使用 IPFS 或 CEPH 作为 StreamNet 后台存储的可行性的报告

1. IPFS 和 CEPH 都是比较复杂的分布式存储系统， 对于原理、功能、性能、部署等的了解都需要较长的时间。短时间内的调研，我们只能从功能点上去分析。

另外资料的获取，由于时间的限制，除了自己手动的尝试部署之外，大多是采用网络上的数据，但是都是经过对比分析的，大概率可信。

2. StreamNet 目前使用的存储情况是：后台的数据库是 RocksDB， 存储形式是Key-Value 形式，有多个 column family。

3. 我们的需求是存储的可扩展性，另外需要较好的性能和稳定性。

关于稳定性，IPFS和CEPH是目前比较热门的开源软件，稳定性基本上是有保证的。

如果IPFS和CEPH能够支持可扩展性，并且性能不错，那么是可以作为StreamNet的后台存储的。

另外，最好是支持K-V形式的存储，否则我们还需要另外想办法。

4. 先看 **CEPH**。

大多介绍CEPH的材料上都会有这么一句话：“Ceph是一个统一的分布式存储系统，设计初衷是提供较好的性能、可靠性和可扩展性。”其实它是官方页面上的一句话的翻译：

“Ceph is a unified, distributed storage system designed for excellent performance, reliablity and scalability.”

“可扩展性”、“稳定性”（“可靠性”）、“性能”看上去都能够满足。

1）扩展性

看到一些资料上说，CEPH可以轻松扩展到PB级别的存储。“扩展性”方面应该是没问题的。

2）KV的支持

CEPH底层使用的是KV形式的存储器，可以选择LevelDB或者RocksDB。但是封装之后没有看到可以直接支持KV存储，它提供三种接口：对象存储、块存储、文件存储。对象存储一般是图片、视频的存储；块存储一般是裸硬盘空间的映射；文件存储就像是linux 上的文件系统。

目前从搜索的资料来看，CEPH不支持直接的KV存储。将来我们需要在这一块做转换处理

3）性能

网络上有很多资料在吐槽CEPH的IO路径过长，会影响到存储性能，另外有很多资料提到了如何优化CEPH，比如使用SSD硬盘、优化架构方案等。这样来看，性能这块确实有需要提升的地方。

CEPH自身也带了一些性能测试工具。CEPH的结构复杂，这样就对性能测试造成了麻烦，对比测试比较困难，测试模型也比较复杂。

从网络上获取的CEPH与原生硬盘的性能对比，数据较多，就不拷贝了，在网页上看更加方便，[**https://yq.aliyun.com/articles/112680**](https://yq.aliyun.com/articles/112680)

5. **IPFS**

之前StreamNet 在收到交易时，是会将交易内容备份一份到IPFS上的，后来有个需求，是将这个功能可以设置为可以选择打开或者关闭，以后就很少涉及到IPFS了。

这个任务是我实现的，所以是使用过IPFS，使用下来的体会就是性能比较差，从发送一份数据进行存储到最后确认，往往需要秒级的时间。

另外，IPFS项目的规划比较宏大，它的目标是形成一个新的网络，作为存储工具来使用可能会有不合适的地方。

另外，不知道是否是因为涉及到代币（FILECOIN），IPFS的网站在国内是被封禁的。

6. 其他产品

除了上面的CEPH和IPFS之外，还有很多其他的分布式数据库产品。

国内有家做分布式数据库的公司PingCAP，有支持KV的数据库TiKV。之前我对TiKV做过小的调研，也使用了一下。好像目前应用的公司还挺多，稳定性应该还可以；它是基于Google 的论文开发的，理论上具有无限扩展的能力；性能测试也比较复杂，总的来说，性能肯定比硬盘读写要差不少。

另外，市场上的分布式数据库有很多，由于时间限制，来不及去做更多的调研。

7. 总结

CEPH 是可以尝试使用的，但是需要设计一套方案来做转换，它是无法直接使用；IPFS使用起来，性能可能会成为瓶颈；TiKV也可以尝试一下。