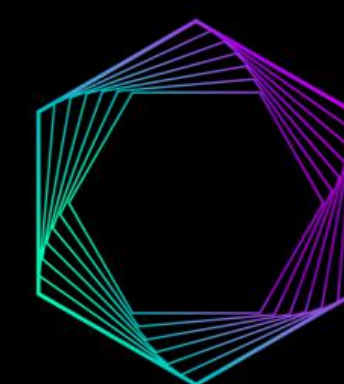




IPFS 带来的变革和意义

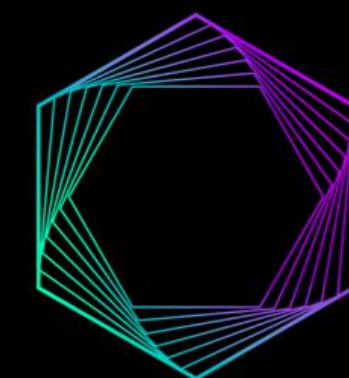
李昕



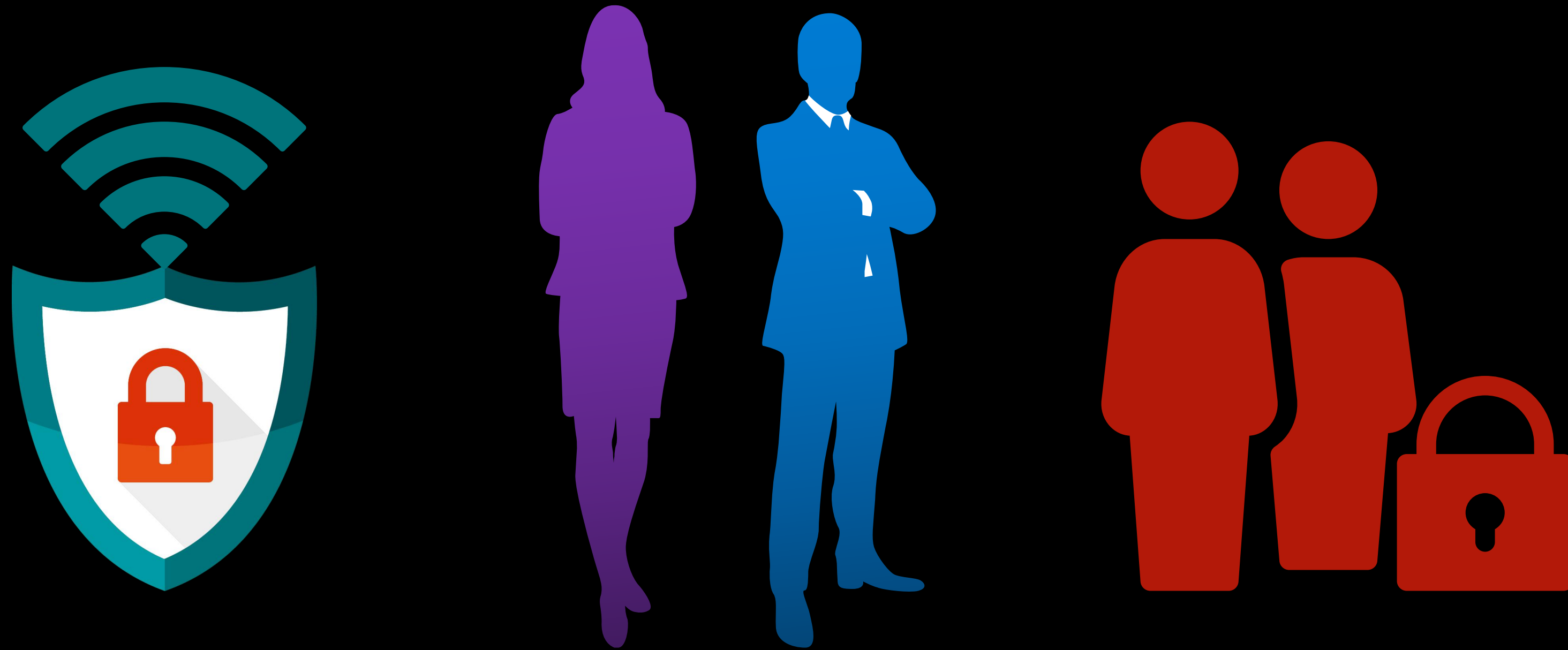
ProtoSchool
Shanghai Chapter

DT 时代

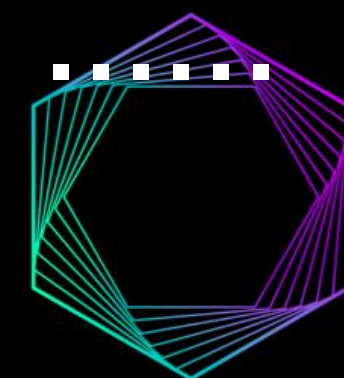
- 数据本身就是生产资料
- 数据的组织和管理则是生产关系
- 算法和系统体现生产力水平



我的数据我做主



安全、完整、可验证、不丢失、易转移、.....



ProtoSchool
Shanghai Chapter

谷歌比利时数据中心遭雷击 部分数据永久丢失

时间: 2015-08-20 15:19 栏目: 主页 > 互联网 >

日前，谷歌位于比利时的数据中心由于遭遇了4次闪电袭击而出现了部分数据丢失的情况。虽然谷歌尝试重新进入受到损害的磁盘，但一些谷歌用户的数据仍旧没能得到挽救，也就是说，他们将永远失去自己的部分个人数据。目前已知的信息只知道受影响的是谷歌计算机引擎(GCE)服务，至于具体哪些客户或哪个国家的数据丢失并不清楚。



对此，谷歌特别强调，即便事态再严重丢失的数据仍非常非常地小，永久被删除的数据只占了该数据中心的0.000001%。



腾讯云给一家创业公司带来的灾难！

