廣東工業大學
Guangdong University of Technology

星际文件系统IPFS 技术架构

刘文印 liuwy@gdut.edu.cn

广东工业大学网络身份安全粤港联合实验室 WIS Lab

个人微信：csliuwy；**微信公众号**：wislab; denglu-1或“**登录易**”

提纲

- HTTP的限制
- 分布式文件系统的发展
- IPFS技术架构
  - IPLD
  - Bitswap
  - Libp2p
- Filecoin

HTTP的限制

- HTTP是脆弱的
- HTTP趋于超中心化
- HTTP的效率不高
- HTTP过度依赖主干网络

分布式文件系统的发展

- 分布式服务器主机和CDN

    1. 将文件分散到多个服务器存储
    2. 在网络各处放置节点服务器构成现有的互联网基础之上的一层智能虚拟网络

- AFS和BitTorrent

    1. 国外以AFS为代表的分布式文件共享系统
    2. 国内以迅雷为代表的P2P下载系统

IPFS技术结构

IPFS的系统架构图， 分为5层：

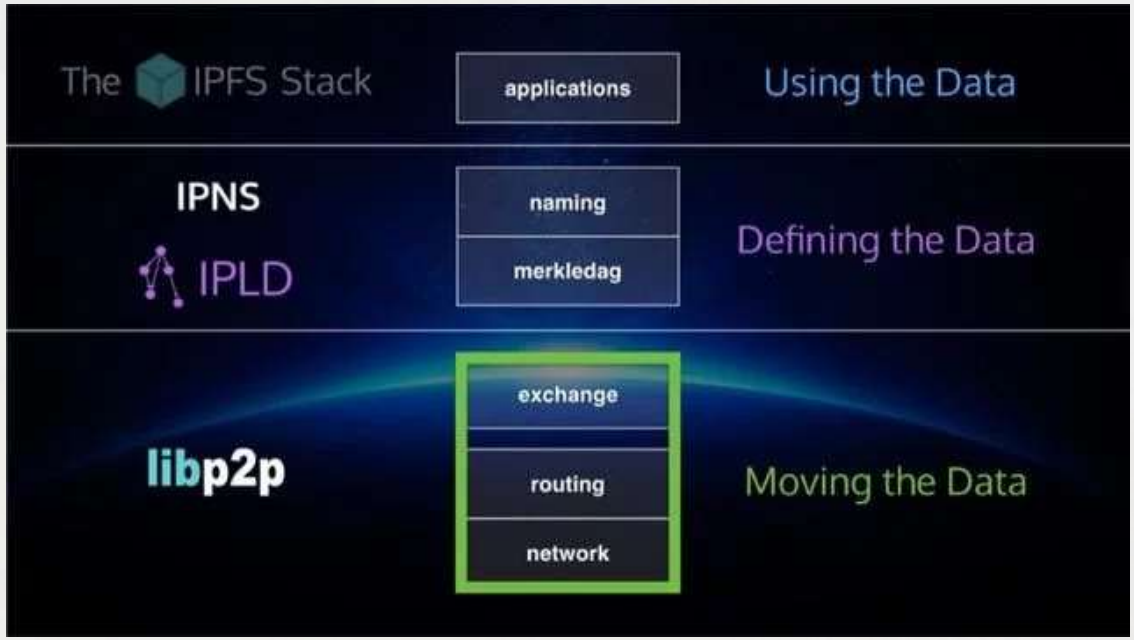
第一层为naming， 基于PKI的一个命名空间；

第二层为merkledag， IPFS内部的逻辑数据结构；

第三层为exchange, 节点之间block data的交换协议，主要是Bitswap；

第四层为routing, 主要实现节点寻址和对象寻址；

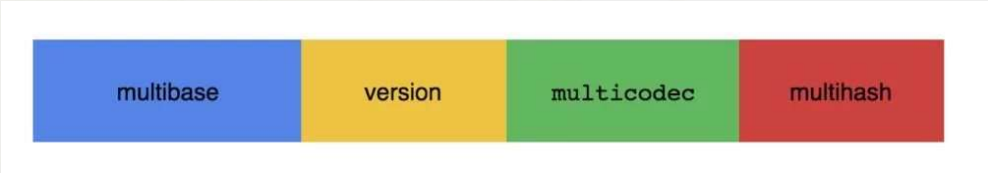
第五层为network, 封装了P2P通讯的连接和传输部分。



IPLD的命名层

- CID（self-describing content-addressed identifiers for distributed systems）基于内容寻址的自我描述表示，内容ID。

