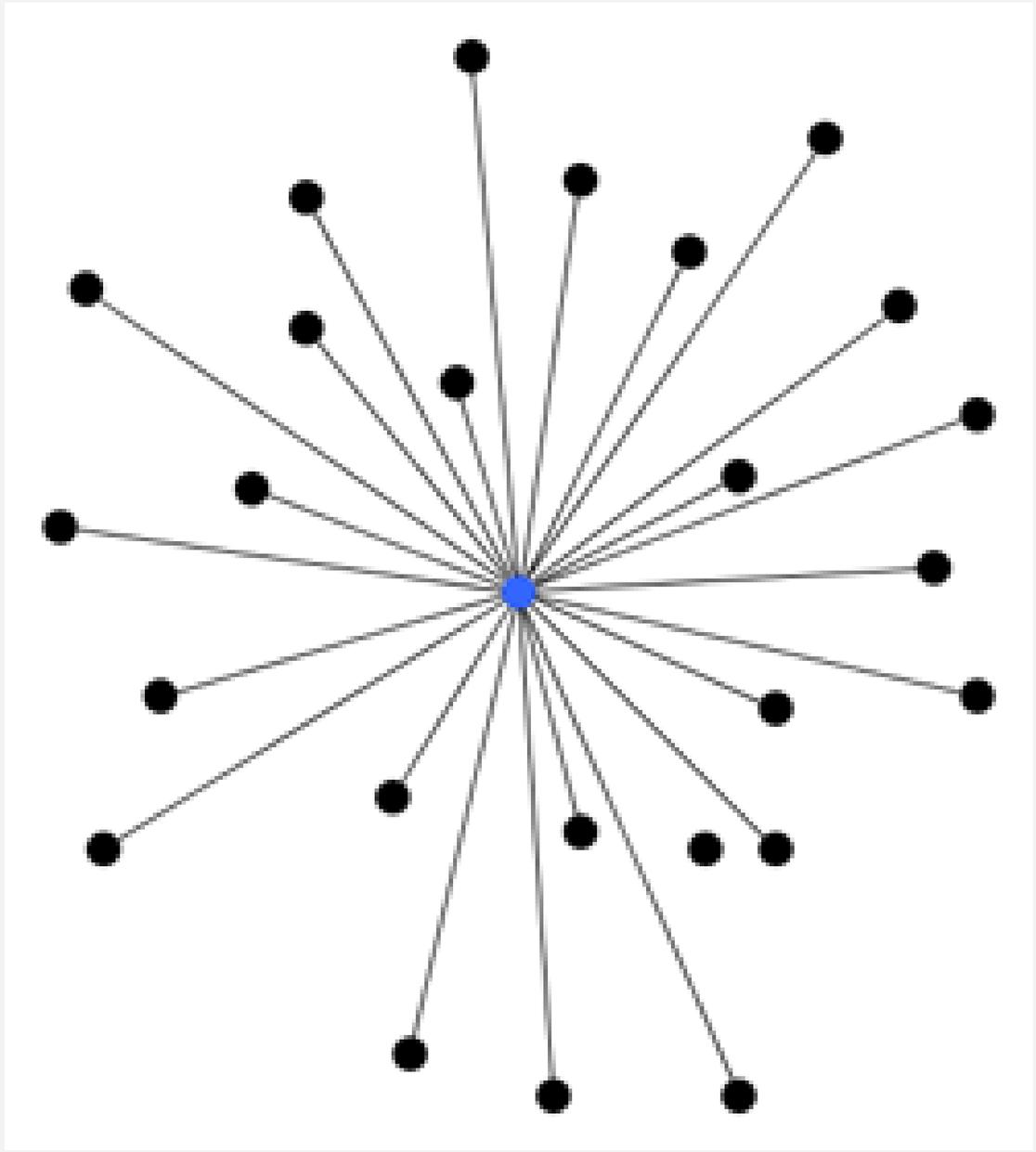
Swarm

A distributed storage by ethereum.

