHT. We want to find out whether this delay is worth by investigating Bitswap's success rate.

Author: [@yiannisbot](#)

Problem Description

Bitswap is involved in the content discovery process and precedes the DHT walk. Nodes ask all of their connected peers for the CID they're interested in, wait for 1sec to receive responses and in case of a negative result resort to the DHT.

Status: **Open**

Amount: 600 FIL

Funder: Protocol Labs

Next Deadline: 5/30/2022 - 10:59 PM

Ecosystem: Libp2p

Discussion Link: [Link](#)

Application Required: Yes

App Deadline: 5/30/2022 - 10:59 PM

Connect to apply!

[Connect Wallet](#)

Grant

IPFS Provider Record Liveness

Provider records are needed in IPFS in order to keep published content live and discoverable. We want to investigate whether provider records stay live for long enough before they are republished, given the high churn of network peers.

Author: [@yiannisbot](#)


Problem Description

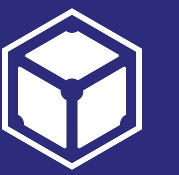
Provider records are replicated in the system to k=20 peers and are re-provided after 12hrs in the hope that, despite network churn, at least one of them will be alive to provide the record throughout the 12hr interval. However, we have not tested whether provider records indeed stay alive for 12hrs. In addition, we have found that the network has very high churn rate (at times in the order of 50% per hour).

- Check [ProbeLab page](#)
 - Lots of context and details
 - Come chat at #probe-lab on IPFS Discord
- Several grants Open at <https://app.radius.space/> ..
 - .. with several more coming!
- Check the Network Measurements GH repo: <https://github.com/protocol/network-measurements>
 - More than 15 Requests for Measurements (RFMs)
 - Many more coming!
 - Feel free to add more!



ProbeLab: Protocol Benchmarking & Optimization

 *Measuring networks is not an end, but a means to an end, which is to use the findings to: i) identify bottlenecks, ii) quantify the available space for improvement, iii) specify benchmarks, and iv) design protocol optimizations.*



Pointers

Get Involved!

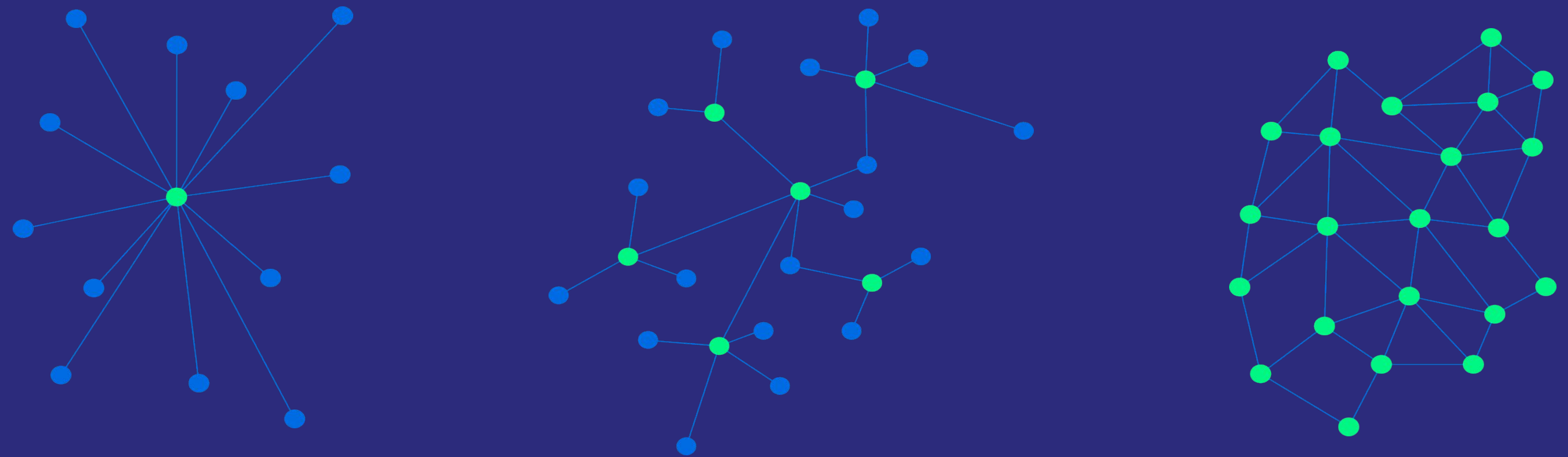
- **What I presented:**
 - Sigcomm'22 paper: [bafybeidbzyvjuzuf7yjet27sftttod5fowge3n3r3ybz5uxxldsdonozq](https://arxiv.org/abs/2205.12345)
 - <https://gateway.ipfs.io/ipfs/CID>
- **General Pointers:**
 - Video tutorials: <https://research.protocol.ai/tutorials/resnetlab-on-tour/>
 - Docs: <https://docs.ipfs.io>
 - Blog: <https://ipfs.io/blog>
 - Forum: <https://discuss.ipfs.io> / <https://discuss.libp2p.io>
 - ProtoSchool Tutorials: <https://proto.school>
- **Store your own files for free:**
 - <https://web3.storage>
 - <https://nft.storage>
 - <https://pinata.cloud>
 - <https://files.chainsafe.io/>