CID 是IPFS分布式文件系统中标准的文件寻址格式，它集合了内容寻址、加密散列算法和自我描述的格式, 是IPLD 内部核心的识别符。目前有2个版本，CIDv0 和CIDv1。



CIDv1的格式

- IPNS（InterPlanetary的命名空间）人类友好名称

在可变和不可变的路径之间建立一个很容易辨认的区别，为了程序也为了人类阅读的便利。

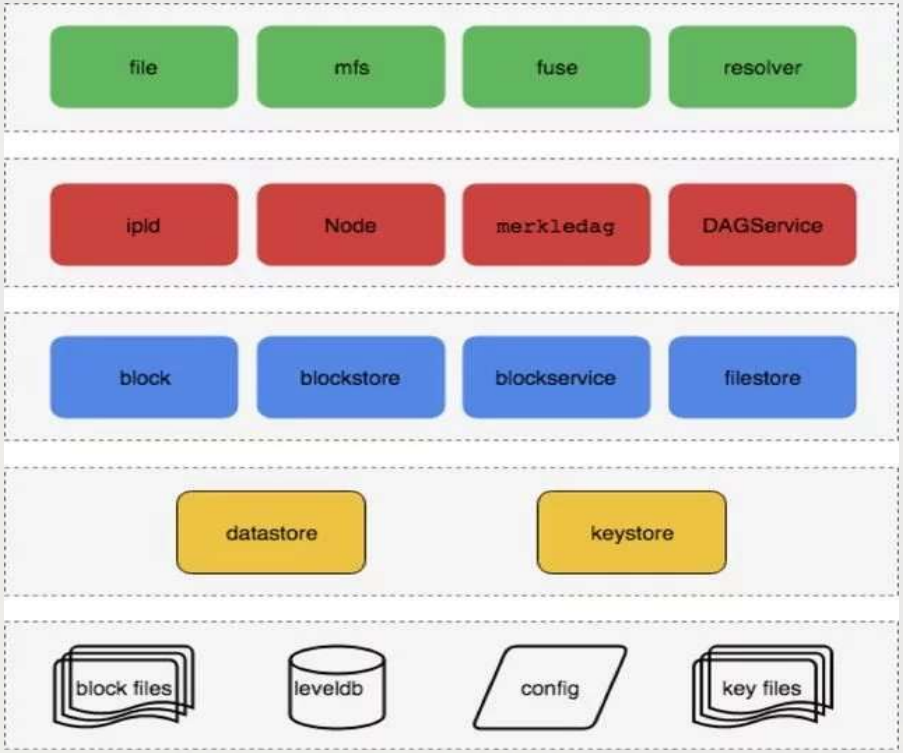
```
1  /ipns/XLF2ipQ4jD3UdeX5xp1KBgeHRhemUtaA8Vm/
2  /ipns/XLF2ipQ4jD3UdeX5xp1KBgeHRhemUtaA8Vm/docs
3  /ipns/XLF2ipQ4jD3UdeX5xp1KBgeHRhemUtaA8Vm/docs/ipfs
```

IPNS的使用

IPLD的MerkleDag层

Merkle directed acyclic graph（默克有向无环图）

1. file，mfs，fuse，resolver是顶层高级接口，将各种异构数据转换成MerkleDag上的DAGNode。
2. MerkleDag，Node管理所有DAGNode之间的增删改和建立连接Link
3. Blockservice将DAGNode和CID包装成文件块block，以适配底层的数据交换和本地存储
4. Datastore和keystore管理本地的文件块，公私钥的存储，以及内存上的节点信息
5. Leveldb，block files是本地数据库



Bitswap协议

易货系统的概念意味着可以创建虚拟货币，Bitswap是一个跟踪货币的所有权和转移的全局分类账本，目前作为节点之间交换文件块的账本，Filecoin的设计来源。

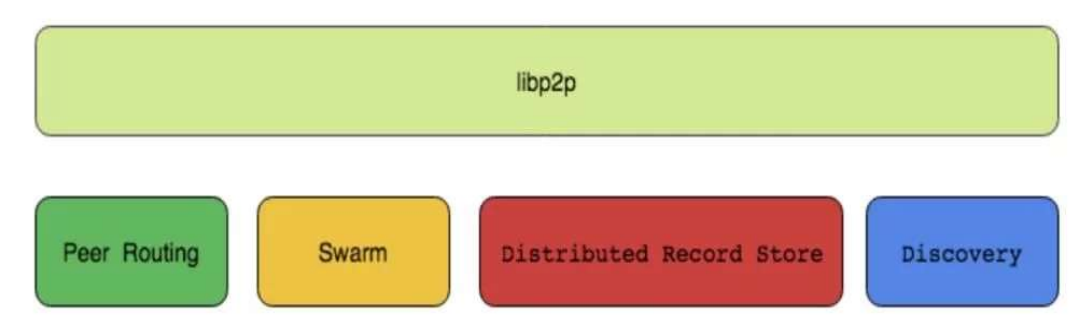
- BITSWAP 信用
  1. 对等节点间会追踪他们的平衡（通过字节认证的方式）
  2. 随着债务人债务增加而降低对等者向债务人发送块的概率

