



DisGraFS

分布式图文件系统

中国科学技术大学 USTC CS19

彭怡腾 彭浩然 黄晋超 朱一鸣 袁玉润

2021.4.26

分布式图文件系统

0

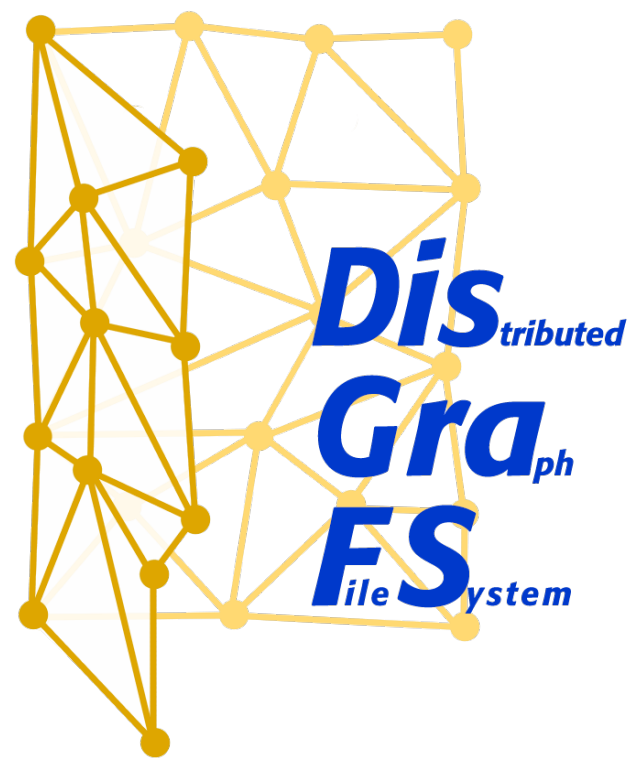
What: 什么是分布式图文件系统

1

Why: 为什么需要分布式图文件系统

2

How: 如何构建分布式图文件系统

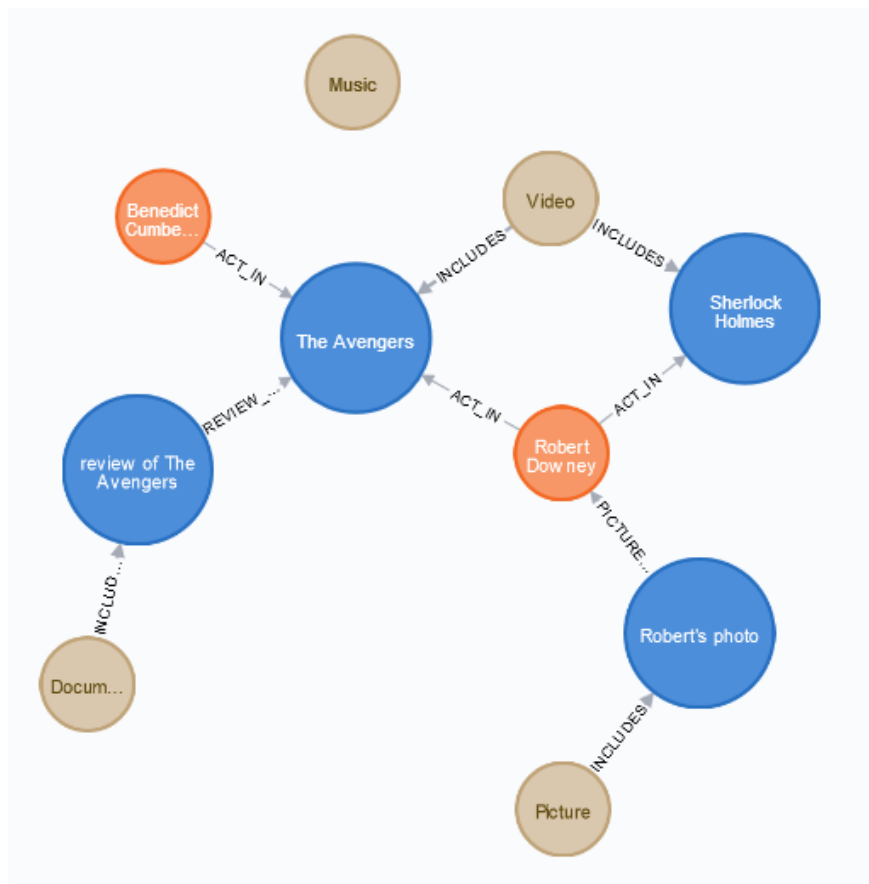




0. WHAT

什么是分布式图文件系统