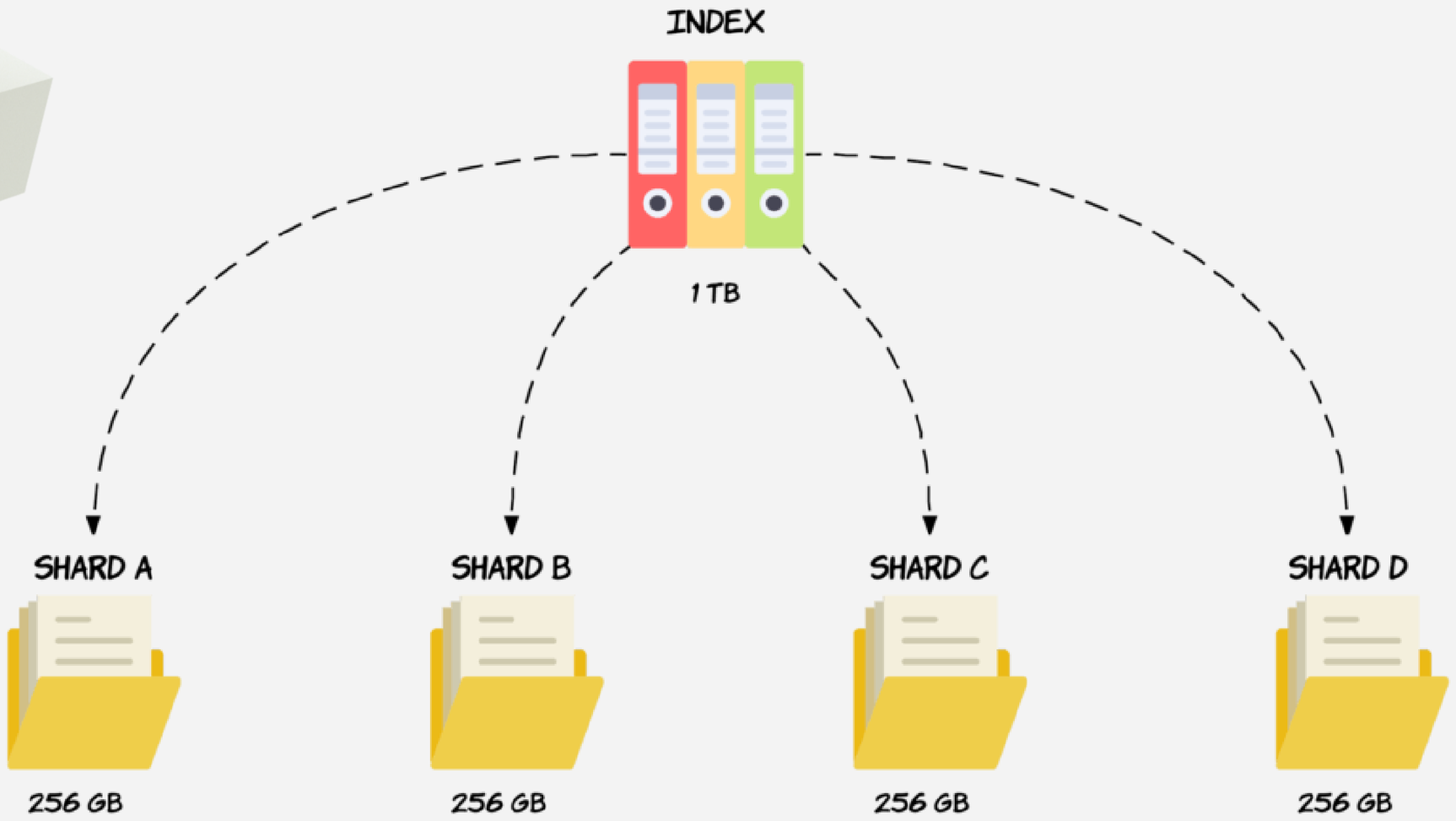
by Abhishek Kedia

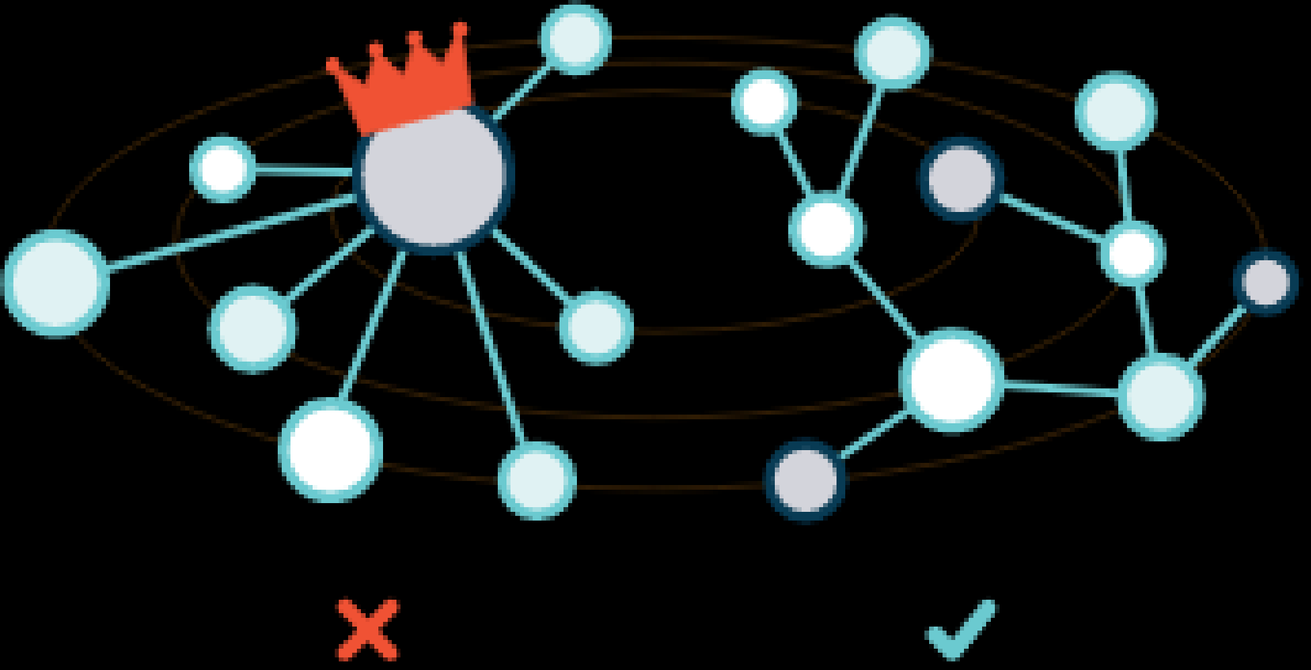
What is distributed storage?

