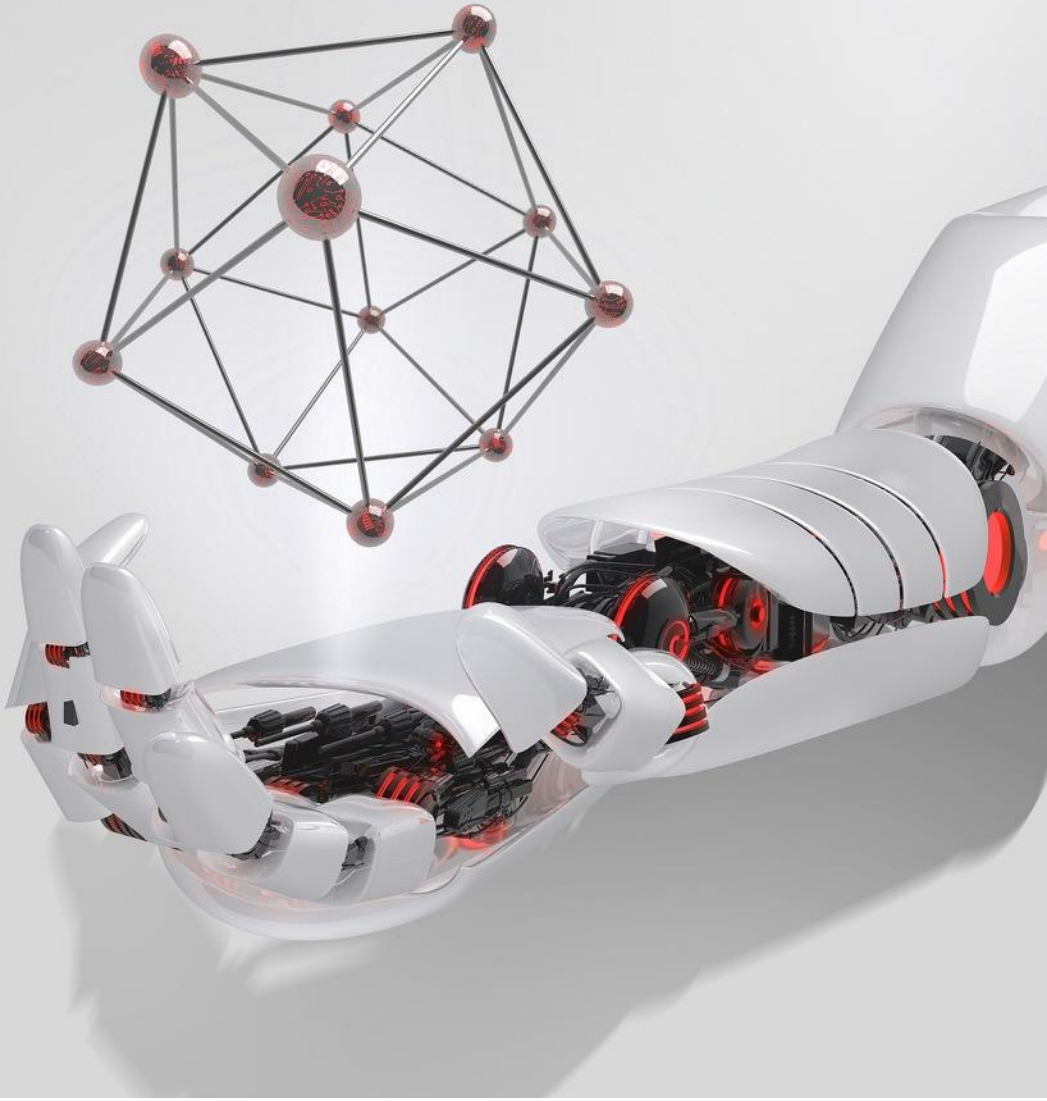
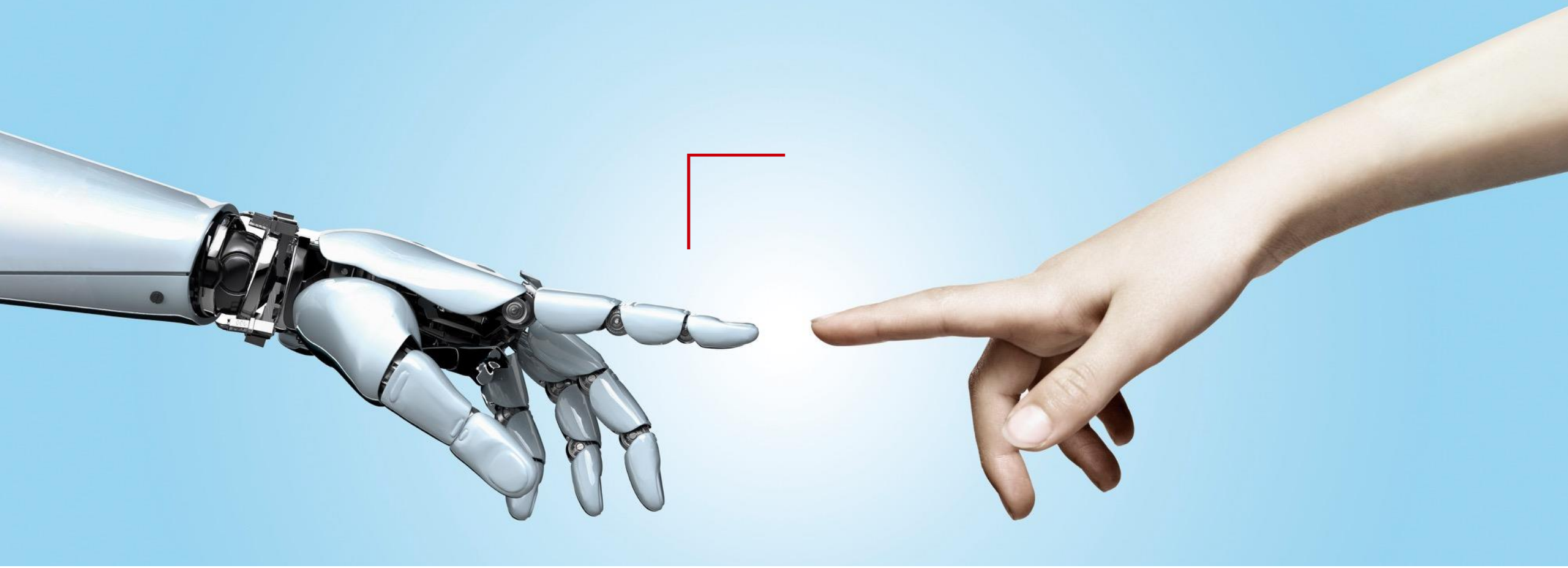
Insights on Impact of Distributed Ledgers in Provider Networks



David Guzman
September, 2022

Outline

Research Questions



Research Questions

Q1: What is the impact of DCSs on provider networks?