什么是图文件系统

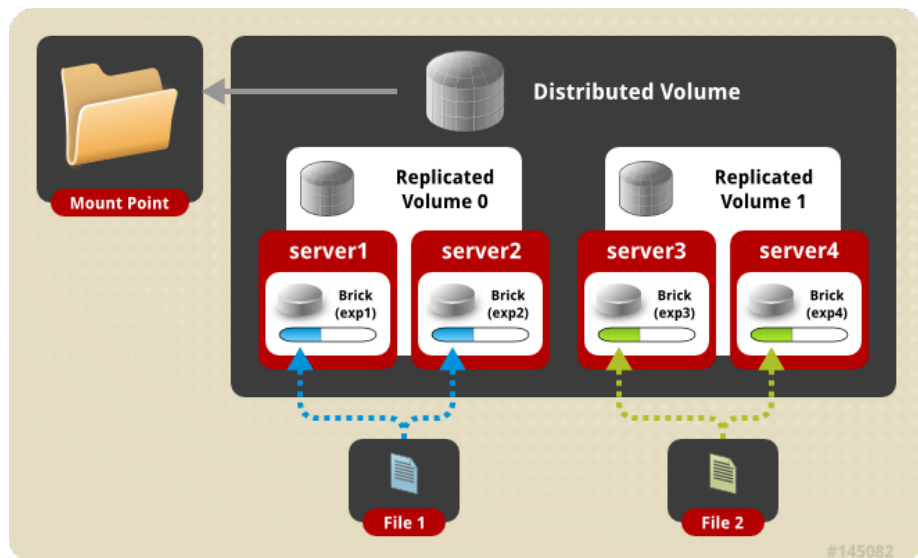


- 逻辑上抛弃树状结构
- 文件之间用“关系”连接
- 基于语义的局部性：有共同特征的文件相连
- 对存储的需求：基于现有文件系统
- 对计算的需求：对文件语义识别

什么是分布式系统

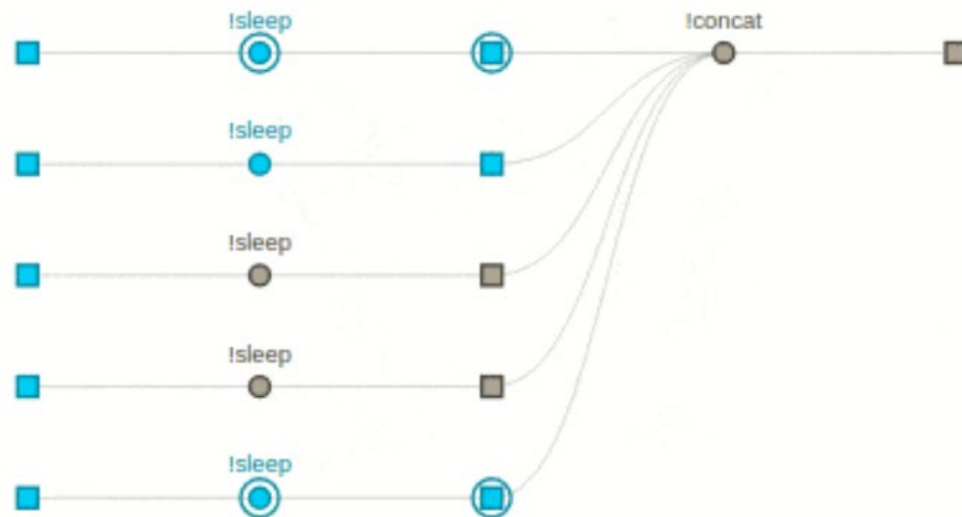
分布式存储

- 块存储、对象存储、分布式文件系统.....
- 将单个文件分块存储在不同的机器上
- 充分利用带宽、冗余备份、高安全性、高可用性.....