腾讯云给一家创业公司带来的灾难！ 发布于 2018-08-05 20:39:21 | 阅读量: 89363

腾讯云到底安不安全？
为什么数据丢了不能恢复？
——腾讯云给一家创业公司带来的灾难！

腾讯云，一个听起来就很牛逼的品牌。其云服务商声称99.999999%的数据可靠性，搭载了云硬盘提供三副本存储策略，也就是说只要把数据放在腾讯云上，只有十亿分之一出现数据丢失的可能性，另外还对数据提供了3个备份，这简直太安全了！

6月27日阿里云故障说明

6月27日下午，我们在运维上的一个操作失误，导致一些客户访问阿里云官网控制台和使用部分产品功能出现问题，引发了大量吐槽。故障于北京时间2018年6月27日16:21左右开始，16:50分开始陆续恢复。

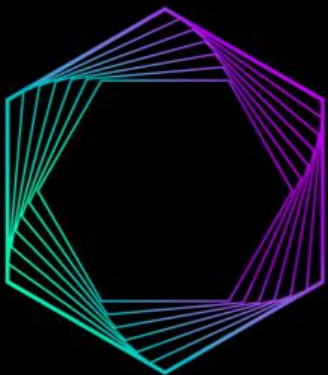
经过紧急技术复盘，故障原因如下：
当天下午，工程师团队在上线一个自动化运维新功能中，执行了一项变更验证操作。这一功能在测试环境验证中并未发生问题，上线到自动化运维系统后，触发了一个未知代码bug。错误代码禁用了部分内部IP，导致部分产品访问链路不通。后续人工介入后，工程师团队快速定位问题进行了恢复。

受影响范围包括阿里云官网控制台，以及MQ、NAS、OSS等产品功能。
对于这次故障，没有借口，我们不能也不该出现这样的失误！我们将认真复盘改进自动化运维技术和发布验证流程，敬畏每一行代码，敬畏每一份托付。