- **Outcomes: (a) insights on mechanisms & (b) quantifications**

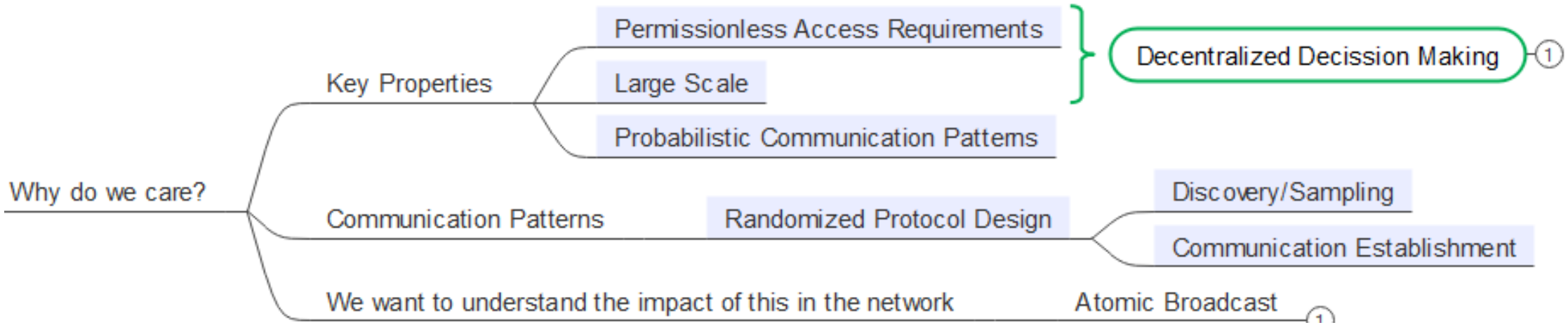
Q2: What network innovations could reduce the impact of the issues identified in Q1?

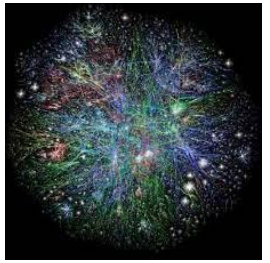
- **Outcomes: (a) design of network mechanisms & (b) quantification of improvements**

Q3: How could the presence of network innovations inform/guide better designs for DCSs?

- **Outcomes: (a) design of new DLT mechanism(s) & (b) quantification**

Why do we actually care?



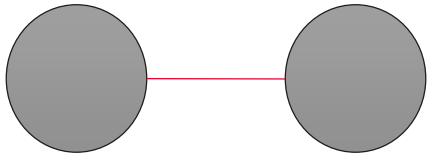


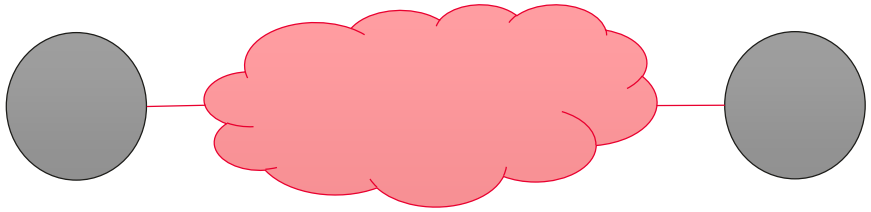
FLP Relaxation

Deterministic consensus in asynchronous networks -> **FLP Impossibility**

Randomized Algorithm: A consensus is always reachable in a **network**, but the time needed to reach consensus may be unbounded [1].

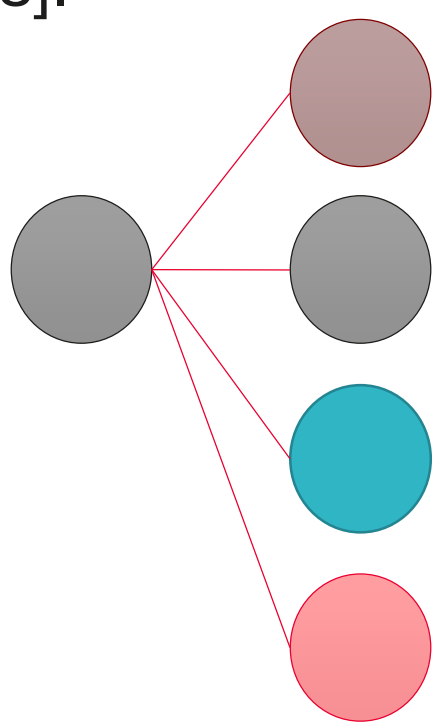
Disseminate information through the network

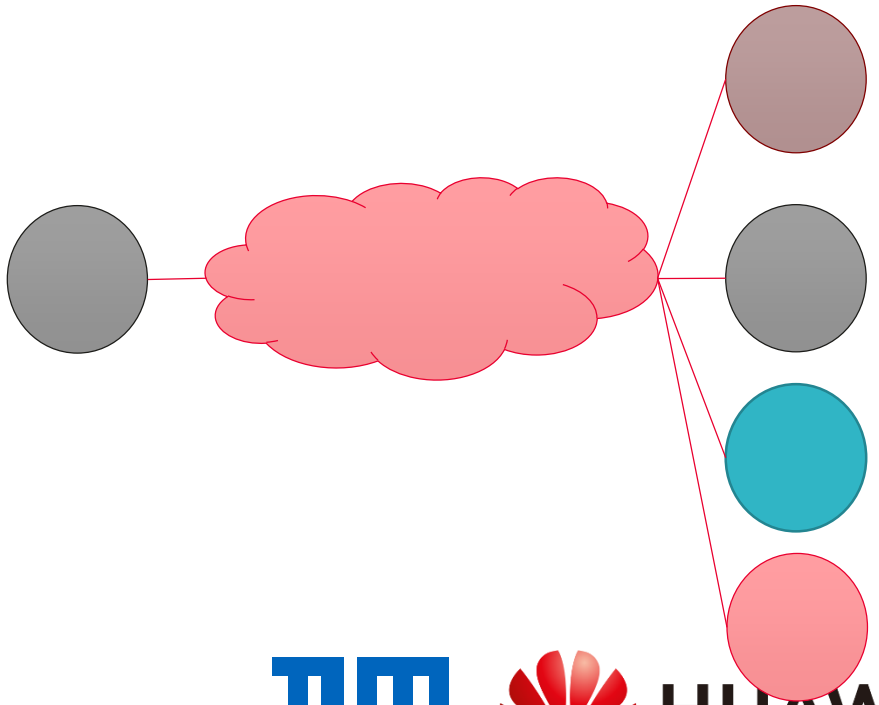




Randomized Distributed Consensus

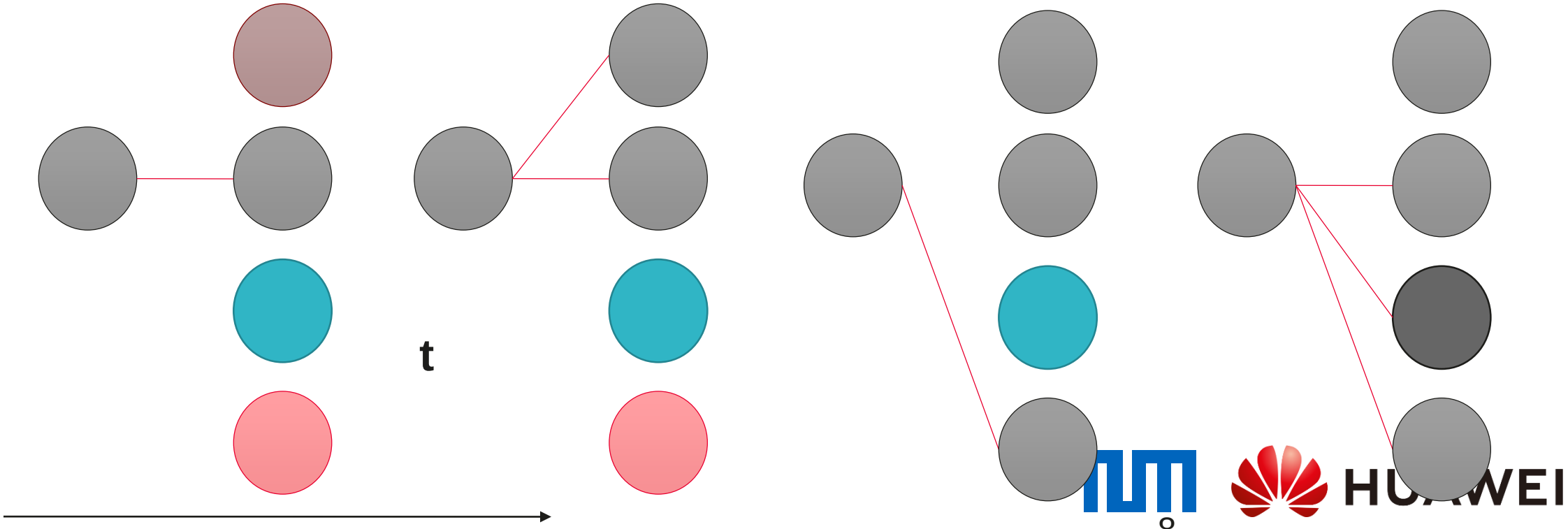
Deterministic atomic broadcast: is a broadcast which guarantees that all participants in a system eventually receive the same sequence of messages [3].





