Design and Evaluation of IPFS

A Storage Layer for the Decentralized Web

ACM SigCOMM 2022, Amsterdam, Netherlands

Authors: Dennis Trautwein, Aravindh Raman, Gareth Tyson, Ignacio Castro, Will Scott, Moritz Schubotz, Bela Gipp, Yiannis Psaras

IETF 118 - IRTF Open Meeting, 2023-11-09

Dennis Trautwein

Research Engineer @Protocol Labs

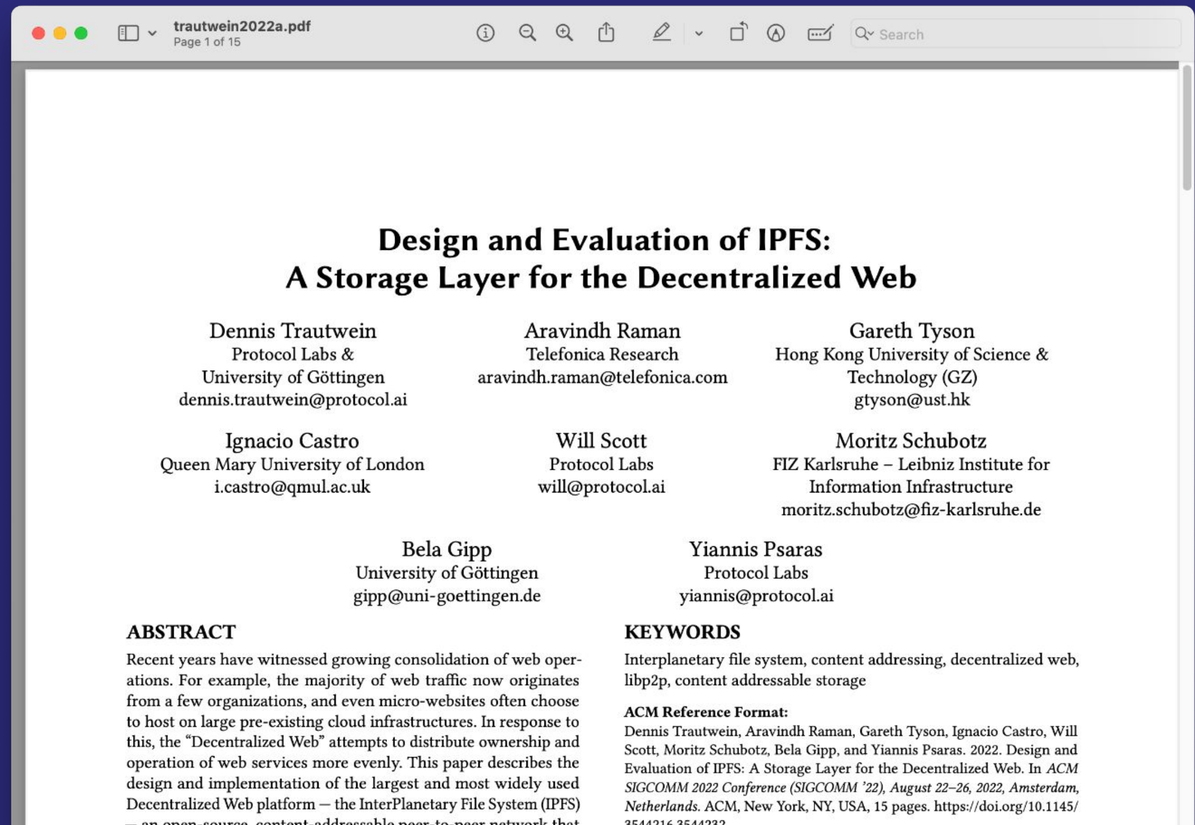
Ph.D. candidate @University of Göttingen

Today's

Agenda

- What is IPFS?
- Design
- Evaluation
- Where to go from here?

QmZtCQPnJ5qnKgRA3C1yWBfbjNxusH5NuRBPYHfrwh8Us1



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, 15 pages. https://doi.org/10.1145/3544216.3544232

WHAT IS IPFS?

What is IPFS?

In Words

stands for the InterPlanetary File System

The IPFS stack is a suite of specifications and tools that share two key characteristics

1) **Content Addressing using CIDs**
2) **Transport Agnosticity**

IPFS is **not** a blockchain.

[1] IPFS Specs *https://specs.ipfs.tech/*
[2] IPFS Docs *https://docs.ipfs.tech/concepts/implementations/*

What is IPFS?

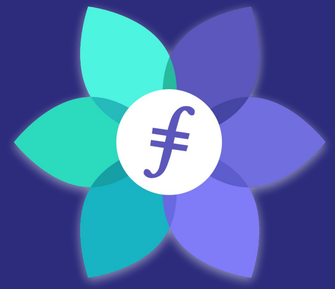
In Numbers

- 10+ implementations
- Operational since 2015
- ~300k nodes / week
- ~3M users / day
- ~120M requests / day

Kubo

Helia

Elastic IPFS



Lotus

and more...

[1] ProbeLab *https://probelab.io/*
[2] IPFS Docs *https://docs.ipfs.tech/concepts/ipfs-implementations/*

DESIGN

Design

Content Addressing

Simplistically: IPFS uses the hash of the content stored in the system as its content identifier (CID)

bafybeigdyrzt5sfp7udm7hu76uh7y26nf3efuylqabf3oclgtqy55fbzdi

Design

Content Addressing

In practice: there is much more sophistication in the structure of a CID

bafybeigdyrzt5sfp7udm7hu76uh7y26nf3efuylqabf3oclgtqy55fbzdi

⇓

<multibase>(cid-version || multicodec || multihash)

Design

Content Addressing

Advantages

- Decouples content from hosts
- Data integrity
- Deduplication
- Alleviate backbone addiction

Challenges

- Access Control
- Discoverability

Design

Peer to Peer Network Structure

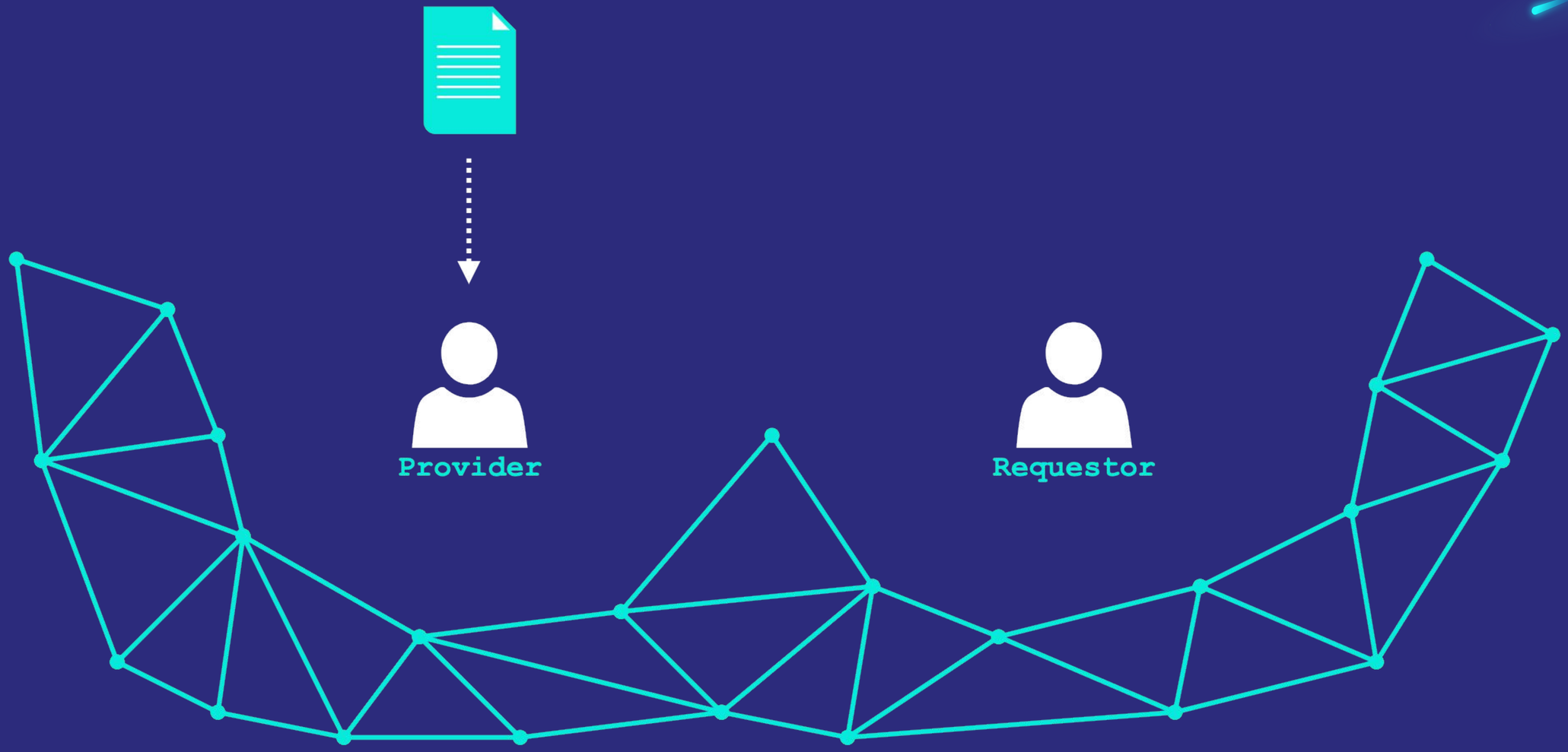
- IPFS uses a Kademlia-based DHT for the P2P network's structure
- It enables the system to be open and permissionless
- Two types of records:
  - *Provider-Records:* CID → PeerID
  - *Peer-Records:* PeerID → Network Addresses