阿里云计算有限公司
2018年6月27日



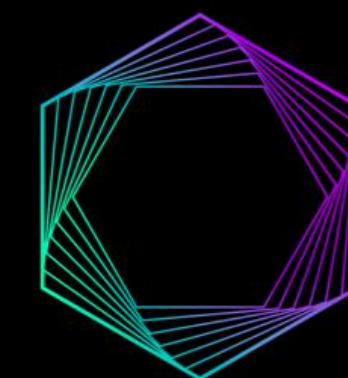
一些中心化存储的灾难性事件



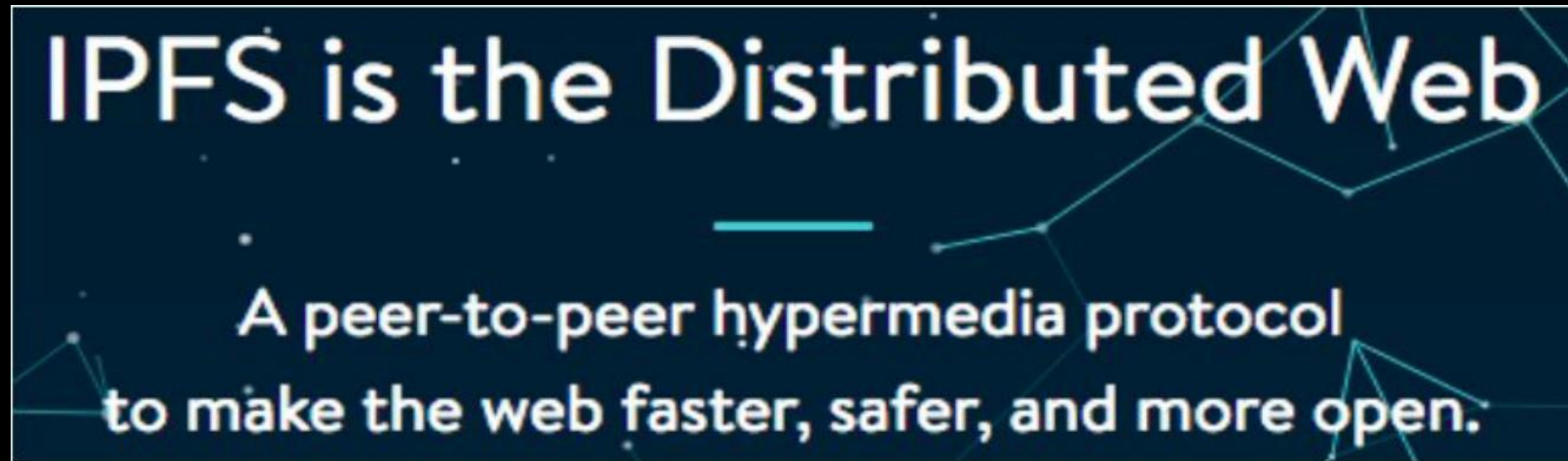
ProtoSchool
Shanghai Chapter

中心化云存储存在的问题

- 1、Single Point Of Failure: 单点故障，影响巨大
 - a) Converged IO: 集中式访问，带宽需求巨大而集中
 - b) Centralized DB: 中心化数据库，一旦数据库出问题，数据难以找回
- 2、unverifiable Storage: 非可验证存储，内容被改变而不知
- 3、Path addressing: 断链问题、内容可达性问题
- 4、Limited Scalability: 有限的扩展能力，难以胜任飞速发展的存储需求
- 5、Data are not natually linked: 数据的关联与存储没有直接关系
- 6、InterData? 数据中心之间互不相通，没有标准