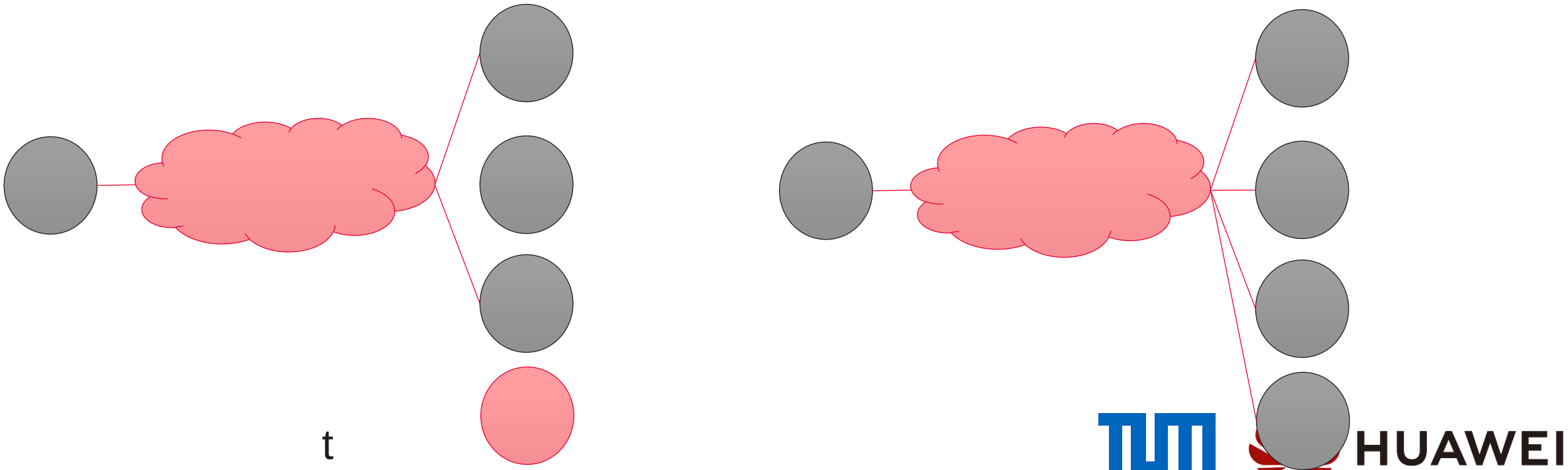
Randomized Distributed Consensus

Randomized atomic broadcast: is a broadcast performed **on random samples of the network** which guarantees that all participants in a system eventually receive the same sequence of messages. Diffusion mechanism.

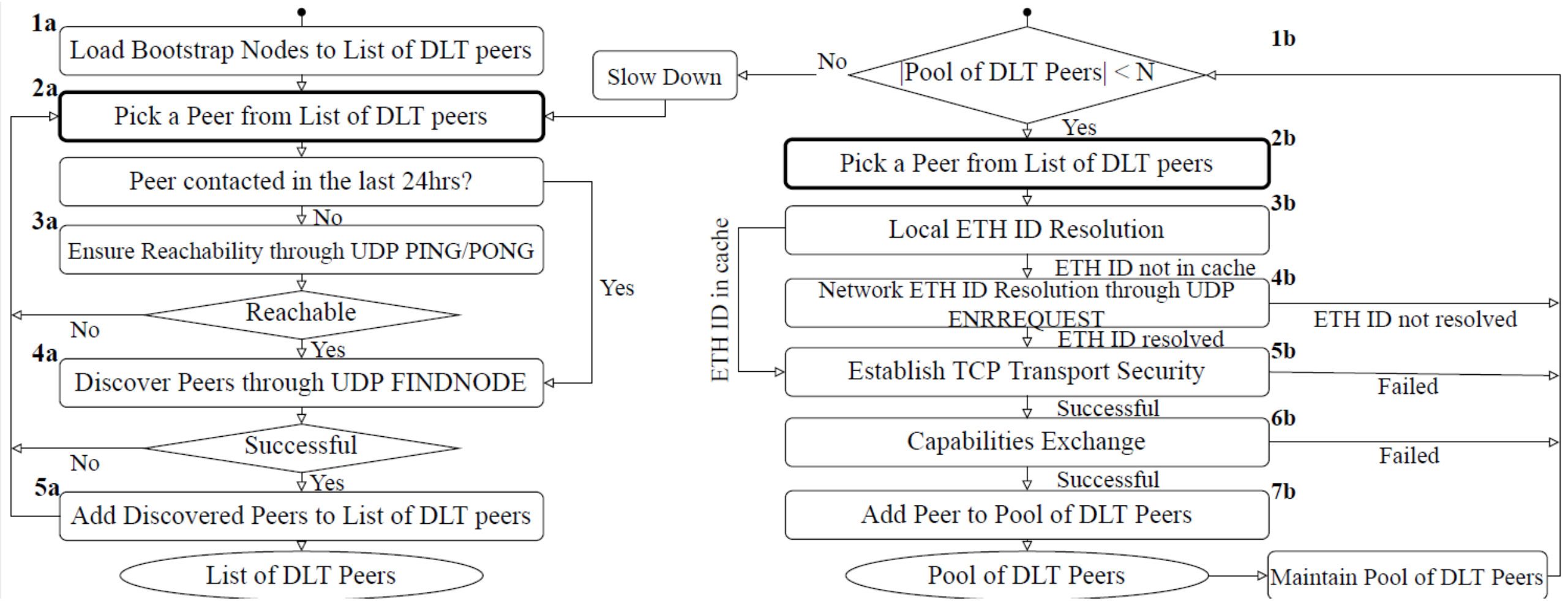


Atomic Broadcast: A Multipoint Randomized Unicast Communication Approach

Large scale overlay systems build on top of IP networks, UDP and TCP based.