Design

Content Lifecycle

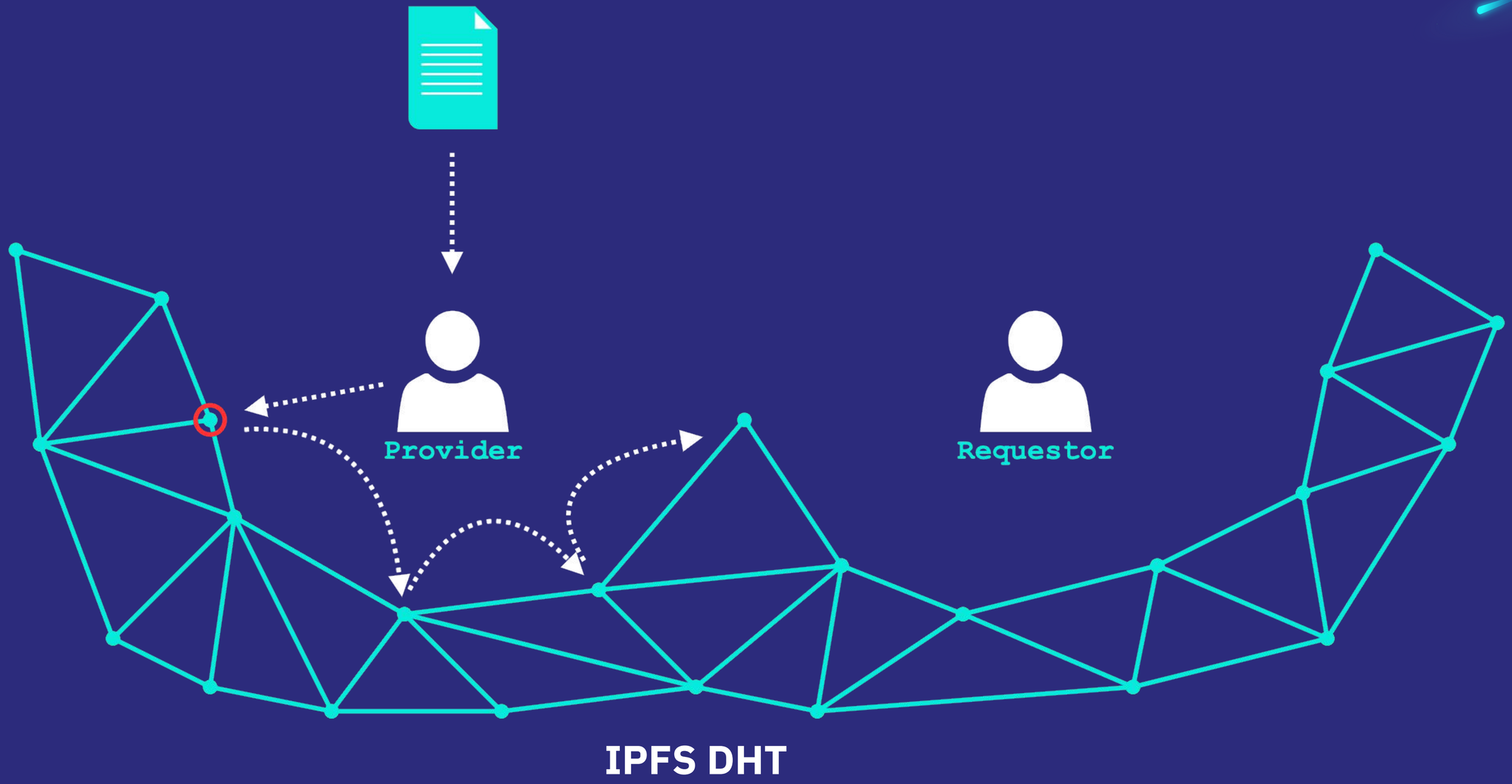
Design

Content Lifecycle



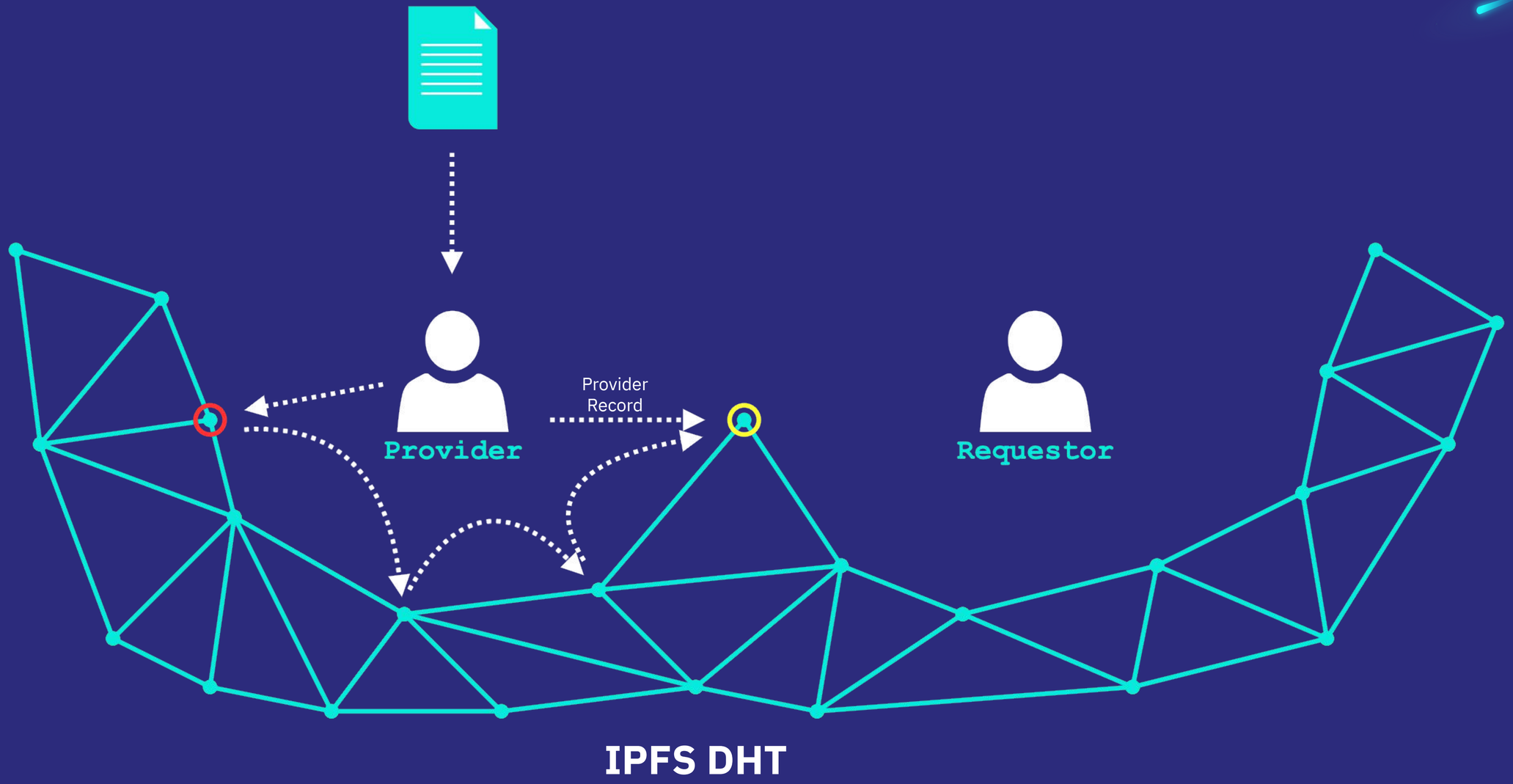
Design

Content Lifecycle



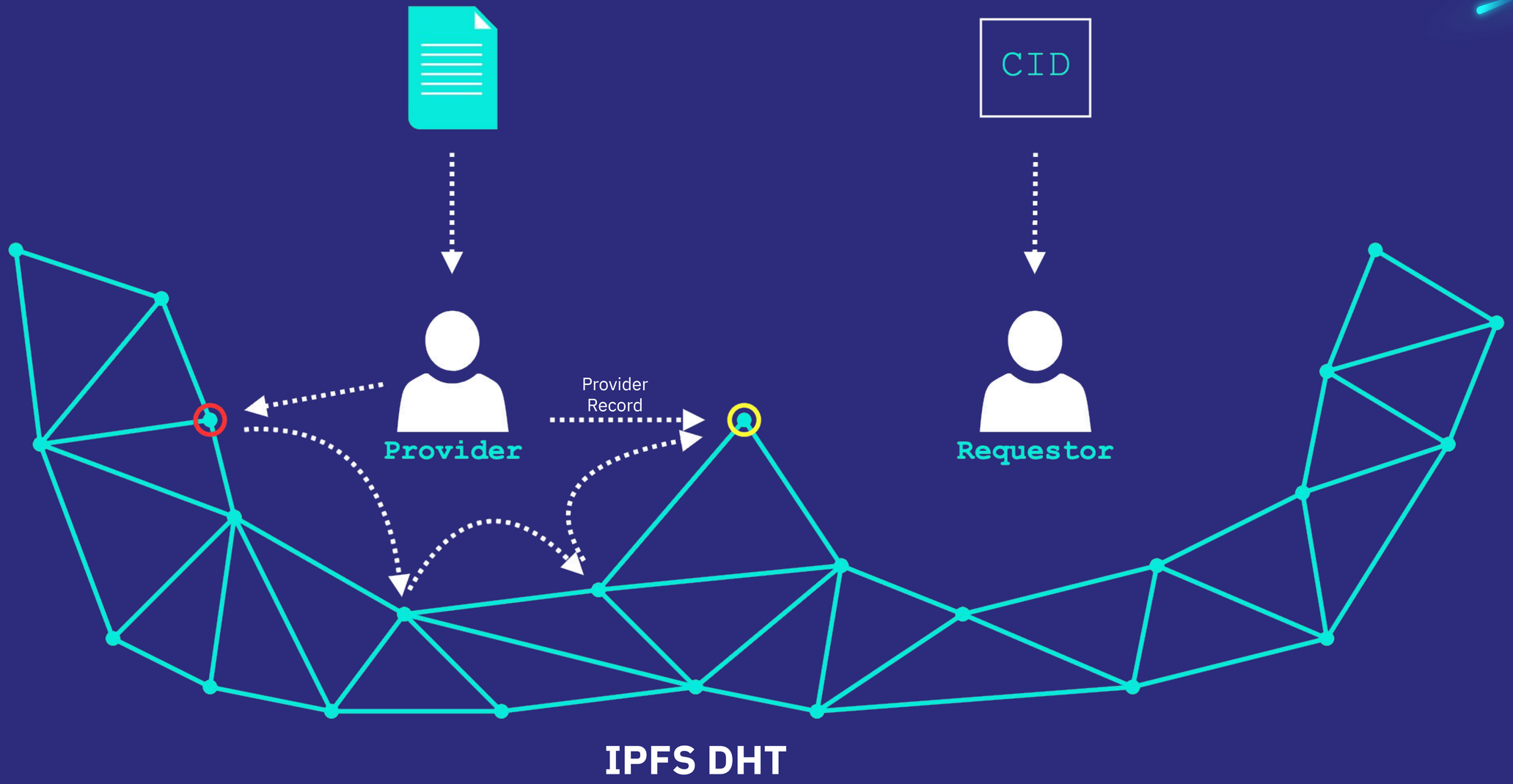
Design

Content Lifecycle



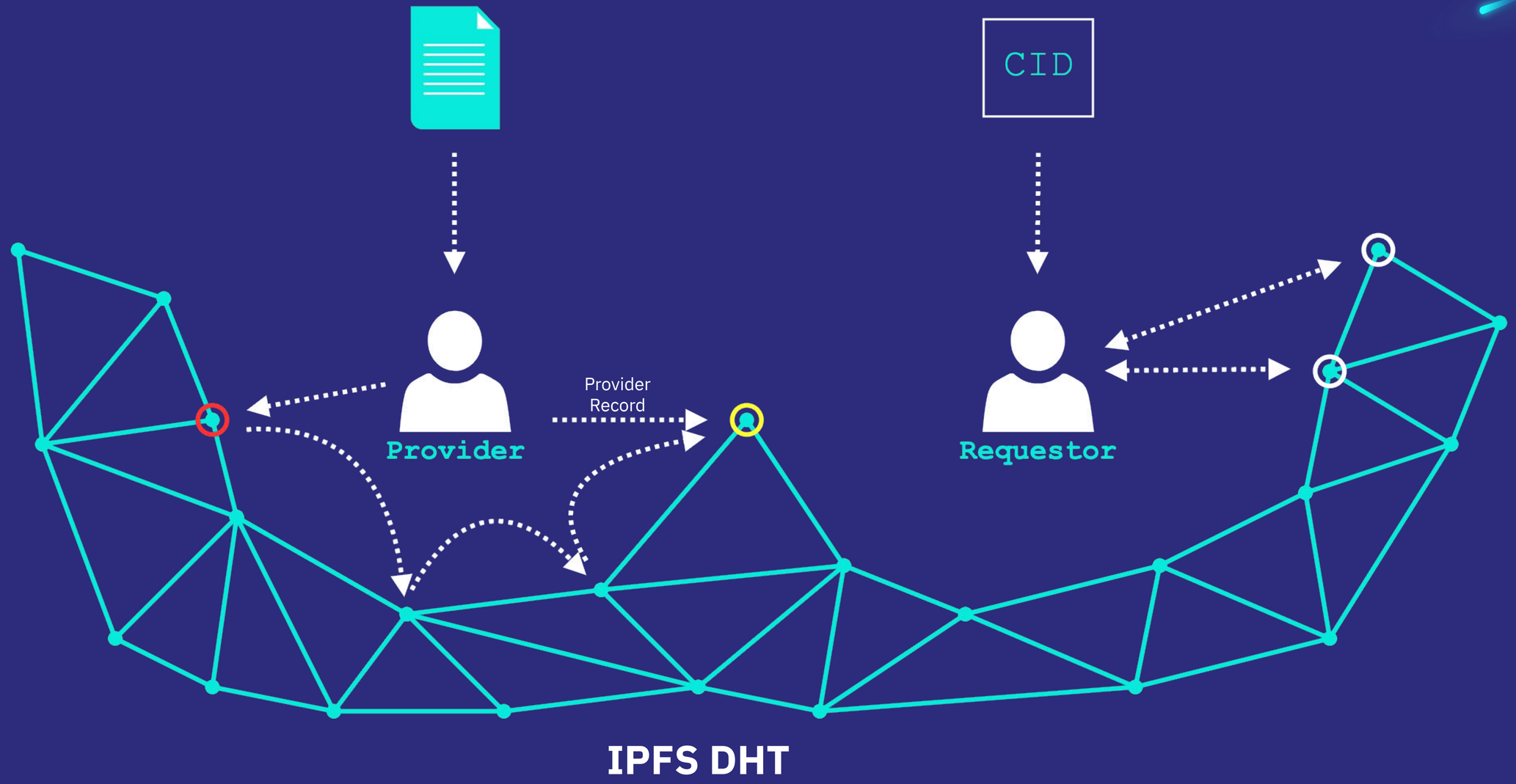
Design

Content Lifecycle



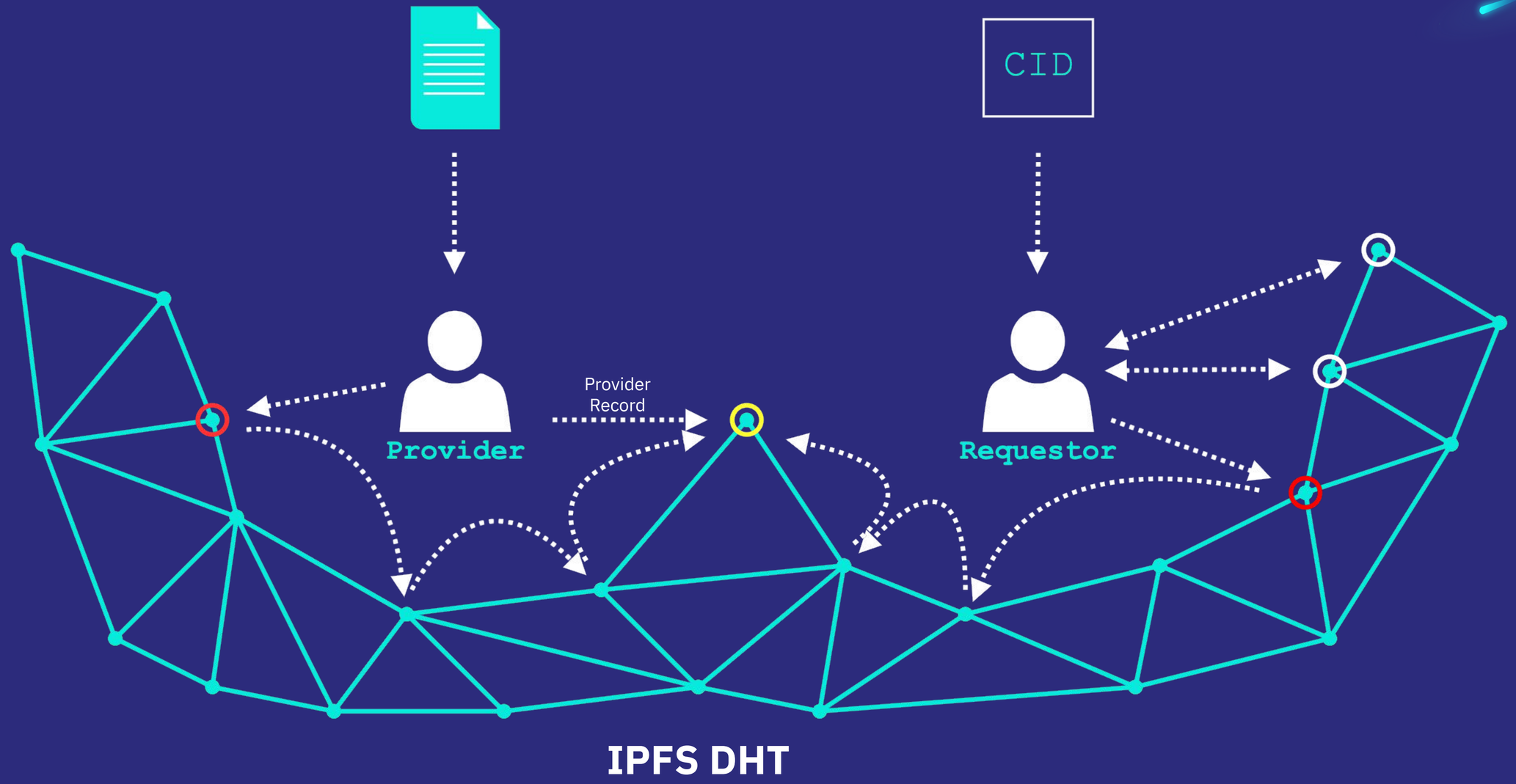
Design

Content Lifecycle



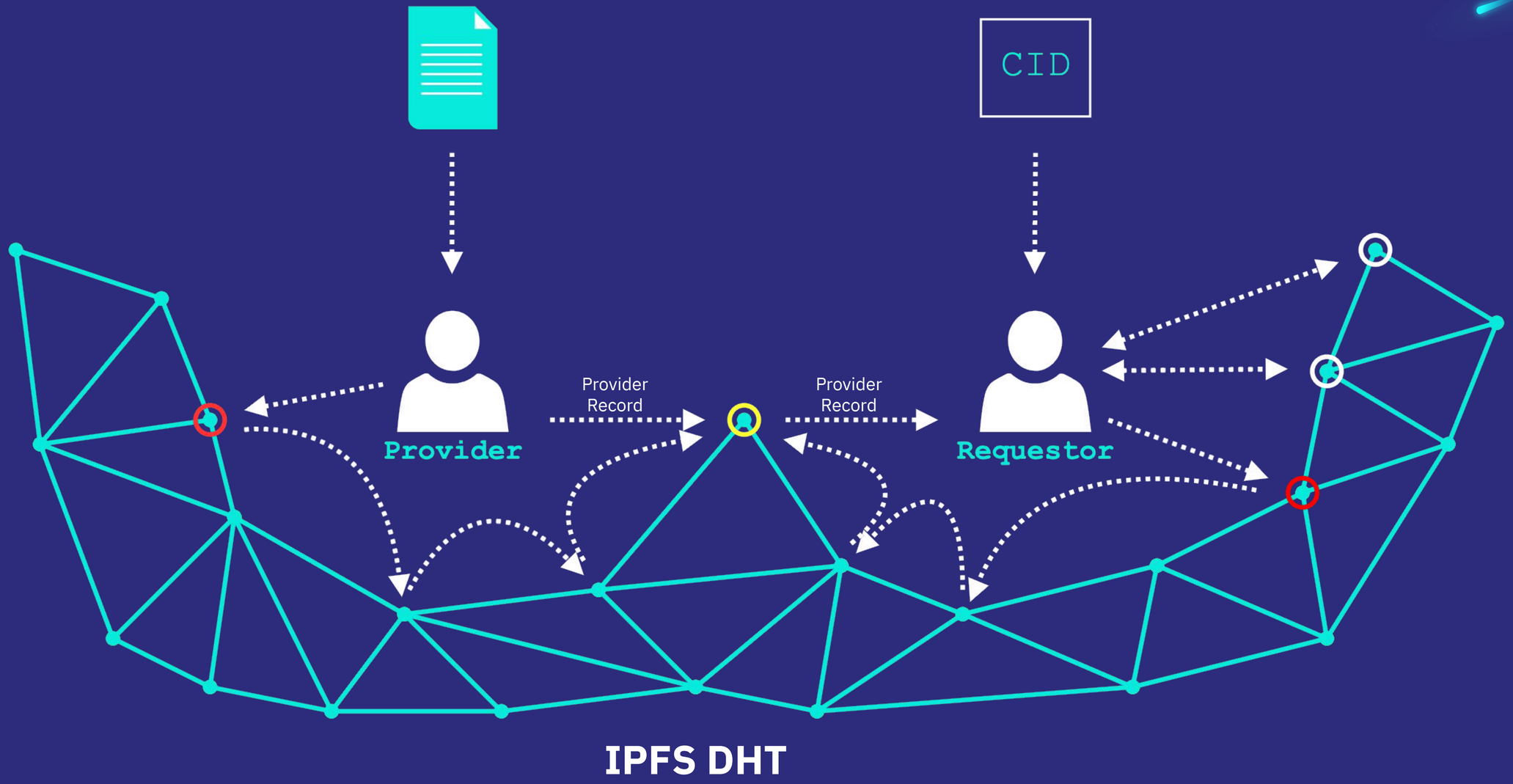
Design

Content Lifecycle



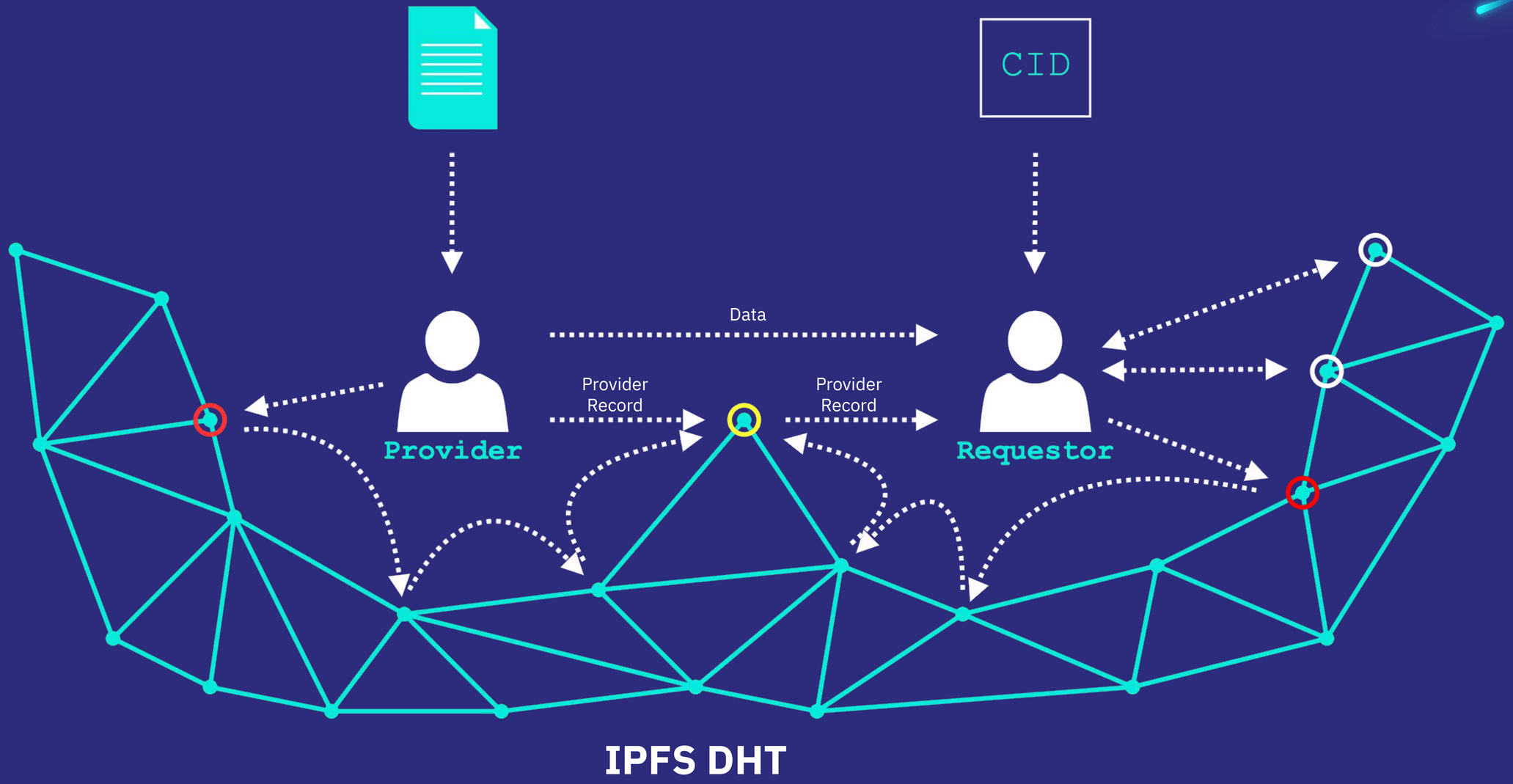
Design

Content Lifecycle



Design

Content Lifecycle



EVALUATION

Evaluation

Methodologies