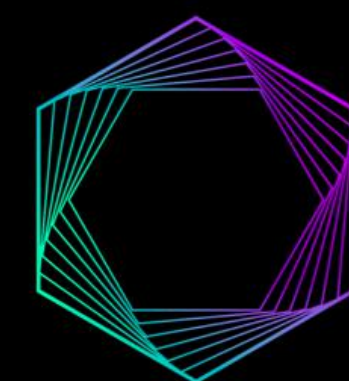
IPFS带来全方位的改变



IPFS是分布式的互联网

- 是一套点对点的超媒体协议，它能使网络更快速，更安全，也更开放

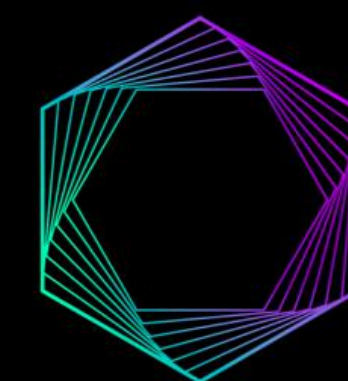
IPFS是通用存储网络的开拓者，是领军者



ProtoSchool
Shanghai Chapter

IPFS之Libp2p - 网络层的去中心化

- 分布式是通用存储网络的基础
 - 网络构建兼容现有网络，去中心化
- 点对点网络的应用完全可以基于Libp2p建立起来
 - Libp2p & MultiFormats 是 IPFS的网络支撑
 - Libp2p已经成为独立的项目
 - Libp2p兼容当前多数网络协议
 - 架构设计扩展性强

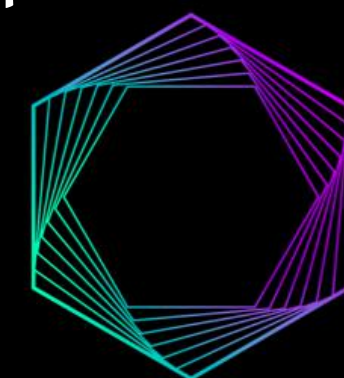


IPFS之内容寻址 - 语义网络

- 内容按照路径寻址，是导致网络逐步中心化的重要技术原因
- IPFS的内容寻址技术，数据不再需要从单个节点读取，自动全网搜索下载
- IPLD - IPFS的数据组织方式，提供了一种通用的数据组织模式
 - IPLD是一个独立的项目
 - 不仅可用于IPFS，也可用于区块链或其他数据的组织和链接方式
 - 通过hash（内容指纹）链接
- 内容寻址和IPLD为数据去中心化存储和语义网络提供了技术基础

~~服务商
依赖性~~

~~断链
问题~~



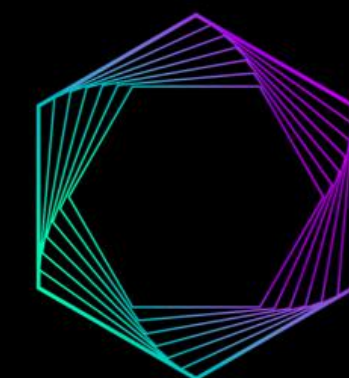
ProtoSchool
Shanghai Chapter

IPFS之自认证文件系统 - 内容安全

- 内容寻址不仅仅带来的是数据存储和组织方式的改变，也同时带来了安全的提升
- 内容不可更改
 - 内容的篡改和损坏立即被知道，无法遁形
- IPNS发布内容必须通过拥有者认证
- 网页的链接通过内容指纹 (hash) 进行，同样不可更改
 - 因此通过超链接的访问与访问原网页一样安全