Communication Patterns

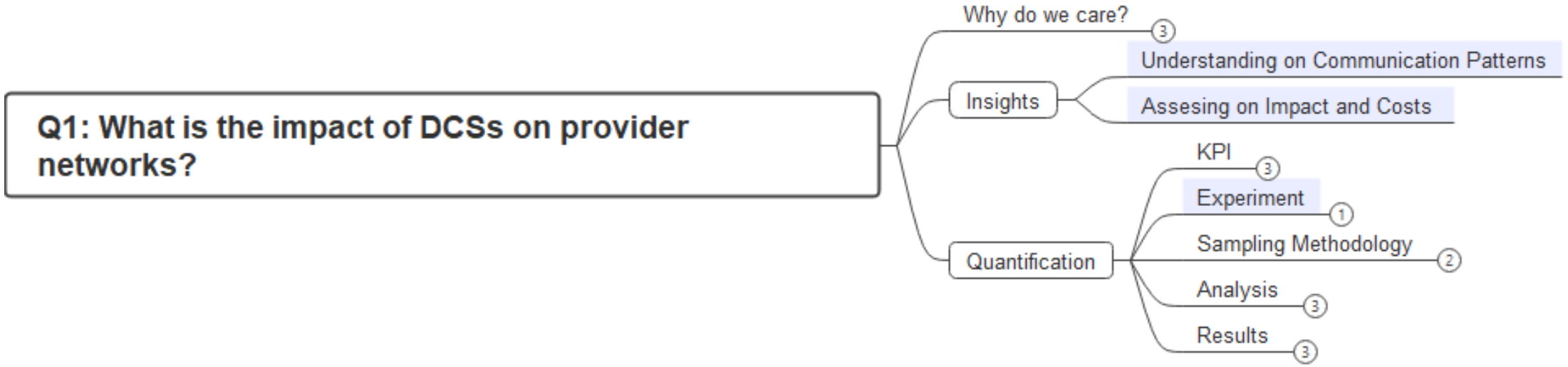


Discovery and Pool Establishment enabling randomized communication algorithms

Challenges for Users and Provider Networks

- Costs for pool maintenance:
  - Peers need to continuously establish and maintain reachability information
  - Each DLT peer maintaining a constantly changing pool (TCP)
- Costs for resilience and reliability:
  - Failing nodes causing latency on pool establishment (hence DC)
  - Timeouts inducing removal of peers, replenishing the pool
- Need to match capabilities:
  - Upper layer capabilities are required for pool.
- Unicast Replication
  - intention to achieve **diffusion** among many DLT peers instead of efficient network-level multicast.
- IP address privacy
  - DLT peers need to expose their IP address

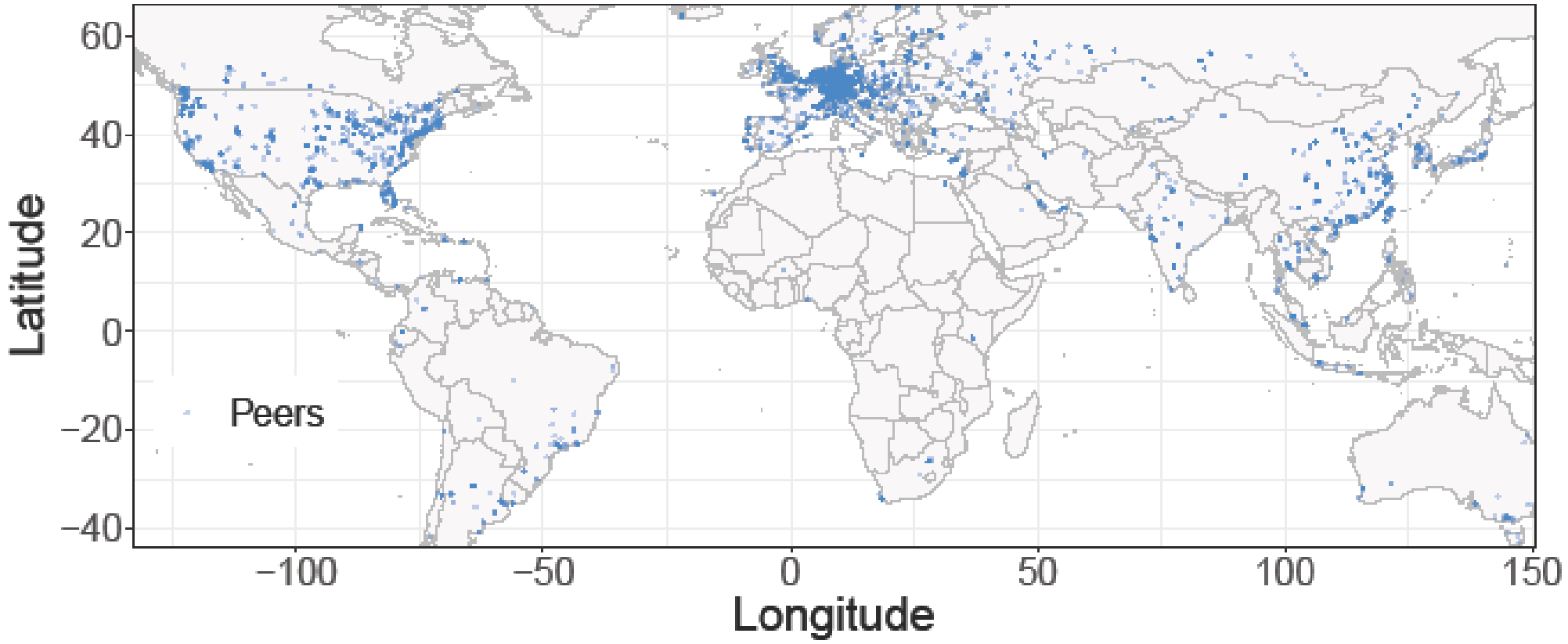
What did we do to answer Q1?



Experimental Insights

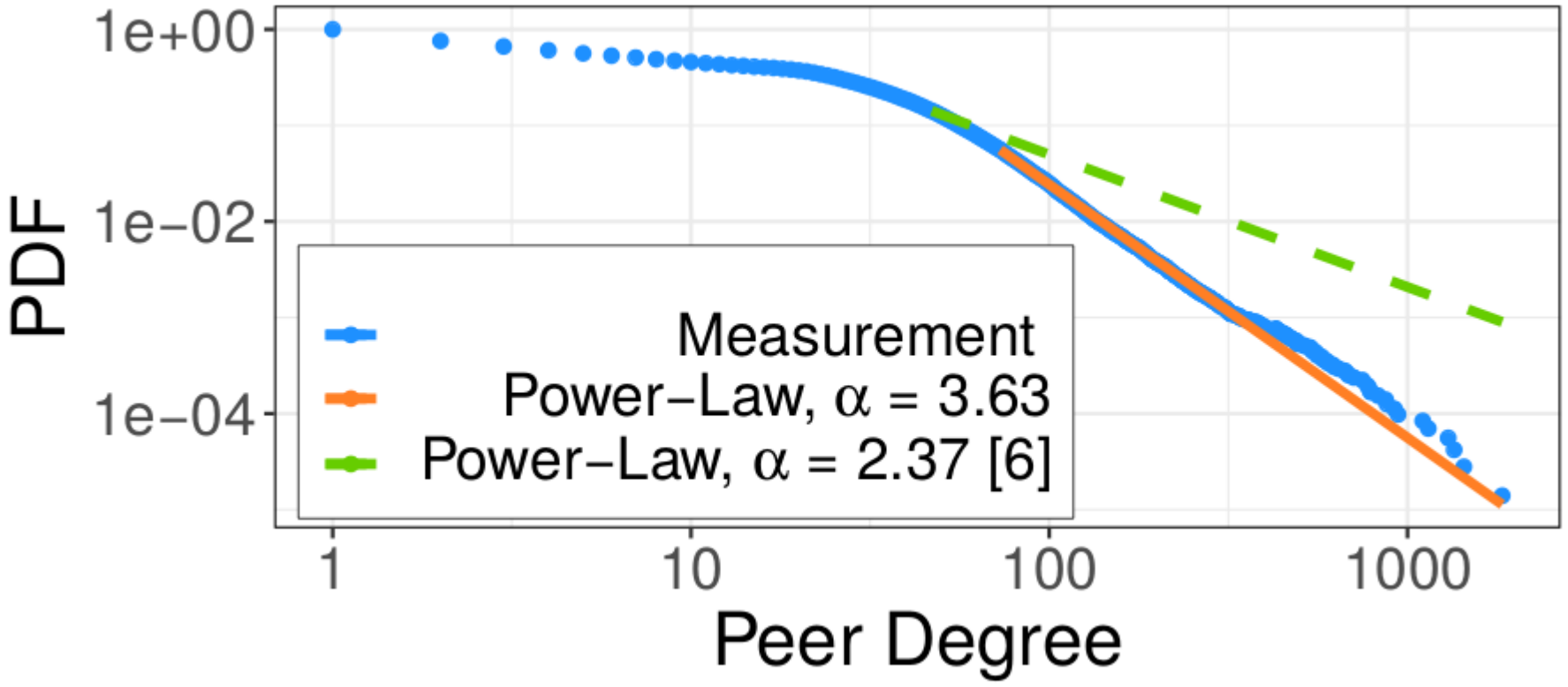
- Go-ethereum client syncronizing to the ethereum mainnet
- Geth/v1.10.2-stable
- Local peer on **: 217.110.131.84**
- UDP/TCP ports : 30303,30313
- Samples taken between : Jan-2022 – Jun-2022
- 8 bootstrap endpoints with v4UDP discover protocol
- Discovered active peers: 72k in comparison to [Gao 2018] 74k
- 100 iterations:
  - SYNC [5hrs] STOP[2mins] SYNC[5hrs] ……
  - Discovery / Pool Establishment [40min] ……..SYNC[4hrs20min]
- Downloaded data : ~280GB