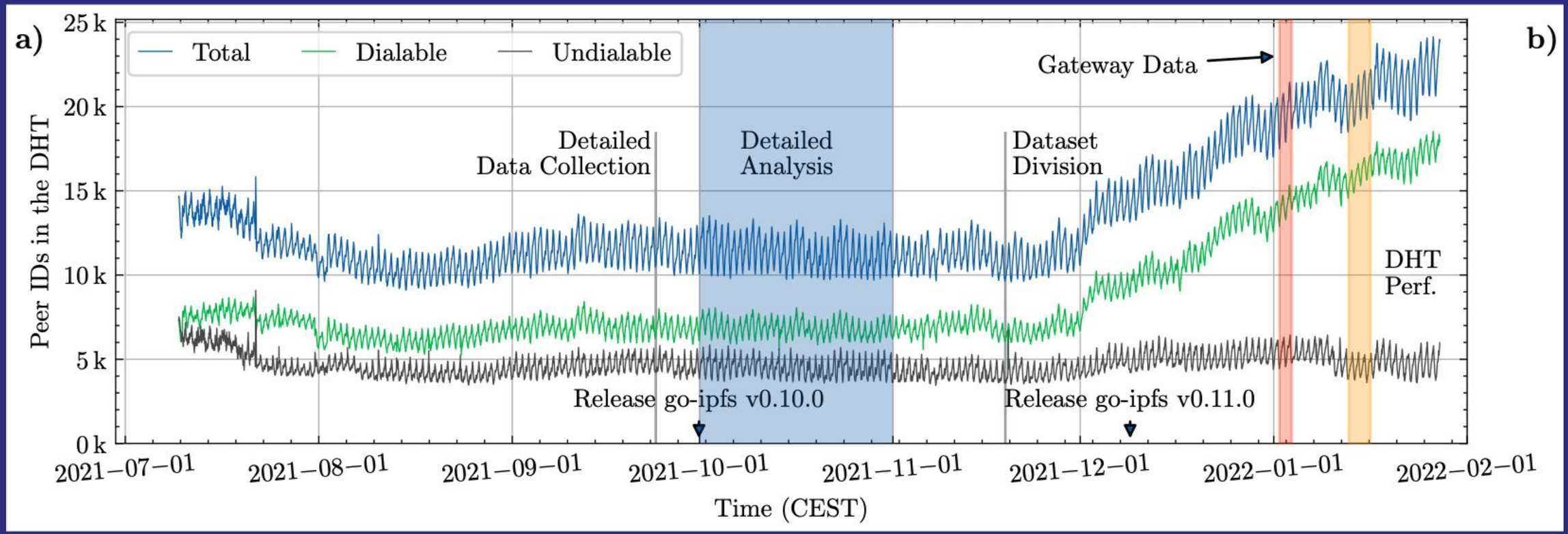
Three complementary methodologies covering the operational spectrum:

- **Crawls:** Continuous crawling and monitoring
- **Probes:** Performance measurements through controlled nodes
- **Logs:** Infrastructure usage log analysis (not in this presentation)

Evaluation

Context

The network is a moving target



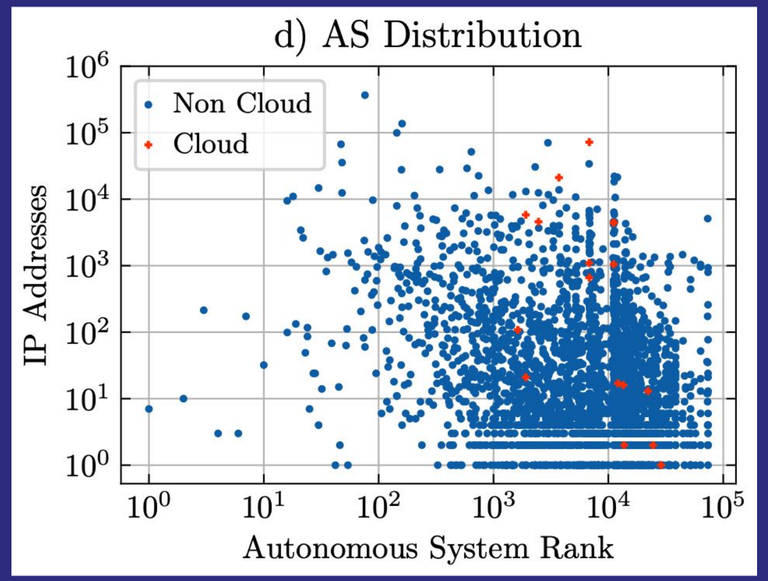
[1] Network Crawls with Nebula: *https://github.com/dennis-tra/nebula*

Evaluation

Crawls

Continuous Network Monitoring

- Full network crawls every 30m
- 9.5k crawls
- **Monitors uptime**



- **~ 464k** IP-Addresses