~~不可验证
问题~~

~~数据被
篡改~~



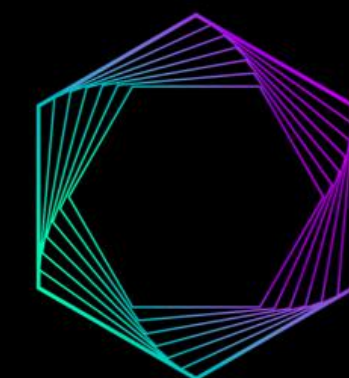
ProtoSchool
Shanghai Chapter

IPFS之分布式Hash表 - 数据库冗余

- Meta-Data完全分散化存储, 没有单点故障
- 存储无上限, 扩展能力更强
- 数据库查询与数据读取分离
- 合理分配带宽

~~扩展能力
不足~~

~~带宽利用
集中~~

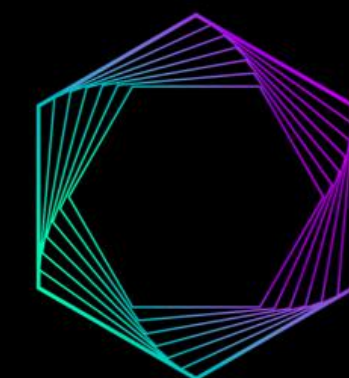


IPFS+Filecoin - 价值与数据的融合

- Web2.0: “羊毛出在猪身上，狗来买单”
- Web3.0: 价值网络，数据流转相结合
 - 让内容体现价值，让内容的转移伴随着价值转移。让世界回归其本来的逻辑。
- IPFS与BlockChain结合，为区块链提供数据存储，提供所有权证明，交易保障；智能合约保证访问控制
- 通过IPLD，IPFS本身可存储区块链的账本数据

~~数据所有
权不明~~

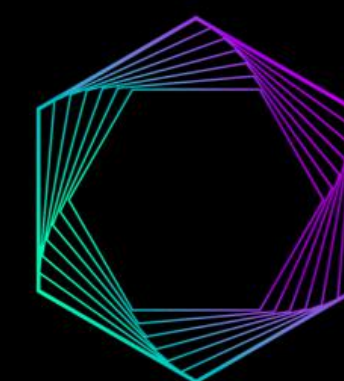
~~数据交易
分离~~



ProtoSchool
Shanghai Chapter

IPFS 所创造的

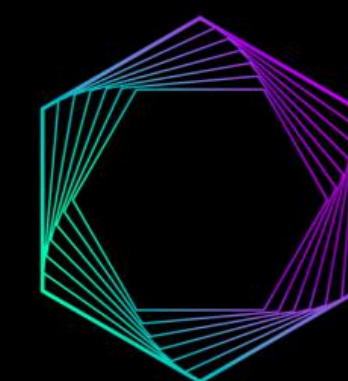
通用存储网络



ProtoSchool
Shanghai Chapter

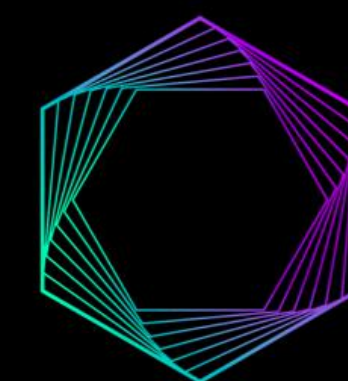
通用存储网络 - 我的数据我做主

- 如何构建一个更加公平的数字世界，让你的数据给你带来价值？
- 分散式存储是一个重要的解决之道。让你的数据存储在你同意存储的地方，或者就让你自己来存储自己的数据。数据的使用需要授权和产生价值。
- 应用应该建立在这些数据存储之上，形成统一的底层存储架构，为各种应用提供服务，应用之间可以共享数据。利用加密体系来保护用户隐私，让数据的随意复制成为不可能。让数据的拥有者来控制数据的访问和交易，真正实现数字所有权机制。
- 数据的存储设施完全可以像网络一样成为数字世界的基础设施，作为底层平台，只有公共协议，大家共建共享共同维护。让数据的存储协议和访问成为标准，这是数据为拥有者掌握的一个路径。这里有很长的路要走，但时候已经到来。