Experimental Insights Topological Meas.



- Clusters

Experimental Insights Topological Meas.



- X_min=71
- Xmax=1844

Experimental Insights
Topological Meas.

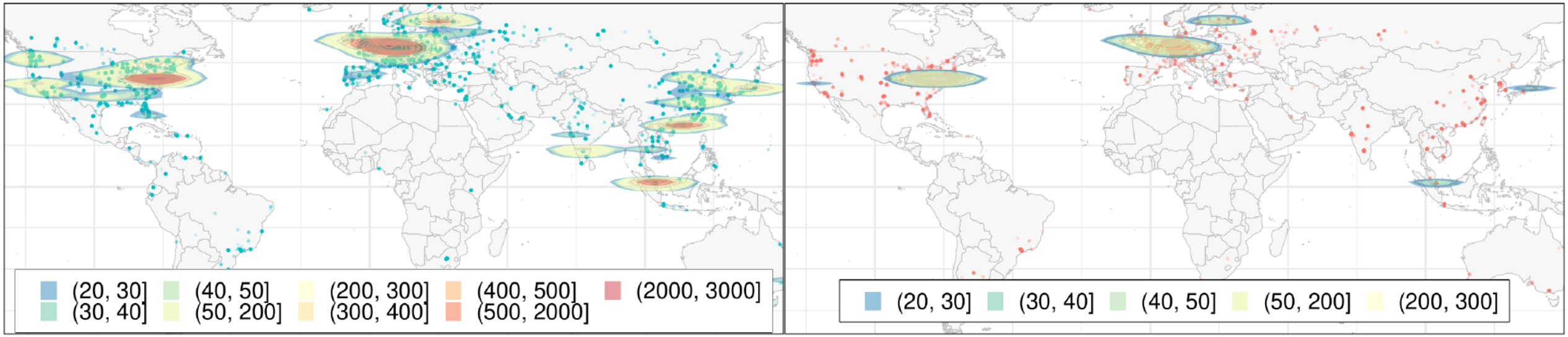


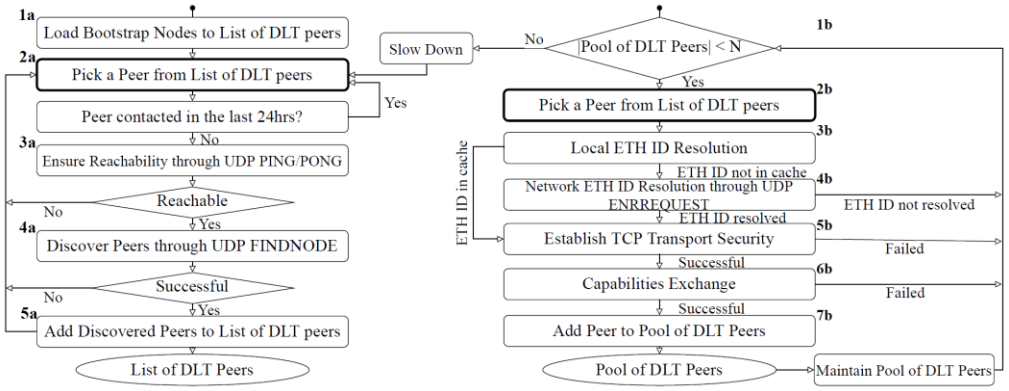
Fig. 3. (a) Incoming Connection Reachability, (b) Outgoing Connection Reachability

- Clustered
- Incoming communication relations 10x

Type of DLT Peers

Table 1. Type of Peers

Peer	Outgoing [%]	Avg.	Incoming [%]	Avg.
Non-Reachable	5.17	353.22	0.01	1.87
Signaling	87.1	6011.29	76.78	10259.56
Potential Data	7.73	540.7	23.19	3098.35
Data	0.002	0.111	0.013	1.667
Dropped Data	0.005	0.302	0.005	0.302



Pool Establishment Time

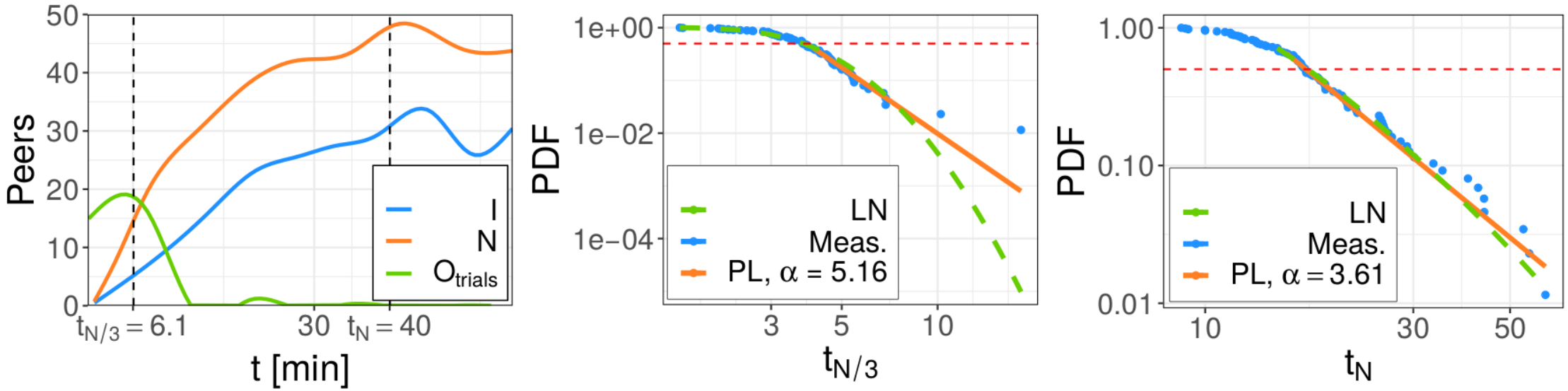


Fig. 4. (a) Single Sample for Pool Establishment Time, (b) $t_{N/3}$, (c) t_N

- tN/3, is reached at 4min with 50%probability, while the time to complete a pool of DLT peers is reached at 20min