- **> 150** Countries
- **> 2700** ASs

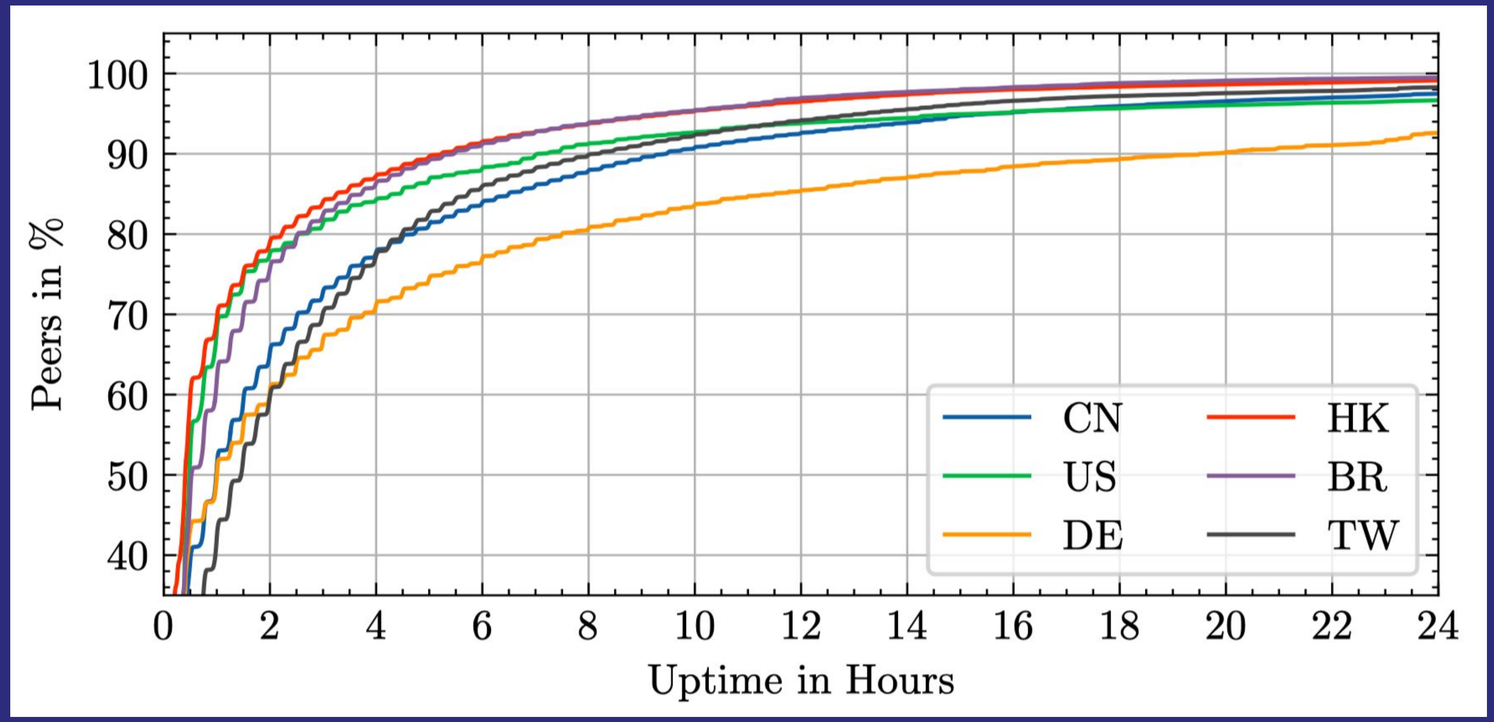
“The Cloud Strikes Back: Investigating the Decentralization of IPFS” – Balduf et al., IMC '23

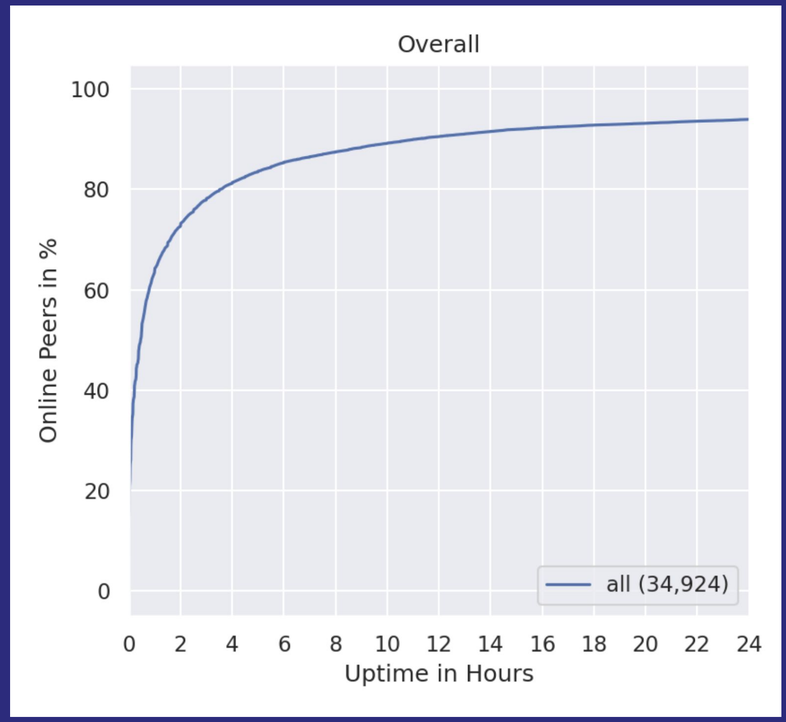
[1] Network Crawls with Nebula: *https://github.com/dennis-tra/nebula*

Evaluation

Peer Churn

Influences several network-wide DHT parameters like record replication or routing table refresh rate





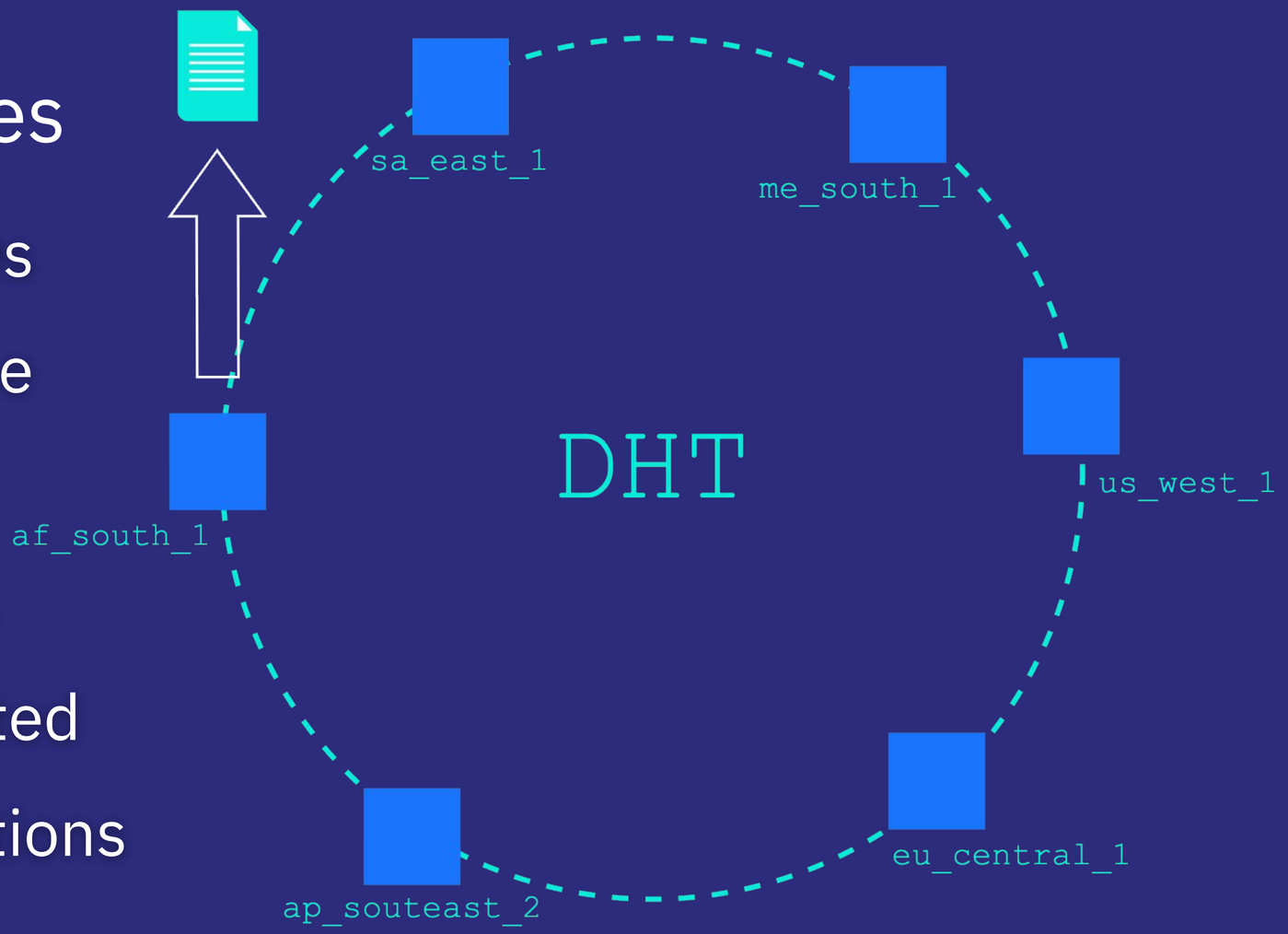
[1] *https://github.com/plprobelab/network-measurements/blob/master/reports/2023/calendar-week-43/ipfs/README.md#churn*