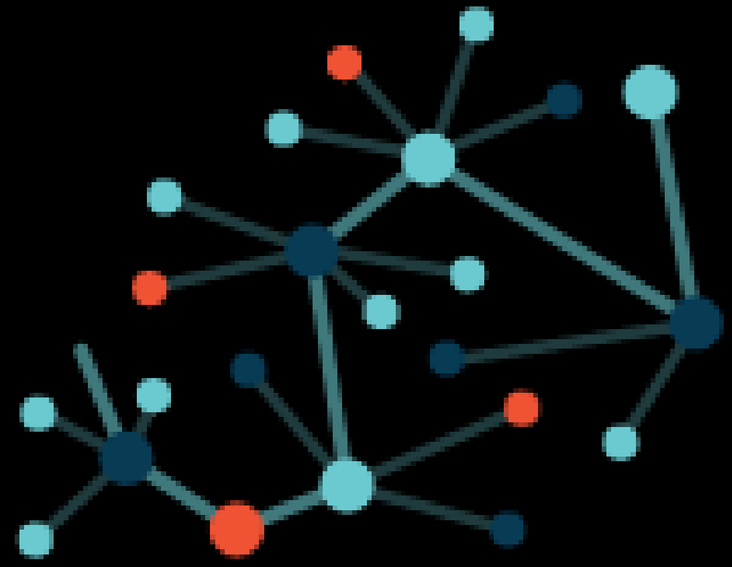
Centralised Storages



SHARDING

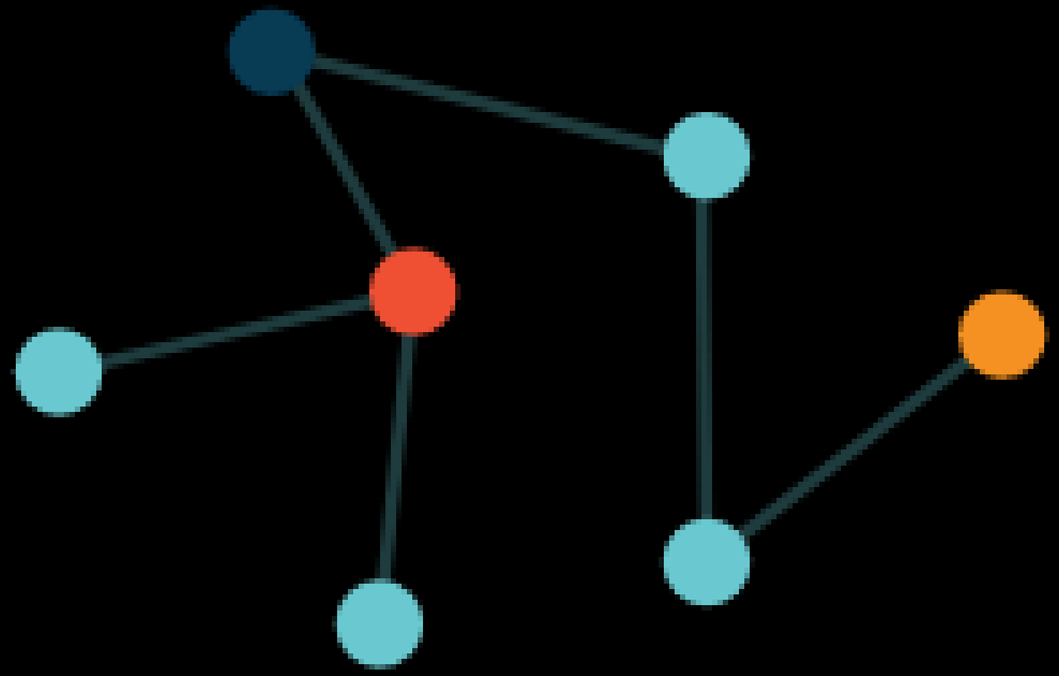

INDEX
1 TB
SHARD A
256 GB
SHARD B
256 GB
SHARD C
256 GB
SHARD D
256 GB





The Rise of distributed storages

Swarm is a censorship resistant, permissionless, distributed storage and communication infrastructure.