Pool Establishment Cost - Discovery

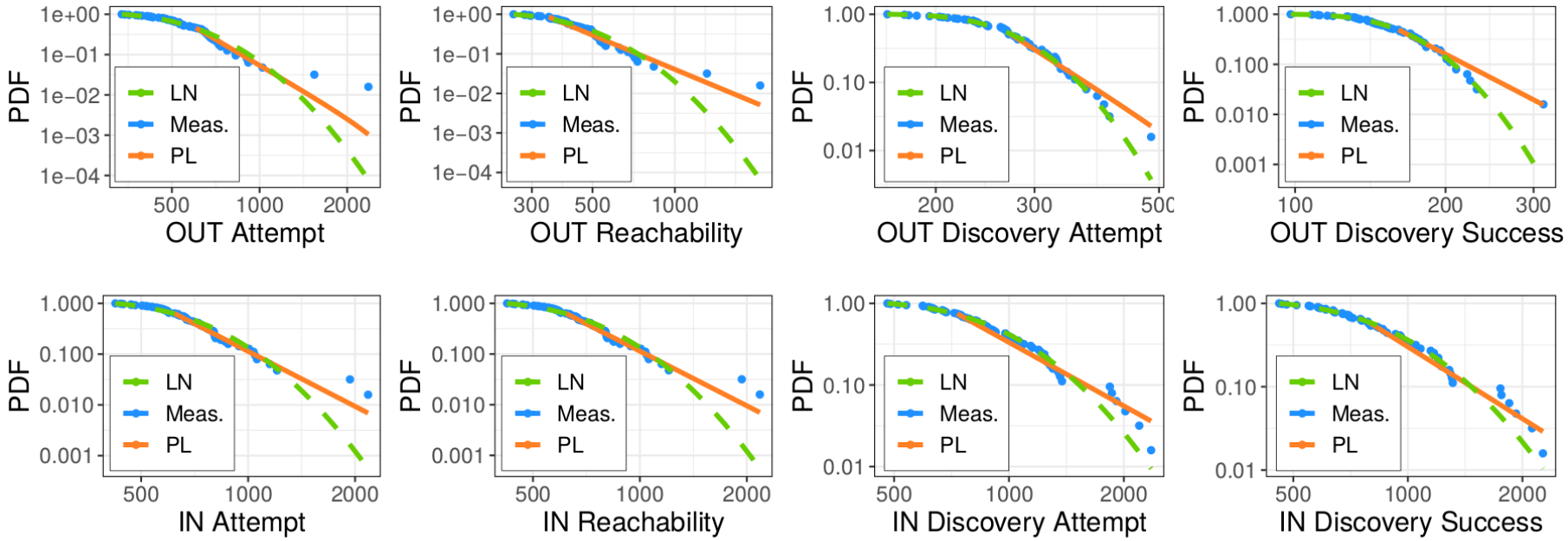


Fig. 5. (a-d) Outgoing, and (e-h) Incoming Discovery KPI Distributions

	Outgoing						Incoming					
	Power-Law			Log-Normal			Power-Law			Log-Normal		
RV	α	x_{min}	p	μ	σ^2	x_{min}	α	x_{min}	p	μ	σ^2	x_{min}
Attempts	5.24	604	0.2	6.31	0.38	336	4.6	620	0.58	6.47	0.37	423
Reachability	3.87	346	**0.03**	6.04	0.41	257	4.58	619	0.59	6.47	0.37	423
Disc. Attempt	6.22	269	0.79	5.61	0.21	164	3.59	732	0.16	6.79	0.41	480
Disc. Success	6.16	161	**0.04**	5.07	0.21	98	3.86	785	0.11	6.72	0.43	459

HUAWEI

Pool Establishment Cost - Discovery

≈77% of the contacted peers are reachable, out of these peers≈63% are used to further topology discovery, and≈57% successfully executed a complete discovery protocol.
almost all the reachability requests were positively replied by our peer, and≈94% of the discovery requests are successfully completed.

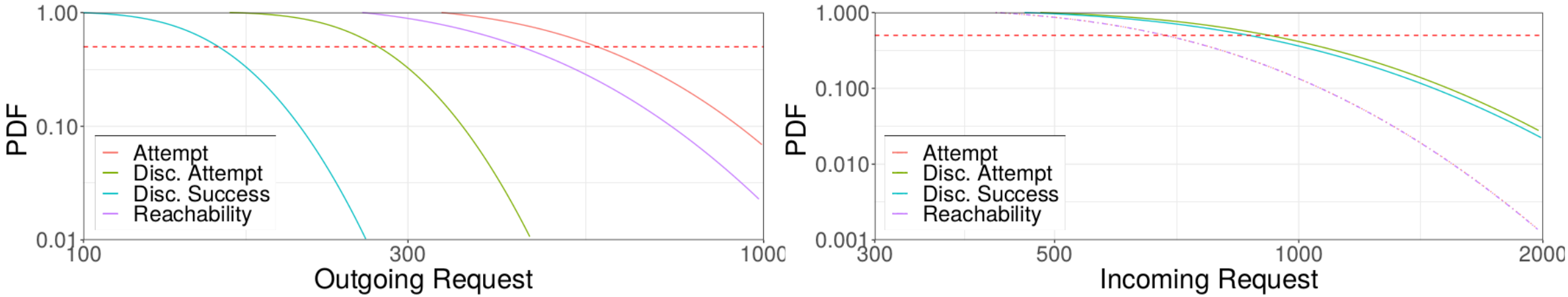


Fig. 6. (a) Outgoing, and (b) Incoming Discovery Cost

Pool Establishment Cost

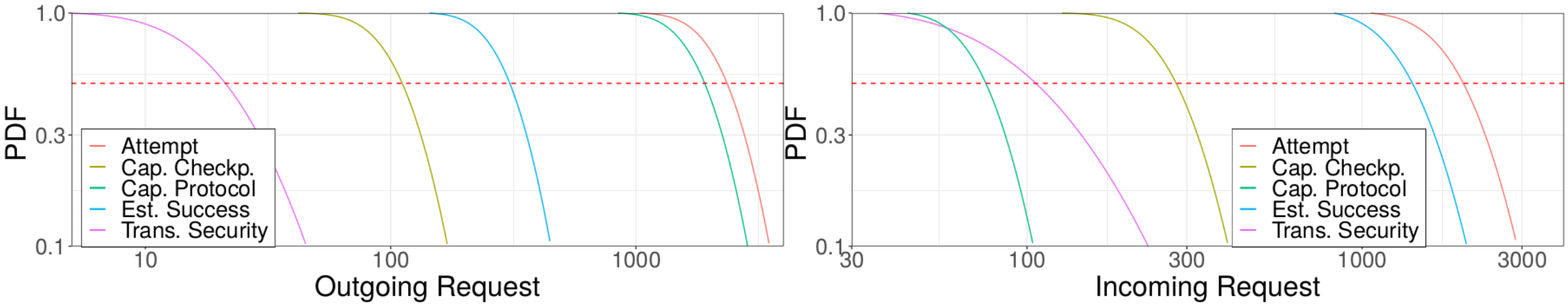


Fig. 8. (a) Outgoing, and (b) Incoming Pool Establishment Cost

- outgoing peers have significantly higher transport and security failures, where 82% of errors occur while trying to decrypt the remote secret and find the proper blockchain checkpoint, and18% I/O errors (invalid ciphertext length, unexpected end of file); we currently explain this issue with stored but outdated cypher information

Impact on Provider Networks: Effective Data Consumption

- Downloaded data amounted 3168G, out of which 892G were useful data added to the local blockchain, while 2276G were dropped.
- We interpret this as an effective data consumption ratio of 28.15%.

Further Work

Further Research Question

Q1: What is the impact of DLT solutions on provider networks?

- **Outcomes: (a) insights on mechanisms & (b) quantifications**

Q2: What network innovations could reduce the impact of the issues identified in Q1?