Evaluation

Probes

Controlled Network Nodes

- DHT Servers publish new CIDs
- DHT Servers communicate the CIDs to the clients
- DHT Clients request the CIDs
- The request process is repeated from several geographic locations

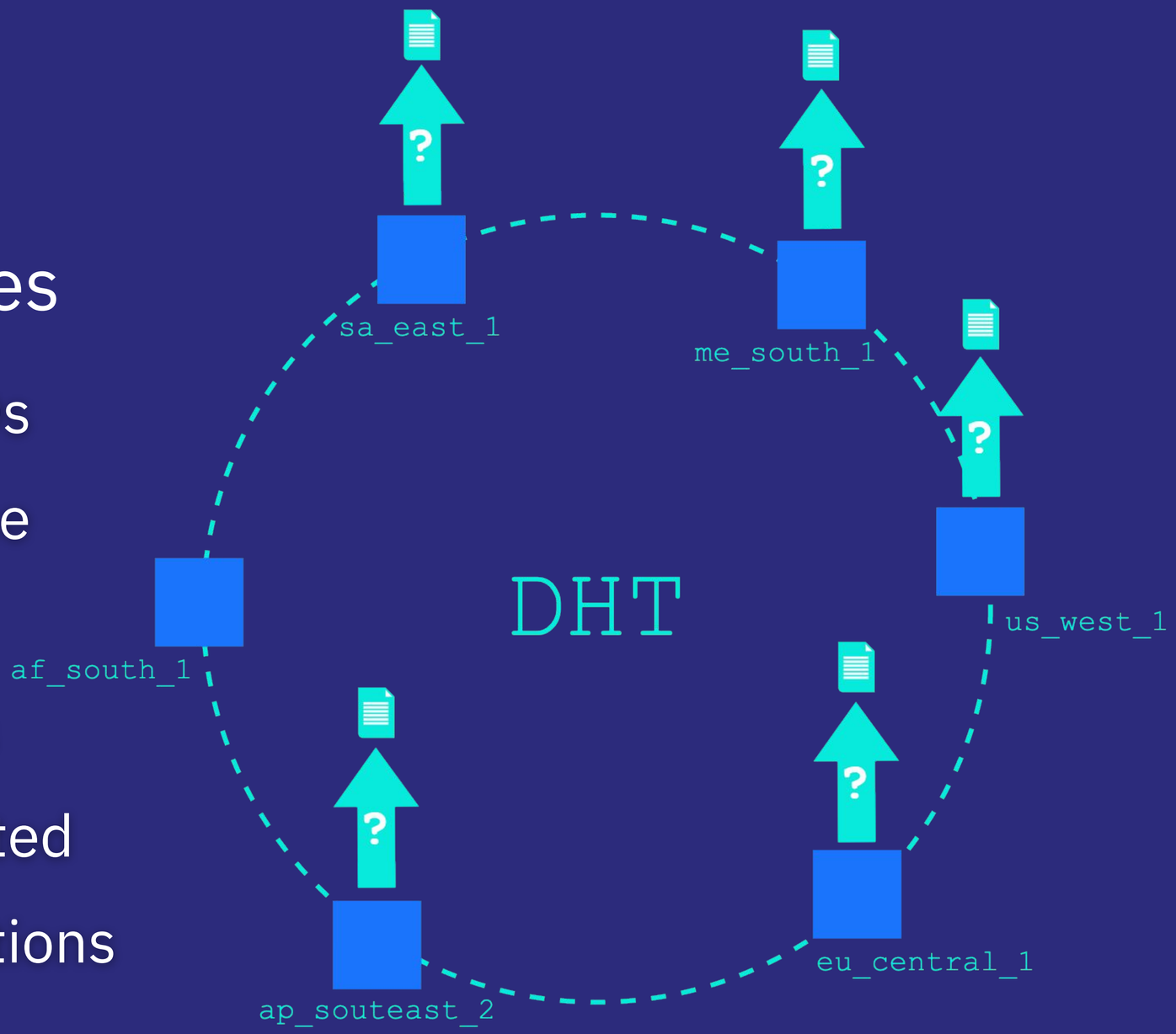


Evaluation

Probes

Controlled Network Nodes

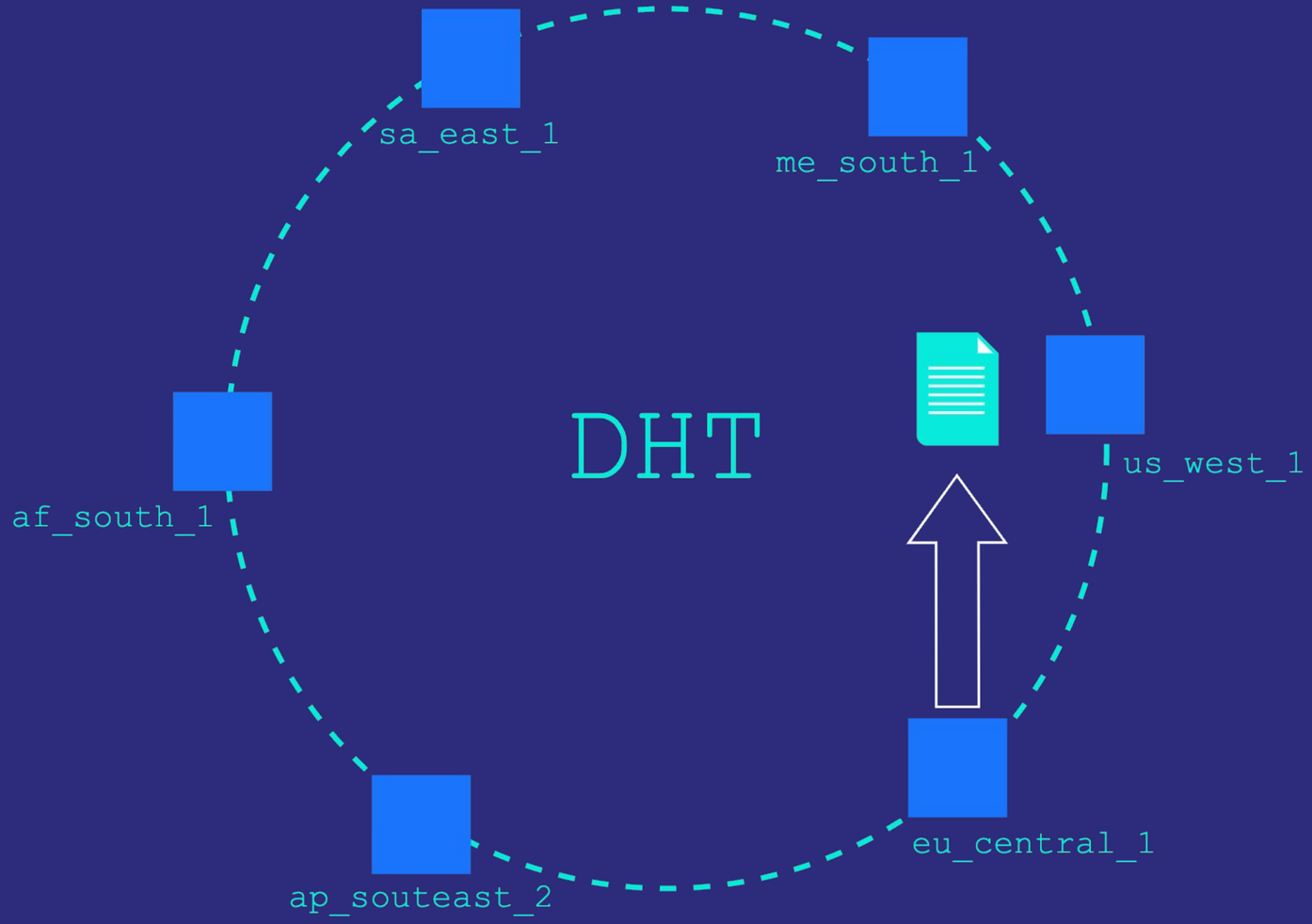
- DHT Servers publish new CIDs
- DHT Servers communicate the CIDs to the clients
- DHT Clients request the CIDs
- The request process is repeated from several geographic locations



Evaluation

Probes

Repeat ...

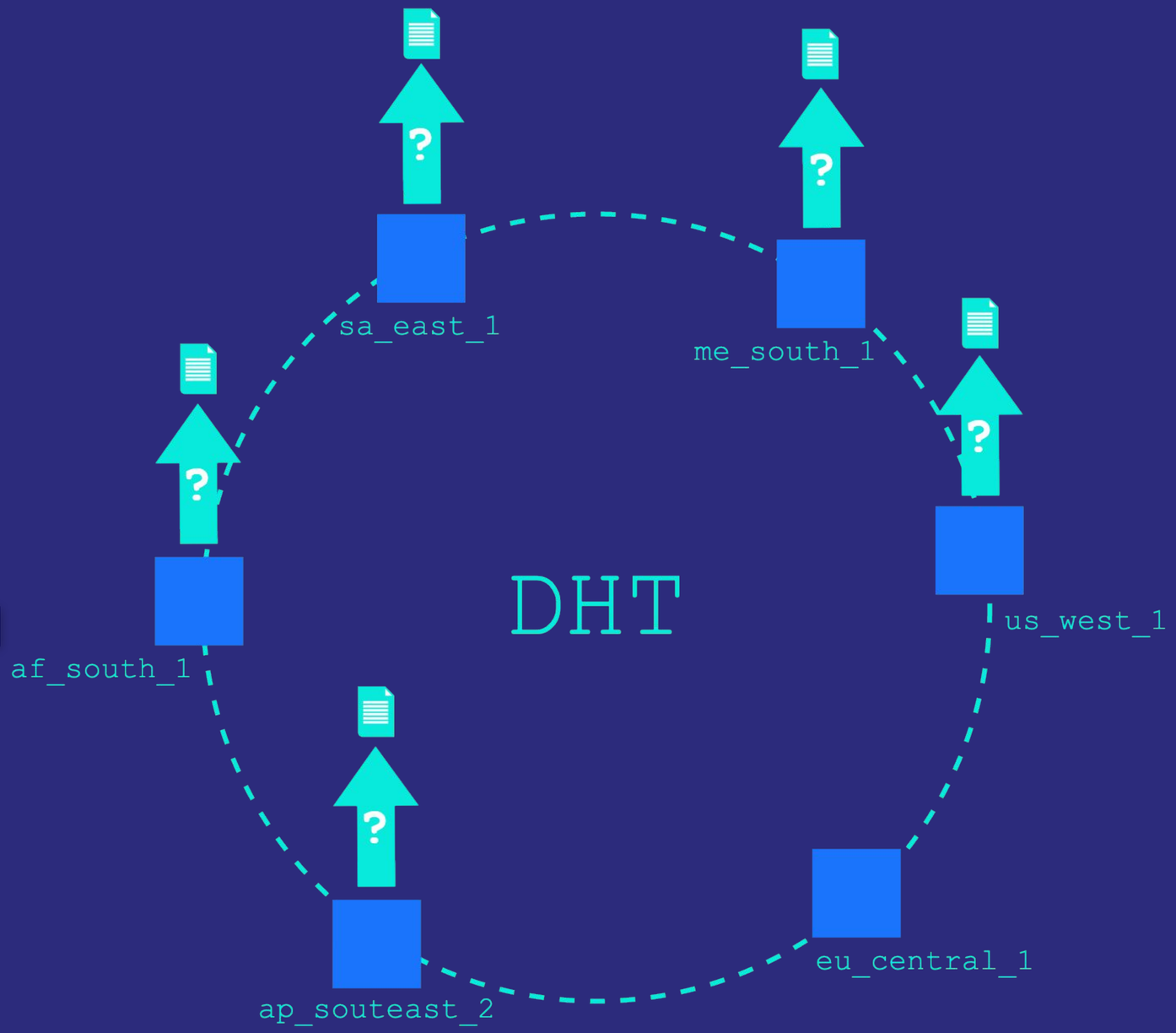


Evaluation

Probes

Repeat ...

