IPFS and the Decentralized Web:
The key to the future of Libraries and Archives in a data-centric world

Matt Zumwalt (@flyingzumwalt) and Richard Littauer (@richlitt)

Protocol Labs

The Future of Libraries

Libraries and Archives will be able to use IPFS to store and serve content using bittorrent-like infrastructure, on top of a file system built around better protocols for linking data. This our imagining of the future:

- **Libraries pin datasets their patrons care about,** serving them to patrons and making sure they are backed up.
- University libraries **seamlessly archive complete histories of research datasets** by simply allowing their stakeholders to submit the IPNS hashes of datasets that should be archived.
- **LOCKSS and DPN use cases become dead simple** to implement and support, and the barrier to entry becomes negligible and overhead for maintenance of the network is radically reduced.
- The **overhead cost of hosting content plummets** due to the Bittorrent-style model of delivery.
- Librarians get to resume their role as facilitators and curators who **engage directly with the community** to help them preserve, discover and access content.

Other Uses

P2P CACHING OF SHARED/COMMON LINKED DATA
IPLD solves some of the core preservation problems that occur with Linked Data (example: Library of Congress Subject Headings - LCSH and all of id.loc.gov) by using content-addressed authenticated data structures instead of semantic web and location-addressed http URIs.

GRACEFUL GEOGRAPHIC REDUNDANCY
IPFS covers the LOCKSS/DPN use cases gracefully with an open, stable, secure protocol that's used across many industries.

BEYOND WEB APPLICATIONS and WEB APIs:
It's time to stop thinking in terms of Web Applications and REST APIs. Instead think of authenticated/functional data structures, p2p replication, and dynamic content on distributed systems.

How IPFS Works

IPFS stores data in a Merkle DAG, a data format used by BitTorrent and Git.

It stores data in blocks, which are then referred to by the cryptographic hash of that data - a reproducible, unique key. These hashes, in turn, link to each other, providing an easy way of indexing content which is fully reproducible, fast, immutable.

MERKLE LINK
/ipfs/QmW98pJrc6FZ6/foo/bar/baz.png
content   path
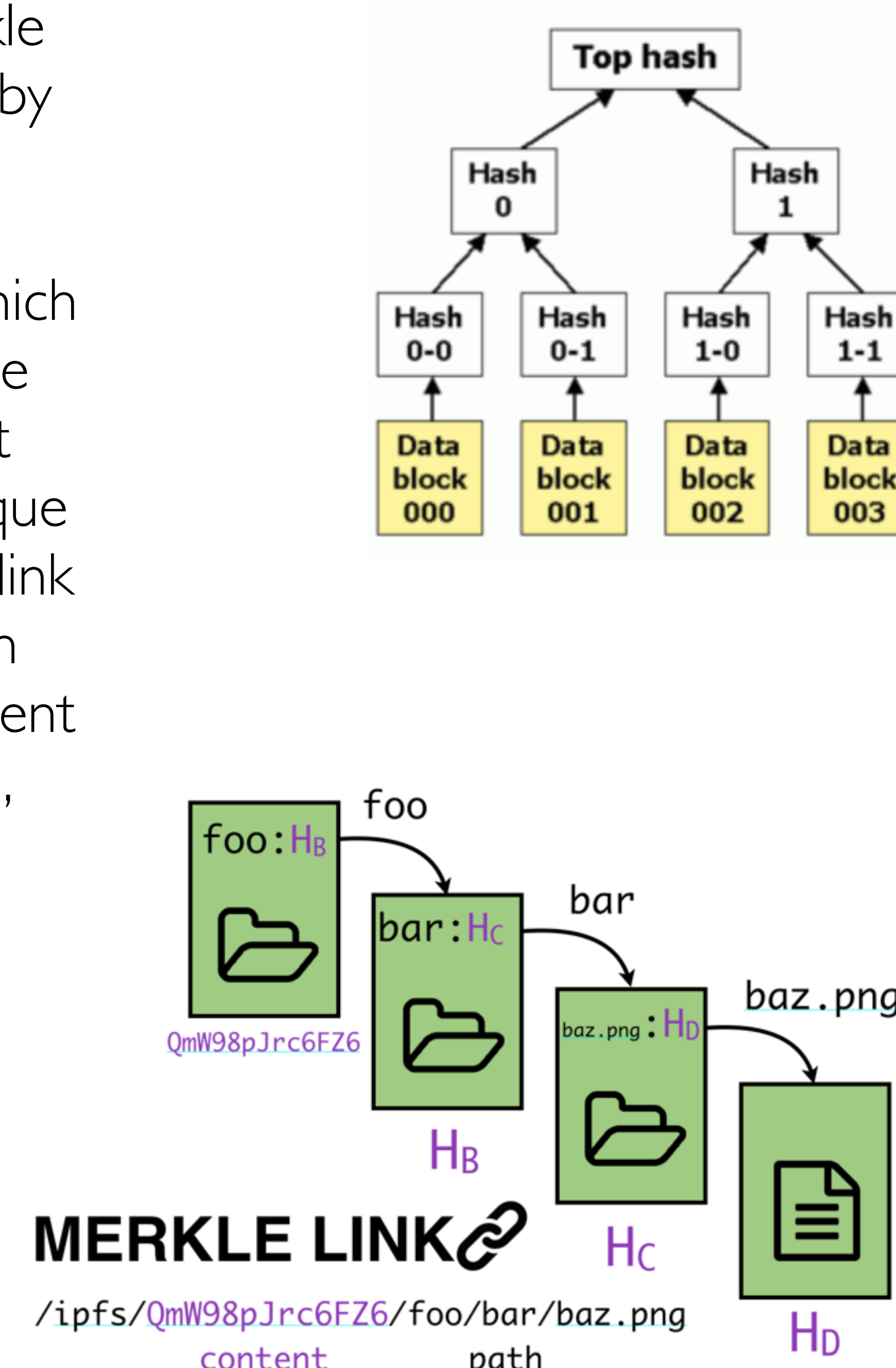
How IPLD Works

InterPlanetary Linked Data is a data format that allows you to express linked data as Merkle DAGs. It can be persisted in IPFS.

Instead of relying on HTTP URIs to resolve links in the graph, use cryptographic hashes. This makes the entire graph an authenticated, immutable dataset that:

- Natively accommodates signing assertions (attribution/provenance of metadata).
- Inherently supports replication and redundancy.
- Scales like BitTorrent.
- Inherently supports compact, efficient versioning.

Reading

- ipfs.io
- ipld.io
- GitHub: github.com/ipfs/ipfs
- Decentralized Web Primer: https://flyingzumwalt.gitbooks.io/decentralized-web-primer/content/
- Decentralized Data Workshop: https://codeforphilly.github.io/decentralized-data/
- @IPFSBOT