- **Outcomes: (a) design of network mechanisms & (b) quantification of improvements**

Q3: How could the presence of network innovations inform/guide better designs for DLTs?

- **Outcomes: (a) design of new DLT mechanism(s) & (b) quantification**

Q2: What network innovations could reduce the impact of the issues identified in Q1?

- **Problem to solve:** Permissionless DCS fundamentally based on utilizing an instantaneously randomized broadcast to a fixed size group among a subset of miners that fit particular constraints
- **Observations**
  1. Miners provide a service capability to other miners and clients in the DLT
  2. Pool creation and maintenance (done at EVERY peer) is core mechanism to enable instantaneously randomized operations
  3. Constraints (as investigated in ETH) include reachability, TLS capabilities, certain HW, checkpoint, …
  4. Fixed group size is defined through heuristics (theoretical bounds) on the probability for converging to consensus among those group members
  5. A group of members is instantaneously randomized to ensure protection against collusion
- **From observations/insights to a proposed design**
  1. Use service-centric abstraction (miners are service instances to DLT service)
  2. Use (service) routes to (pool of) service instances as key concept to enable instantaneously randomized operations
  3. Replace pool maintenance by encoding constraints that ensure successful communication as naming structure
  4. Provide a forward multicast capability to a fixed size subset of constrained named service instances
  5. Ensure that fixed size is randomized with every request (which excludes the use of IP multicast)
- **Reducing impact on network (compared to ETH over IP)**
  - From requiring pool maintenance to using service route announcements
  - From waste in pool maintenance due to, e.g., lack of reachability, mismatching capabilities, to service route convergence
  - -> Baseline here will be our insights (in terms of convergence latency, pool latency, pool maintenance cost) in ETH over IP

Questions

Thank you.

Bring digital to every person, home and organization for a fully connected, intelligent world.

Pool Establishment Cost

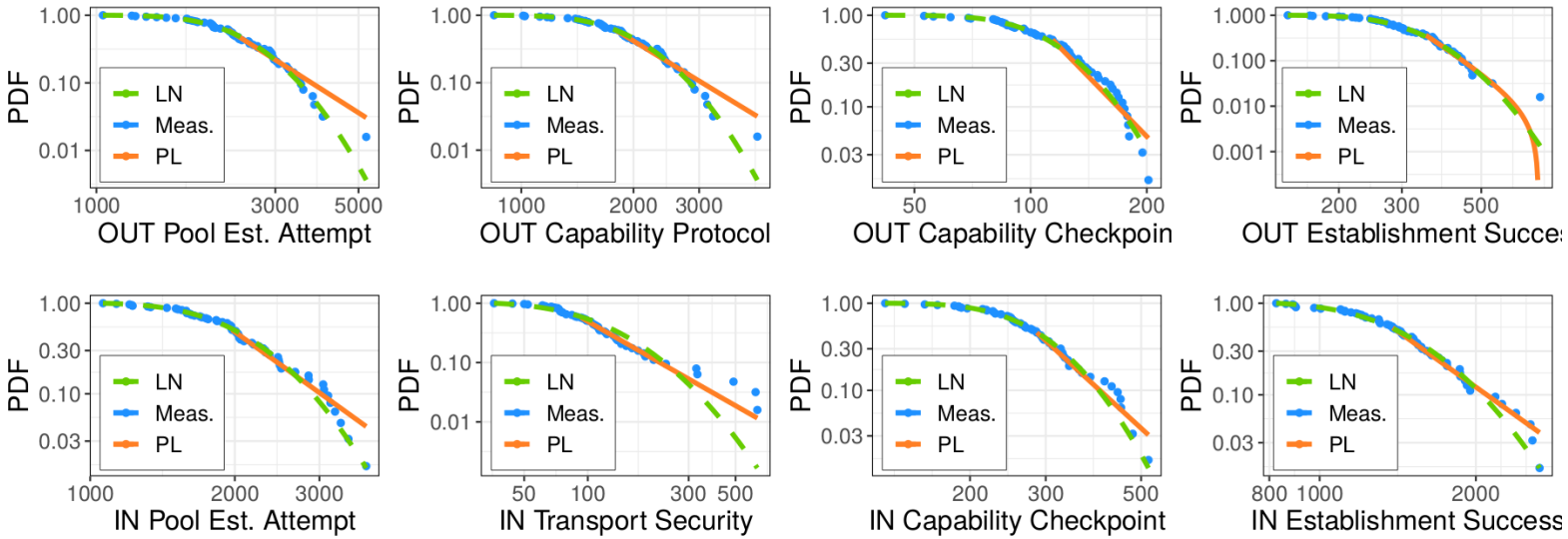


Fig. 7. (a-d) Outgoing, and (e-h) Incoming Pool Establishment KPI Distributions

	Outgoing						Incoming					
	Power-Law			Log-Normal			Power-Law			Log-Normal		
RV	α	x_{min}	p	μ	σ^2	x_{min}	α	x_{min}	p	μ	σ^2	x_{min}
Attempts	4.52	2250	0.33	7.78	0.29	1041	4.81	1973	0.43	7.6	0.29	1063
Transport Security	3.6	23	0.69	3.09	0.56	5	3.03	93	0.62	4.63	0.62	36
Capability Protocol	4.37	1777	0.25	7.57	0.3	845	6.71	83	0.86	4.32	0.26	44
Capability Checkpoint	5.31	114	0.25	4.73	0.31	42	5.52	286	0.46	5.63	0.28	127
Establishment Success	6.21	350	0.26	5.74	0.28	144	5.04	1383	0.75	7.24	0.3	825

HUAWEI

A Primer on Distributed Consensus

Consensus: several processes/peers need to agree on a single value

Some processes may be malicious: <=f out of n=3f+1

State-machine replication (SMR): processes/peers agree on a sequence of values – commands to change the replicated state

Blockchain (a.k.a DLTs): using Byzantine SMR to agree on a sequence of blocks in a ledger. [2]

Byzantine Consensus

Byzantine consensus: hard decision finality, but **permissioned** system – fixed set of participants [2].

Blockchain consensus (PoW): no hard (deterministic) finality guarantees, but **permissionless** system – anyone can participate [2].

Using Byzantine consensus for blockchain: elect a committee to finalize decisions via Byzantine consensus.

Byzantine Consensus

At its heart: network (understanding its features, characteristics)

Synchronous network: there is a known fixed upper bound **D** on the time required for a message to be sent from one processor/peer to another, and a known fixed **f** upper bound on the relative speeds of different processors [1].

Asynchronous network: no fixed upper bounds **D** and **f** exist [1].

Partial synchrony: fixed bounds **D** and **f** exist, but they are not known a priori [1].

Byzantine Consensus - Fisher Lynch Paterson (FLP) Impossibility

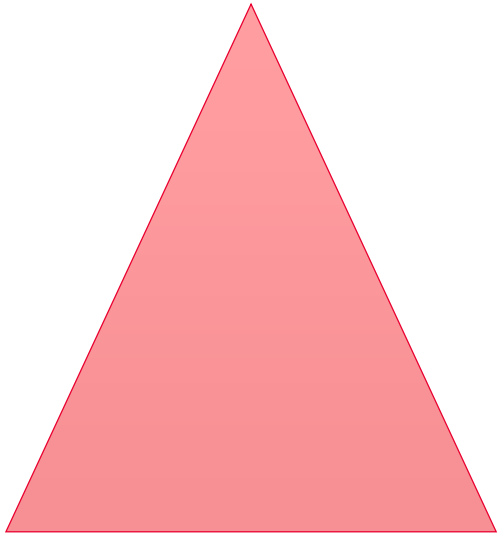
Fisher Lynch Paterson (FLP) Impossibility: a consensus protocol that works in an asynchronous model also works in a synchronous model.