Research Workshop Invitation

July 10th - Bologna



DINPS Workshop @ ICDCS 2022

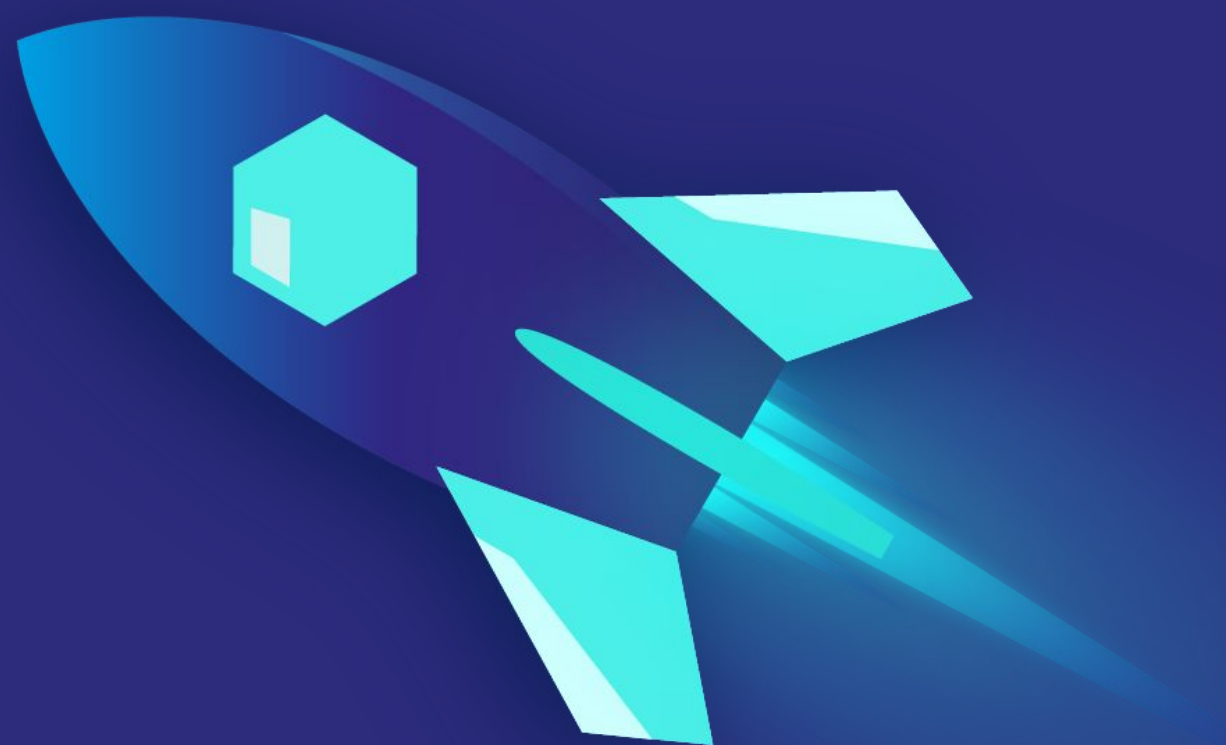


Decentralized Internet, Networks, Protocols, and Systems - DINPS

A research workshop organized alongside a top conference in the area of Networks and Distributed Systems!

Your opportunity to showcase your work to an academic audience and attract the attention of researchers and PhD students!

Check all the details at: <https://research.protocol.ai/sites/dinps>





Thank you!

Let's chat over a drink 🍷

yiannis@protocol.ai