- BITSWAP的策略
  1. 为整个交易和节点最大化交易能力
  2. 为了防止空负载节点利用和损害交易
  3. 对可信任的对等节点更宽容

- BITSWAP 账本
  1. 每个节点都拥有自己的账本
  2. 在交换文件块数据前对比账本内容
  3. 恶意空账本没有获取文件的权利

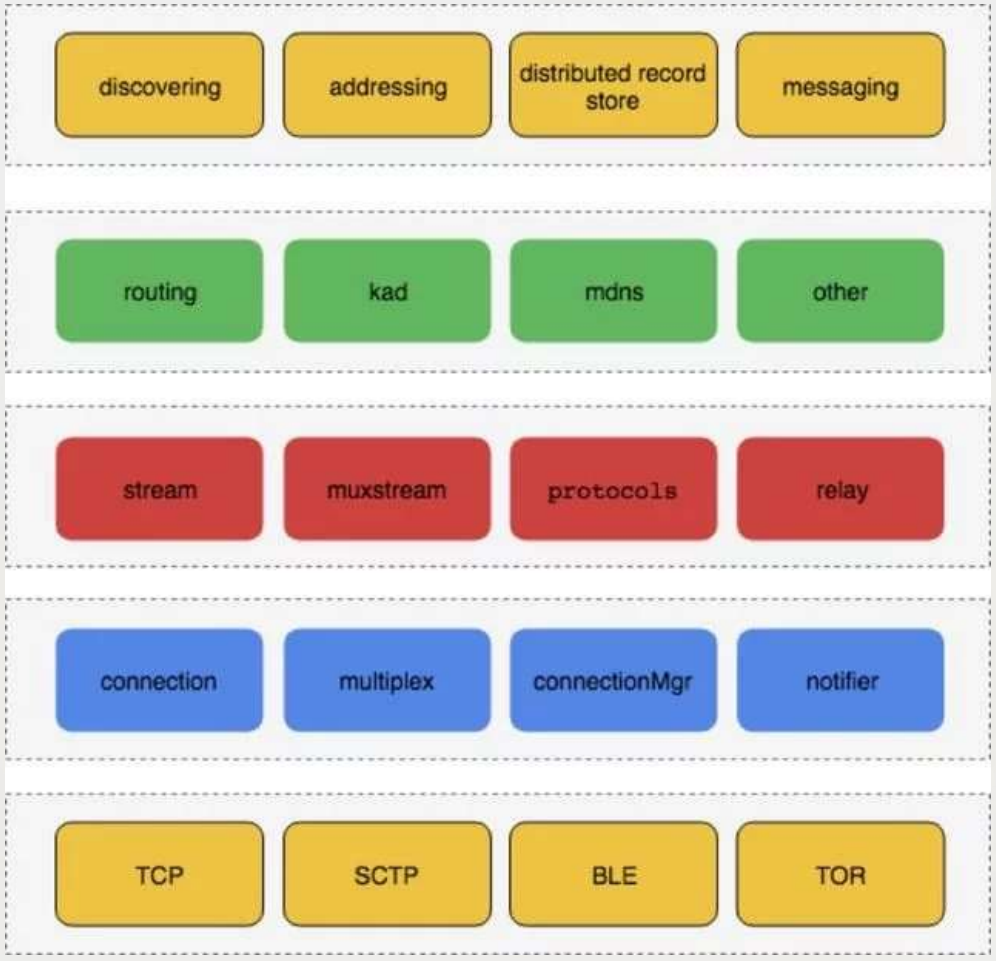
Libp2p的架构

- Peer Routing 路由协议。主要包括KAD routing和 MDNS routing。
- Swarm 传输和连接。有以下接口：
  1. transport 网络传输层的接口
  2. connection 处理网络连接的接口
  3. stream multiplex 同一connection 复用多个stream的接口
- Distributed Record Store 分布式节点的记录存储。包括内存中的datastore和LRU Cache