A synchronous model has modifications and restrictions on an asynchronous model so that the synchronous model is closer to the real scenarios and it is **possible to solve the consensus problem** in practice.

FLP indicates that consensus is not always **reachable in bounded** time in asynchronous networks. [4]

Byzantine Consensus - Pick 2 out of 3

Fault tolerance requires that a protocol must also effective in case of node failures.



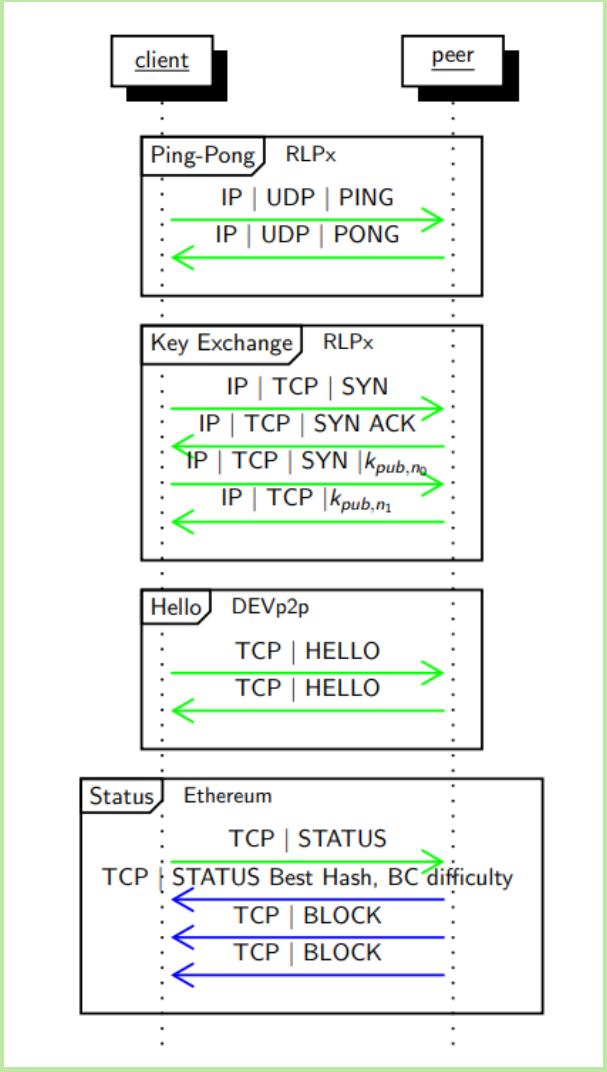
Agreement (Safety) means that the values reached across nodes in a system are consistent and valid

Termination (Liveness) indicates that individual nodes in a system must reach an agreement (in bounded time), that is, the system must move forward and cannot always be in the inconsistency state.

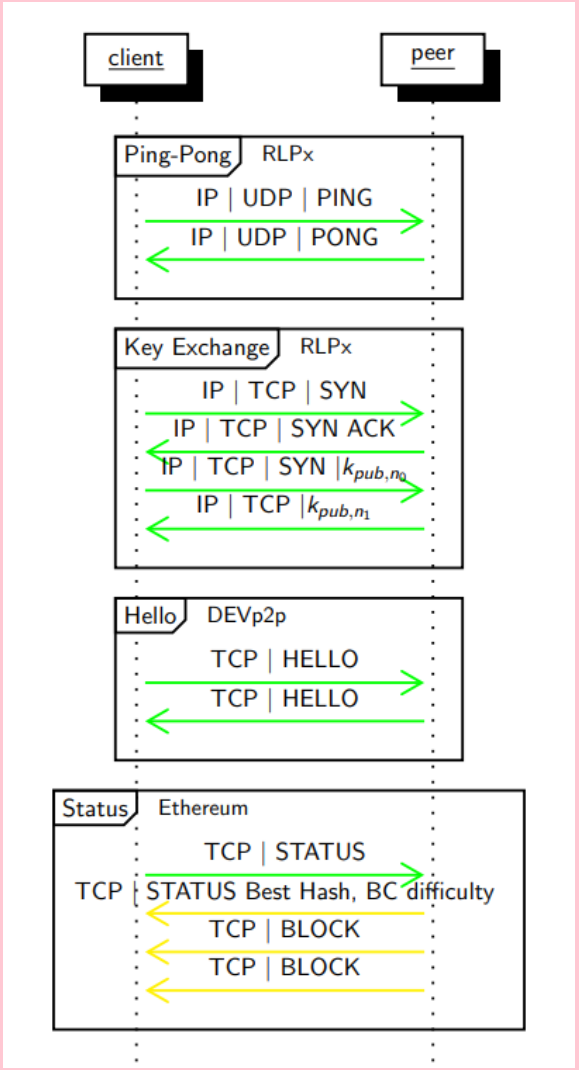
Example: we can sacrifice a certain degree of safety, which means that the system can always reach an agreement quickly but the agreement is not very reliable.

Types of Endpoints in a Deployed DLT

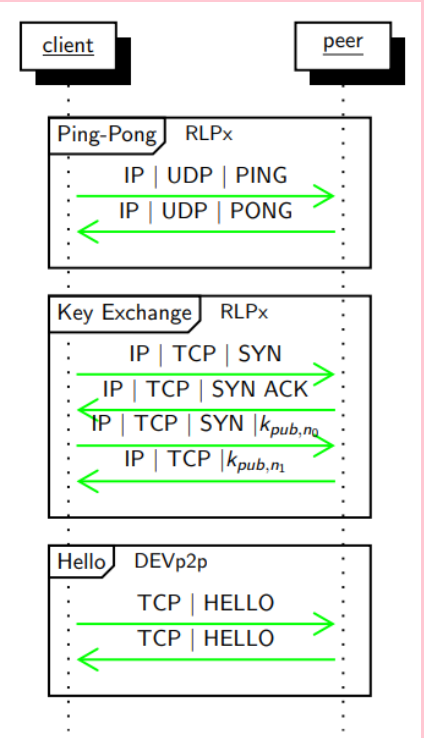
Successful Transaction



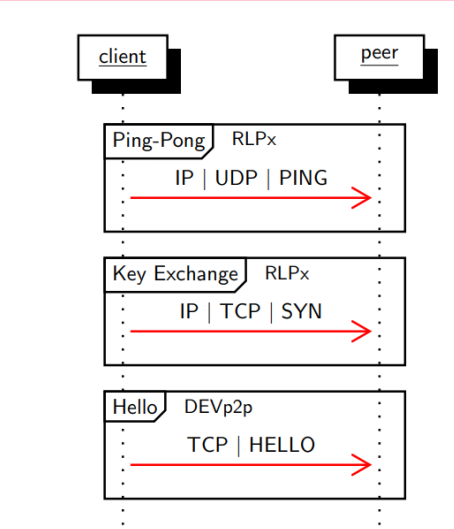
Failed Transaction



Just signalling



Non routable



- [1] Consensus in the Presence of Partial Synchrony. Dwork, Lynch and Stockmeyer. ACM on Distributed Computing. 1984
- [2] Liveness in PBFT. Gotsman Alexander. 2022
- [3] Distributed consensus revised. Howard Heidi. Cambridge. 2019
- [4] From Distributed Consensus Algorithms to the Blockchain Consensus Mechanism. 2019
- [5] A Survey and Comparison of Peer-to-Peer Overlay Network Schemes. Eng Keong Lua, Jon Crowcroft, Marcelo Pias, Ravi Sharma and Steven Lim. IEEE Survey. 2004