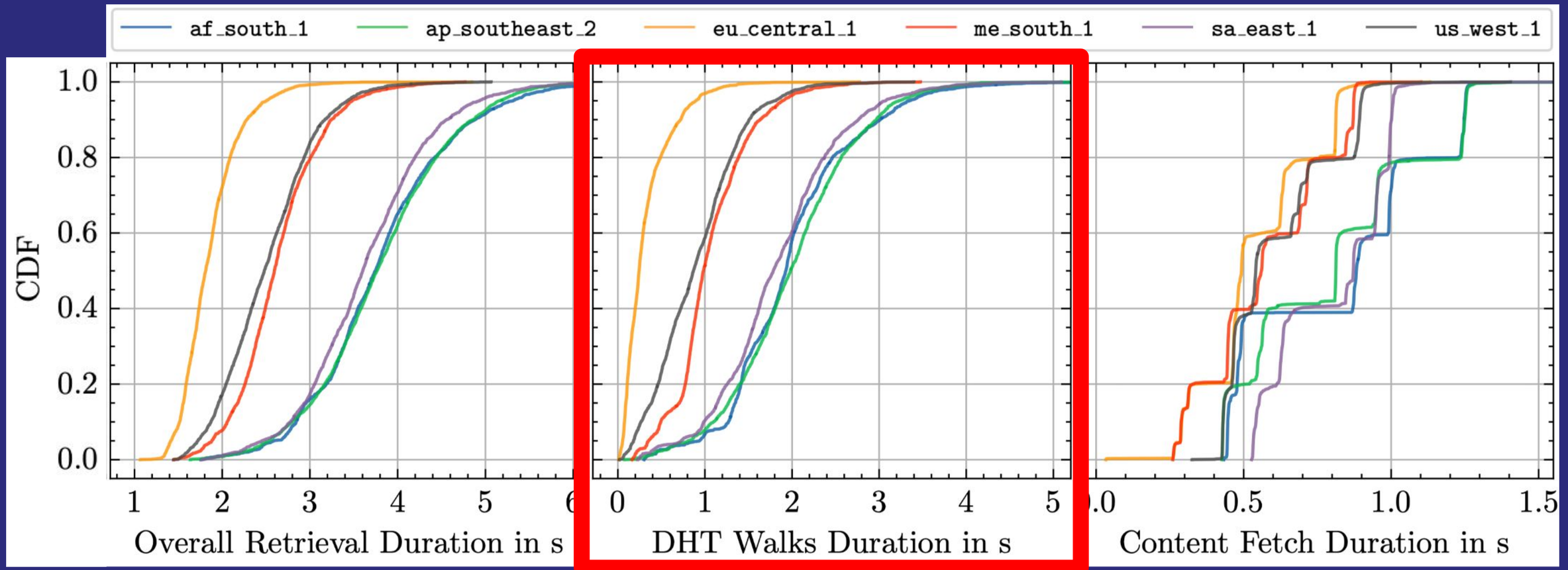
- ... more than 3k CIDs published
- ... more than 14k CIDs retrieved



Evaluation

DHT Lookup Latency

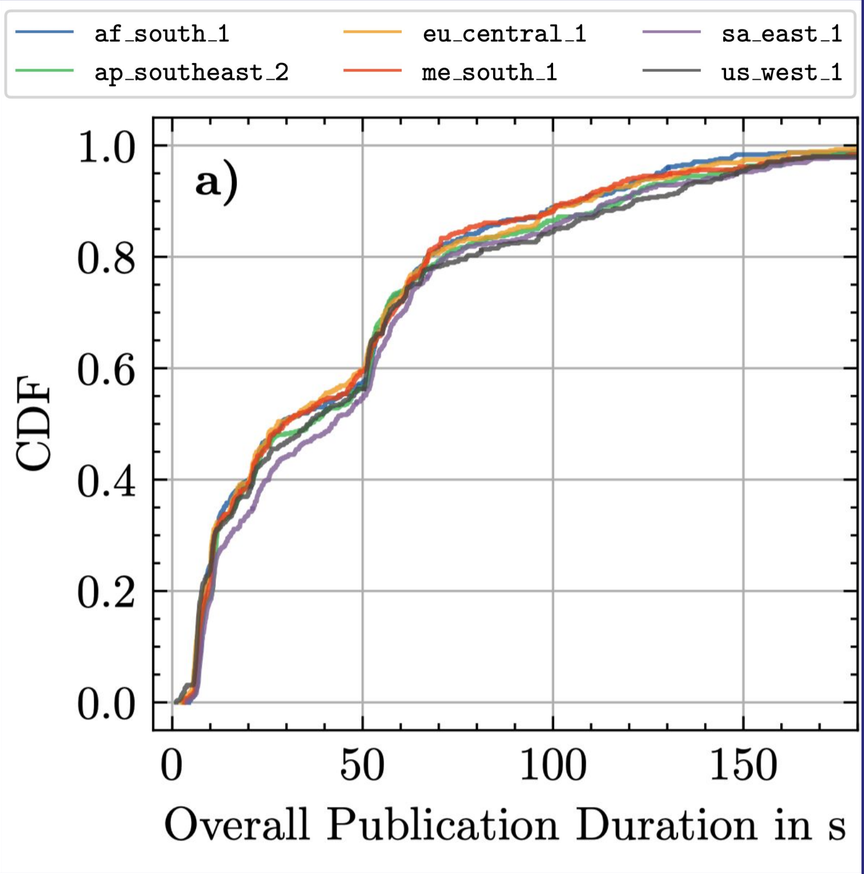
80% of requests from EU resolve in < 500ms

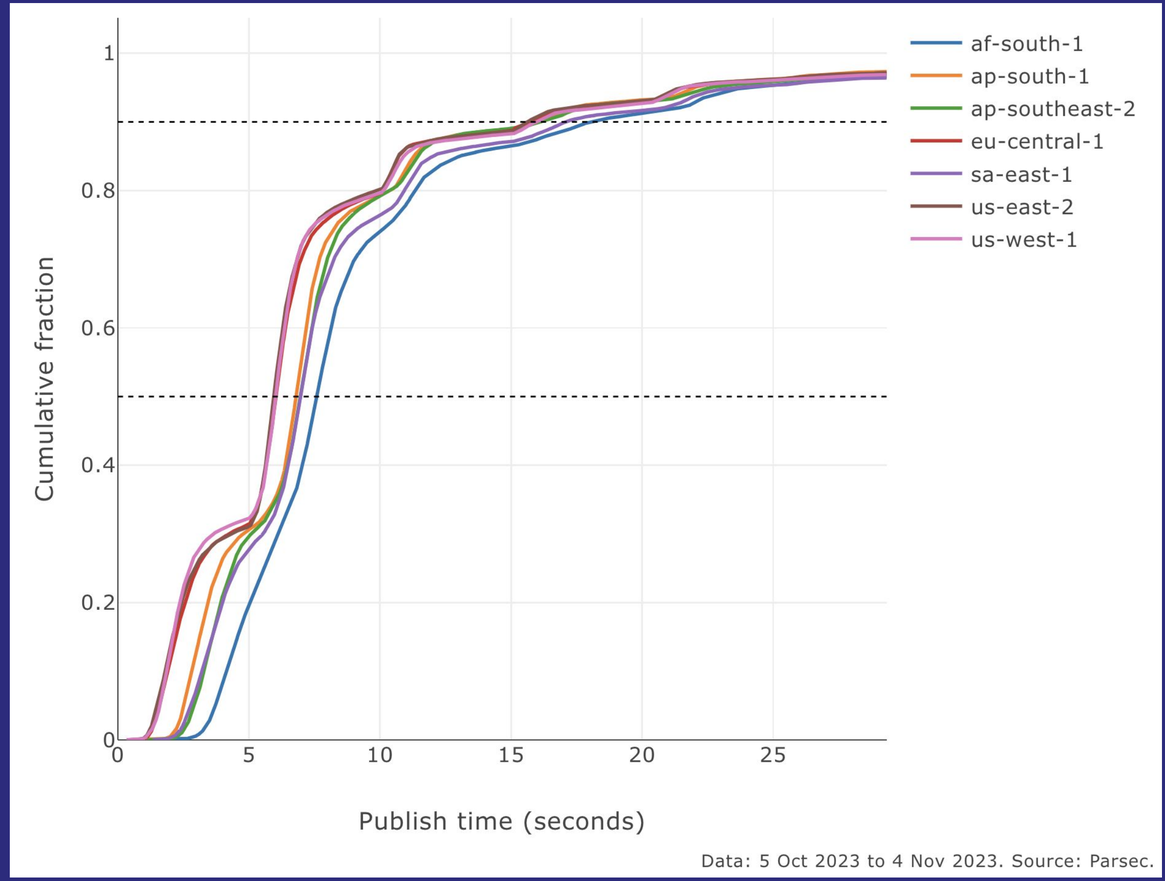


Evaluation

DHT Publication Latency

Orders of magnitude larger than the lookup latency





Evaluation

Cliffhanger

More detailed analysis in our paper:

- IPFS Network Design Details
- Public Gateway Usage Log analysis
- Cloud Provider Dependence
- Geographical Distribution of network Participants
- Lookup performance compared to HTTPS – “Request Stretch”

WHERE TO GO FROM HERE?

Where to go from here?

Datasets

Use our datasets!

- Network Crawls

  `bafybeigkawbwjxa325rhul5vodzxb5uof73neszqe6477nilzziw5k5oj4`

- Probe Performance Data

  `bafybeid7ilj4k4rq27lg45nceq4akdpetav6bcujgiym6vch5ml24tk2t4`

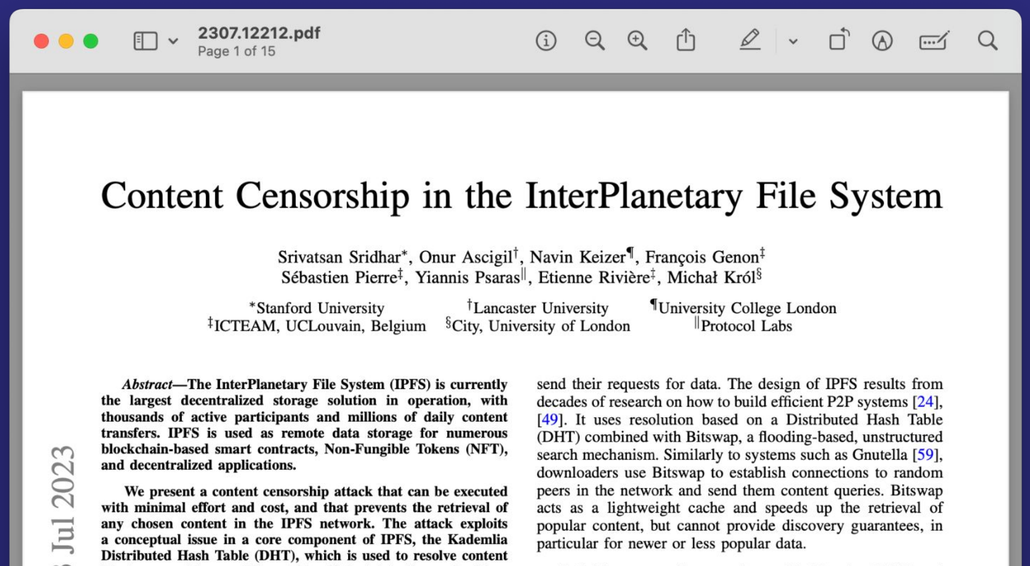
- Infrastructure Usage Logs

  `bafybeiftyvcar3vh7zua3xakxkb2h5ppo4giu5f3rkpsqgcfh7n7axxnsa`

Where to go from here?