How is it different from IPFS or bittorrent?

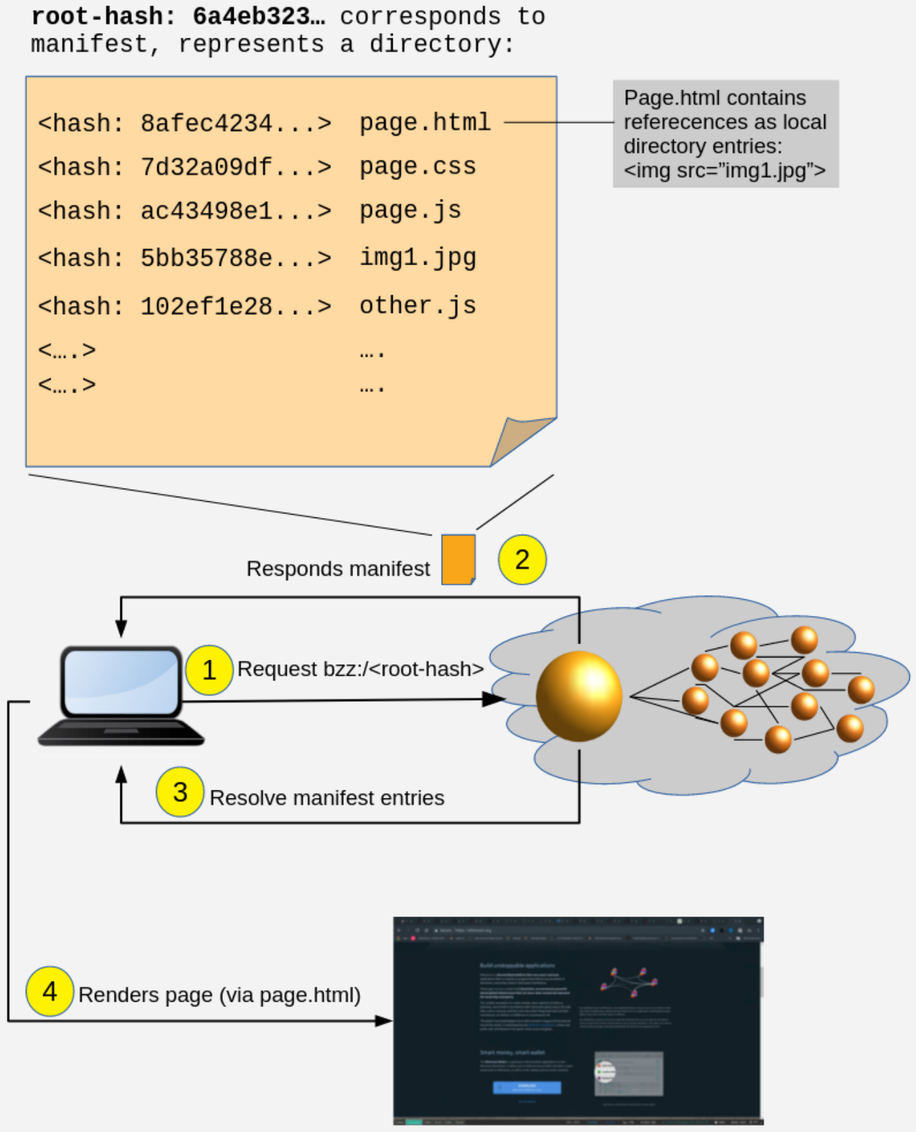
How does Swarm work?

Swarm defines 3 crucial notions:

Chunks
Chunks are pieces of data of limited size (max 4K), the basic unit of storage and retrieval in the Swarm. The network layer only knows about chunks and has no notion of file or collection.

Reference
A reference is a unique identifier of a file that allows clients to retrieve and access the content.

Manifest
A manifest is a data structure describing file collections; they specify paths and corresponding content hashes allowing for URL based content retrieval.



The nodes constituting the Swarm all dedicate resources (diskspace, memory, bandwidth and CPU) to store and serve chunks.

Who stores the data?

The process by which chunks get to their address is called syncing and is part of the protocol. Nodes that later want to retrieve the content can find it again by forwarding a query to nodes that are close the the content's address. Indeed, when a node needs a chunk, it simply posts a request to the Swarm with the address of the content, and the Swarm will forward the requests until the data is found.
Swarm is similar to a traditional distributed hash table (DHT).

Let's Try it.