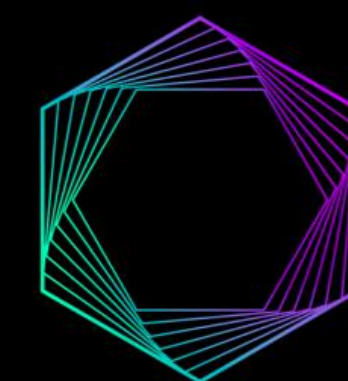
通用存储网络的特性 - 1

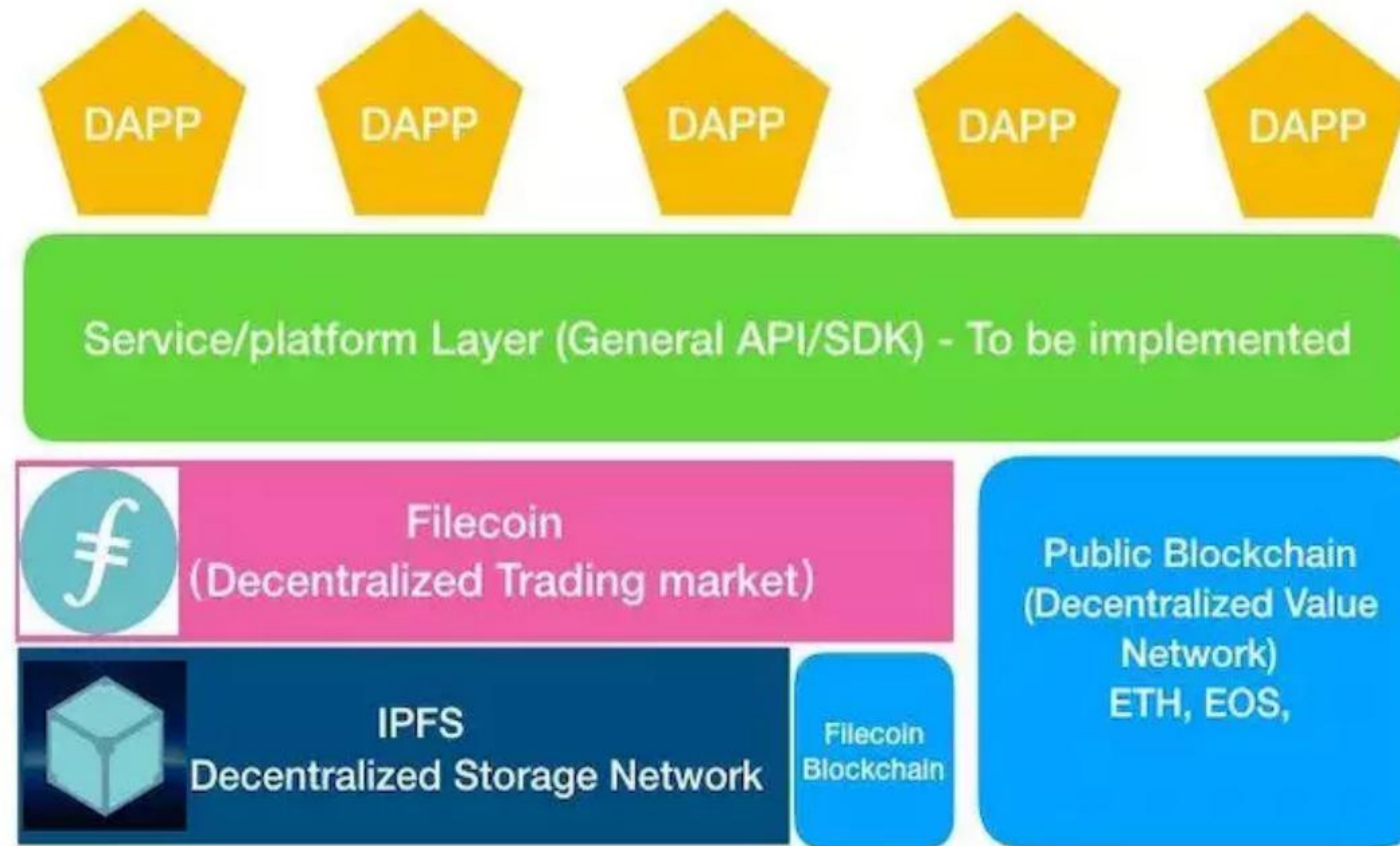
- 区块链确权：解决数据的归属性问题
- 基于脚本的访问控制：与智能合约相结合，解决数据的转移和价值交换的问题；可以有标准的通用的智能合约和数据访问服务，也可以有定制的智能合约。这给数据的访问控制的灵活性和创造性带来前所未有的巨大空间。
- 基于语义网络的数据关系处理：打破传统的简单的文件层级模式，采用灵活的数据关系模型，可以描述复杂的数据之间的关系
- 基于密码学的数据安全管理和数据共享：用户不用在一个服务商那里注册和身份认证，而是使用全网统一的基于密码学的身份系统进行身份认证和鉴权