Reading Recommendations

Qmbu34GKt1Z5npMiBcsMPTo5VFNuSADNctxZU4QT2iZuGj



Content Censorship in the InterPlanetary File System

Srivatsan Sridhar*, Onur Ascigil†, Navin Keizer¶, François Genon‡
Sébastien Pierre‡, Yiannis Psaras‖, Etienne Rivière‡, Michał Król§

*Stanford University †Lancaster University ¶University College London
‡ICTEAM, UCLouvain, Belgium §City, University of London ‖Protocol Labs

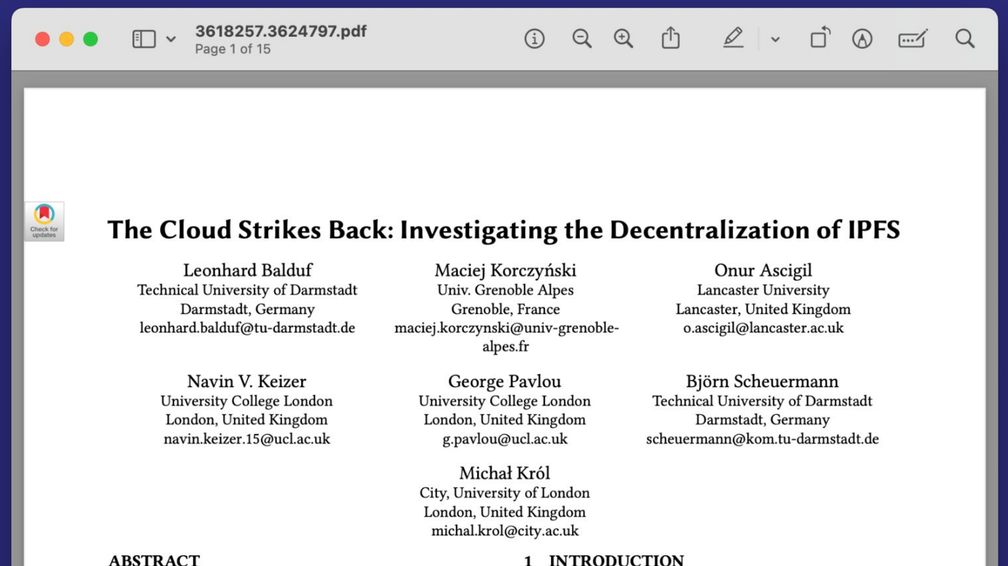
***Abstract*—The InterPlanetary File System (IPFS) is currently the largest decentralized storage solution in operation, with thousands of active participants and millions of daily content transfers. IPFS is used as remote data storage for numerous blockchain-based smart contracts, Non-Fungible Tokens (NFT), and decentralized applications.**

We present a content censorship attack that can be executed with minimal effort and cost, and that prevents the retrieval of any chosen content in the IPFS network. The attack exploits a conceptual issue in a core component of IPFS, the Kademlia Distributed Hash Table (DHT), which is used to resolve content

send their requests for data. The design of IPFS results from decades of research on how to build efficient P2P systems [24], [49]. It uses resolution based on a Distributed Hash Table (DHT) combined with Bitswap, a flooding-based, unstructured search mechanism. Similarly to systems such as Gnutella [59], downloaders use Bitswap to establish connections to random peers in the network and send them content queries. Bitswap acts as a lightweight cache and speeds up the retrieval of popular content, but cannot provide discovery guarantees, in particular for newer or less popular data.

6 Jul 2023

QmVU2rxWtbHT3vUgFAjbwVyNJAb29gFwz6VP9QRHfVtDKz



The Cloud Strikes Back: Investigating the Decentralization of IPFS

Leonhard Balduf
Technical University of Darmstadt
Darmstadt, Germany
leonhard.balduf@tu-darmstadt.de

Maciej Korczyński
Univ. Grenoble Alpes
Grenoble, France
maciej.korczynski@univ-grenoble-alpes.fr

Onur Ascigil
Lancaster University
Lancaster, United Kingdom
o.ascigil@lancaster.ac.uk

Navin V. Keizer
University College London
London, United Kingdom
navin.keizer.15@ucl.ac.uk

George Pavlou
University College London
London, United Kingdom
g.pavlou@ucl.ac.uk

Björn Scheuermann
Technical University of Darmstadt
Darmstadt, Germany
scheuermann@kom.tu-darmstadt.de

Michał Król
City, University of London
London, United Kingdom
michal.krol@city.ac.uk

ABSTRACT

1 INTRODUCTION

Where to go from here?

ProbeLab

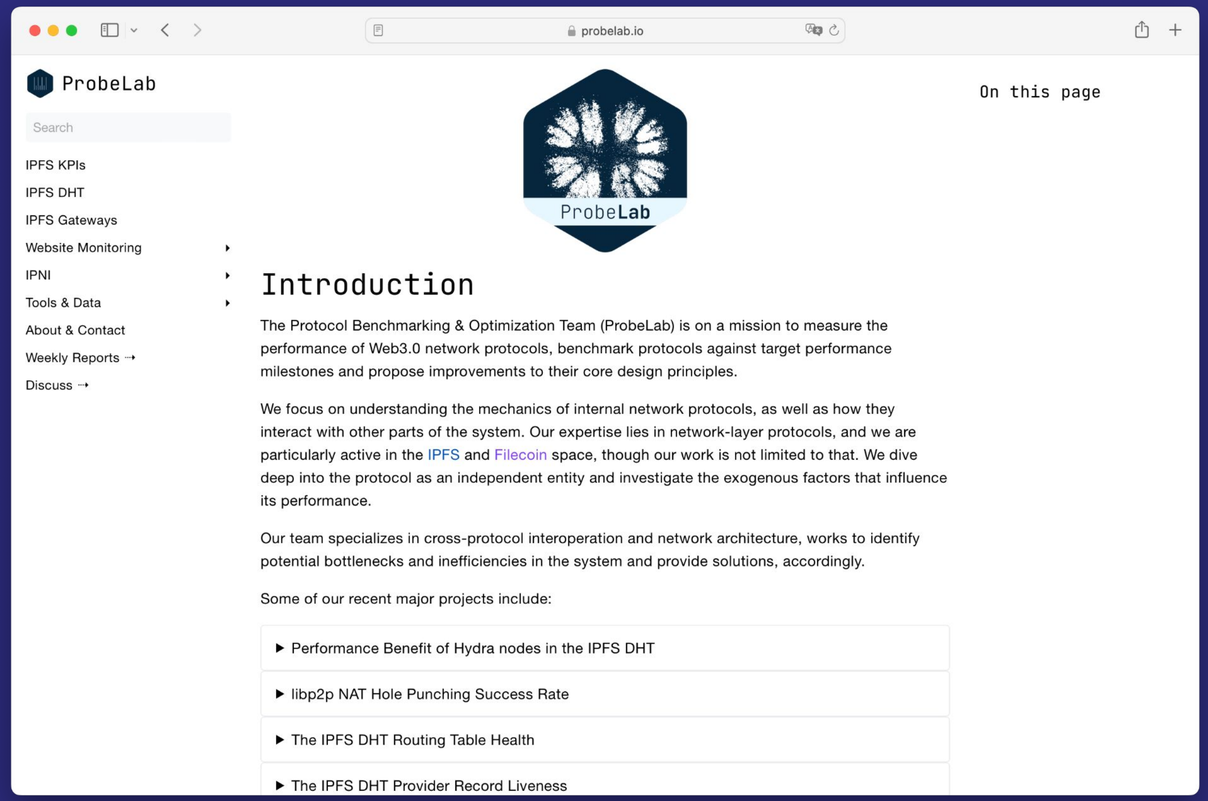
Visit

- ProbeLab: **https://probelab.io**
- Weekly Reports at **https://stats.ipfs.network/**

Future Work

- Content availability, severe network conditions, content routing latency, broaden focus

https://probelab.io



IETF 118 - IRTF Open Meeting, 2023-11-09

Dennis Trautwein

@dennis-tra on *GitHub*
@dtrautwein.eu on *Bluesky*
https://dtrautwein.eu on the *Web*
dennis@protocol.ai via *Email*

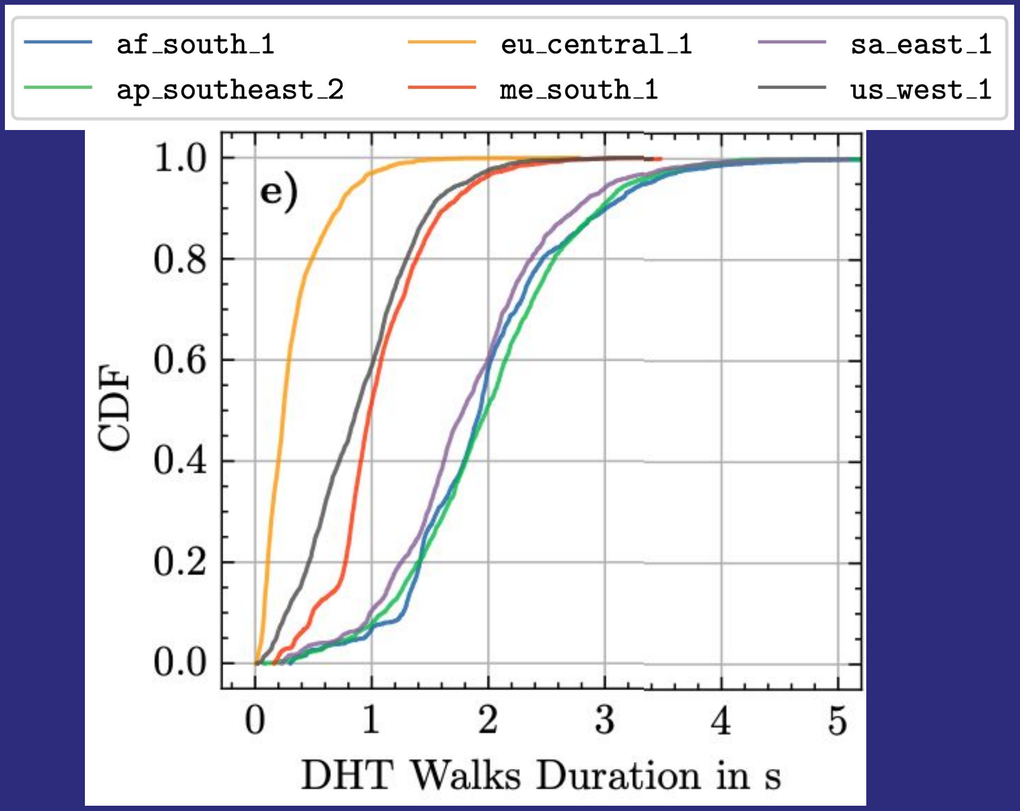
Thank you!