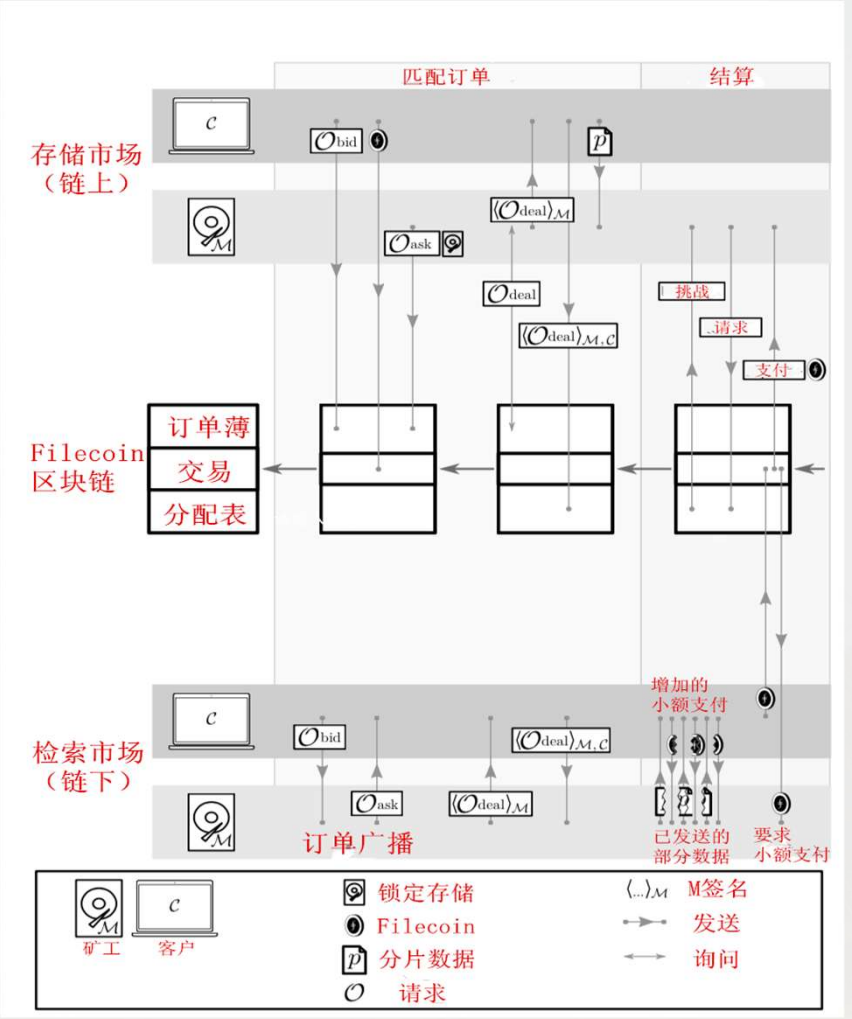
- Discovery 节点发现协议。3种方式：
  1. bootstrap 通过配置的启动节点发现其他的节点
  2. Wsstr 通过查询随机生成的peerID， 从而发现新的节点
  3. mdns 通过multicast 广播发现局域网内的节点

Libp2p模块集合


discovering
addressing
distributed record store
messaging
routing
kad
mdns
other
stream
muxstream
protocols
relay
connection
multiplex
connectionMgr
notifier
TCP
SCTP
BLE
TOR

Filecoin

- 每个新区块
  1. 检查交易、订单等是否处于有效格式并打包，即将订单进行分类，分别为deal、bid和ask，然后添加到数据库中
  2. 对数据库中每个订单进行检查是否过期（或取消）
  3. 验证打包存储提供商提供的存储证明
- 客户
  1. 提交新的存储bid订单或检索bid订单
  2. 从存储矿工接收Odeal并签署，然后提交
- 存储提供商（矿工）
  1. 质押filecoin以保证自己能提供相应的存储量
  2. 提交ask订单，选取合适的存储bid订单并签署发给客户
- 检索提供商（矿工）
  1. 提交ask订单，选取合适的检索bid订单并签署发给客户



总结

- LibP2P是IPFS核心中的核心，面对各式各样的传输层协议以及复杂的网络设备，它可以帮助开发者迅速建立一个可用P2P网络层，快速且节约成本。
- 整个系统由非常多的模块组成，而且是可拔插的，耦合性很低，方便新协议的开发和新功能的加入
- 解决了当前HTTP存在的缺陷，为大型区块链项目提供底层基础。
- 设计Filecoin作为激励层的token，构建了一个可信的去中心化存储网络及市场