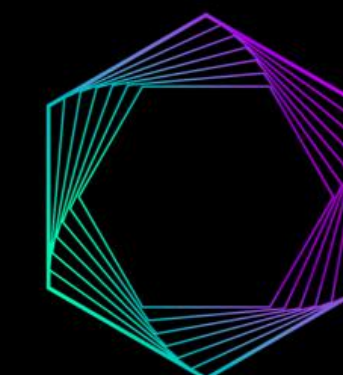
通用存储网络的特性 - 2

- 数据与应用分离：促进数据的自由流动。数据可服务于不同的应用，用户拥有自己的数据，而不是服务/应用提供商拥有数据
- 数据自由迁移：用户有权随时迁移数据至不同的服务商或服务节点。数据服务商仅仅无差别地服务用户数据，而与数据内容以及关联关系无关
- 自主分发网络：服务提供商可以考虑为数据加速来提高服务质量和响应速度，从而提高收益



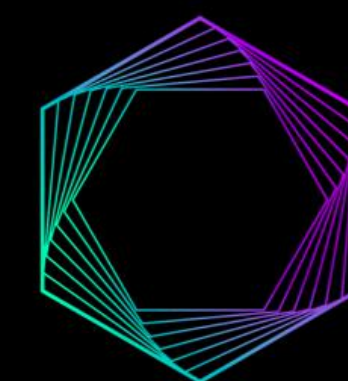


利用 IPFS/Filecoin 构建通用存储网络，实现Web3.0



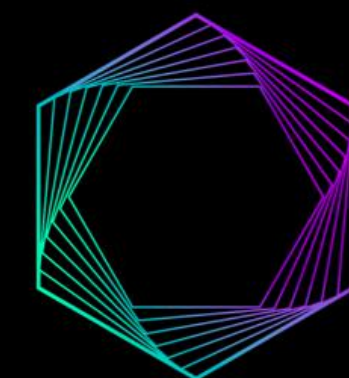
IPFS+IoT+5G+EC -> 无限可能

- IPFS 组建网状网络，存储走向边缘
- 物联网设备选择就近存储；物联网设备也可以提供存储
- 5G网络使数据的自由流动，点对点网络称为现实
- 存储到边，赋能边缘计算，存储无处不在，计算无处不在



和现有存储有哪些不同

- 网络智能大幅提升：比如区块链确权，密码学为基础的身份认证和访问控制，自主分发网络的实现等等，都要以网络智能的大幅提升为基础
- 去中心化：之前由中心化处理的问题，要由网络实现
- 打破局限性的数据关系和访问控制管理：传统的文件系统以文件夹文件的模式存储，体现一种层次感，但这种模式已经不能反映数据之间的复杂管理，应该以一种更加自由的数据关系管理方式供用户选择，以体现数据的任意连接；同时，传统的基于用户组的管理模式的局限性也应该打破。使用脚本为基础的访问的模式给扩展提供了极大的想象空间
- 数据与应用分离：应用开发者或服务商不再拥有数据，而是直接访问用户产生的数据。举个例子，短视频应用将采用通用的接口来引用用户的数据，而用户使用短视频应用时产生的视频存储在通用的存储网络之中，数据归用户所有，仅仅是授权此应用访问，同时，用户可以把这个视频文件自由地用于其他应用。
- 价值交换与数据交换相结合：现有的多数应用，是用户用数据换取应用的使用权，同时应用服务商利用广告获取利润。在通用存储网络中，用户的数据交换直接与价值交换挂钩，无需通过其他形式转换。



共同探讨学习

- [Web3.0通用存储网络将非常不同](#)
- [物联网时代，边缘存储将大行其道](#)
- [我的数据我做主](#)
- [IPFS如何助力第三代互联网](#)
- [存储的演进之路 - 分久必合、合久必分](#)
- [IPFS不是一个单纯的产品，而是一个体系](#)
- [IPFS存储七问](#)