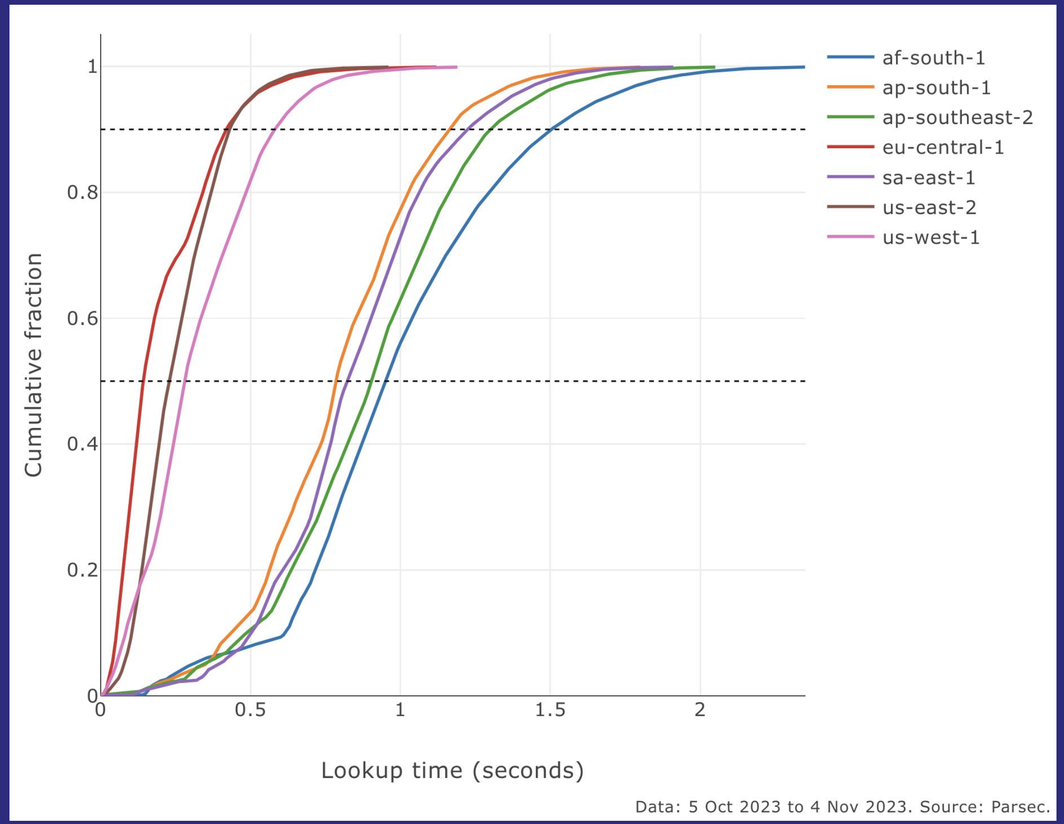
BACKUP SLIDES

Evaluation

DHT Lookup Latency

80% of requests from EU/NA resolve in < 500ms

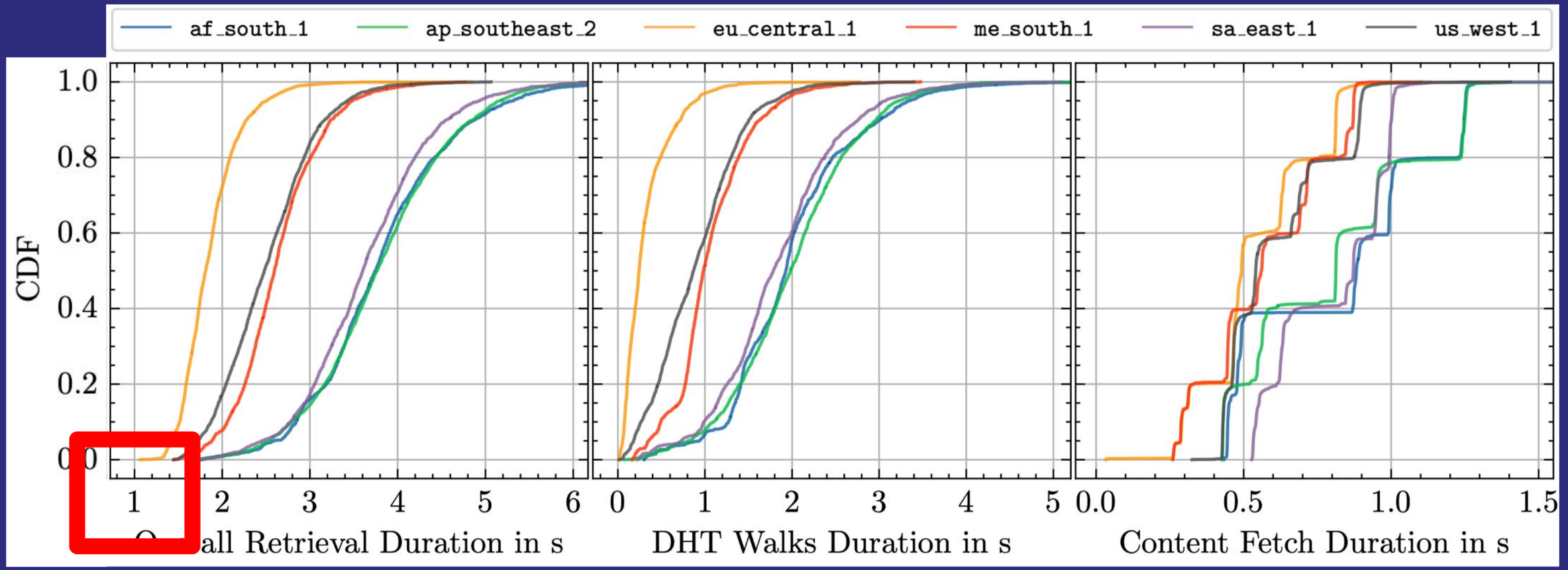




Evaluation

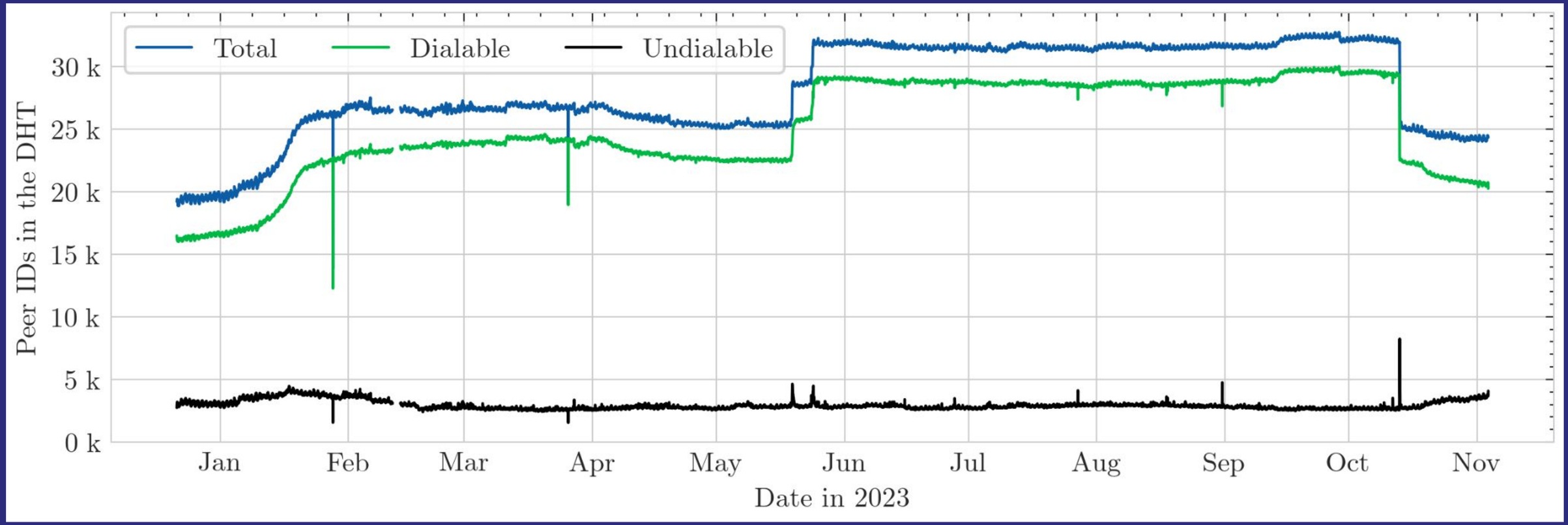
DHT Lookup Latency

Constant 1s lookup delay



Evaluation

Context



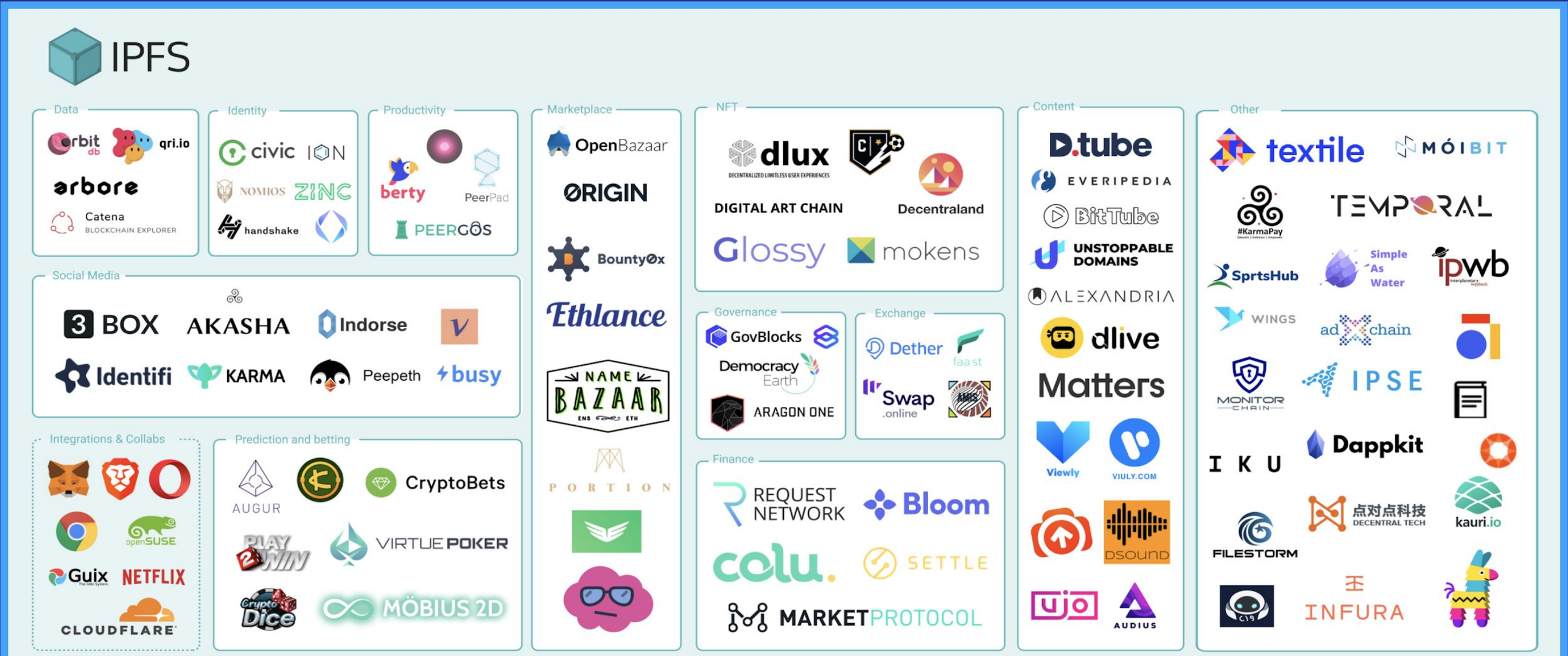
[1] Network Crawls with Nebula: *https://github.com/dennis-tra/nebula*

Principle I

Decentralized

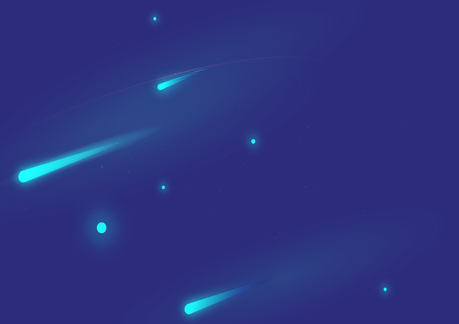
A part of the..

IPFS Ecosystem



Design Fundamentals

Content Addressing



Example of a CIDv1:

bafybeigdyrzt5sfp7udm7hu76uh7y26nf3efuylqabf3oclgtqy55fbzdi

⇓

<multibase>(cid-version || multicodec || multihash)

v1	dag-pb	sha2-256	32 bytes	SHA256 hash
00000001	01110000	00010010	00100000	110010010 ...
CID-Version	Multicodec	Multicodec	Length	Actual Hash
		Multihash		

Crawler Results

Cloud Provider Dependency

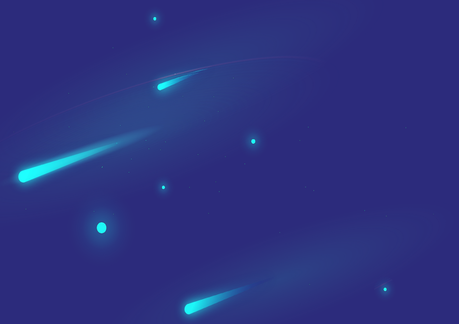
- Very small minority of nodes hosted on centralised cloud infrastructure!
  - At least on providers whose IP addresses are public.

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 %

Design Fundamentals

Peer Addressing



Metrics & Statistics

Agent Version Uptake