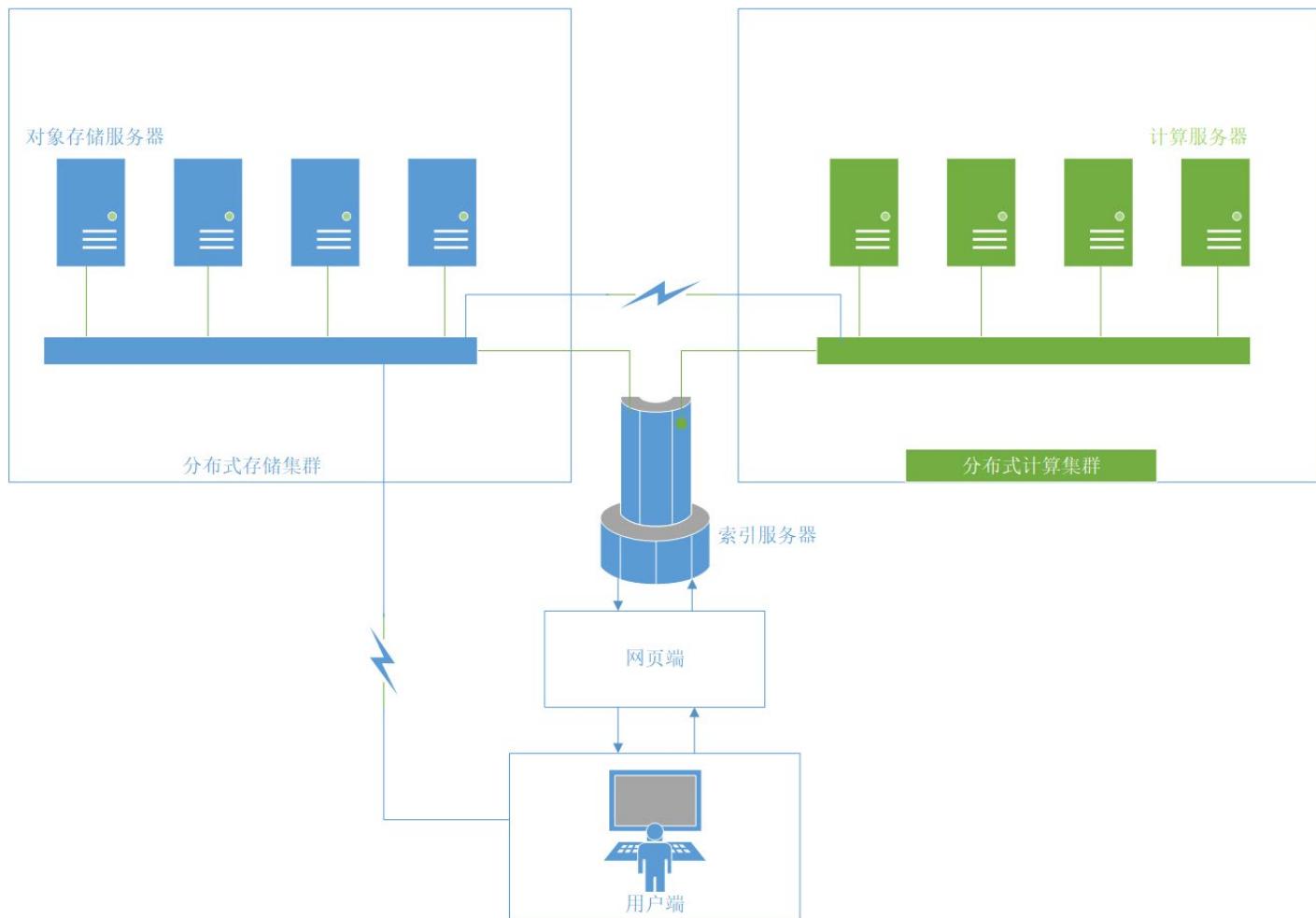


分布式计算

- 将计算任务分配给集群中的不同机器，有序执行
- 可靠性、可扩展性、高性能



什么是分布式图文件系统



- 底层存储采用分布式存储
- 语义识别采用分布式计算
- 使用图结构描述文件之间的关系



1. WHY

为什么需要分布式图文件系统



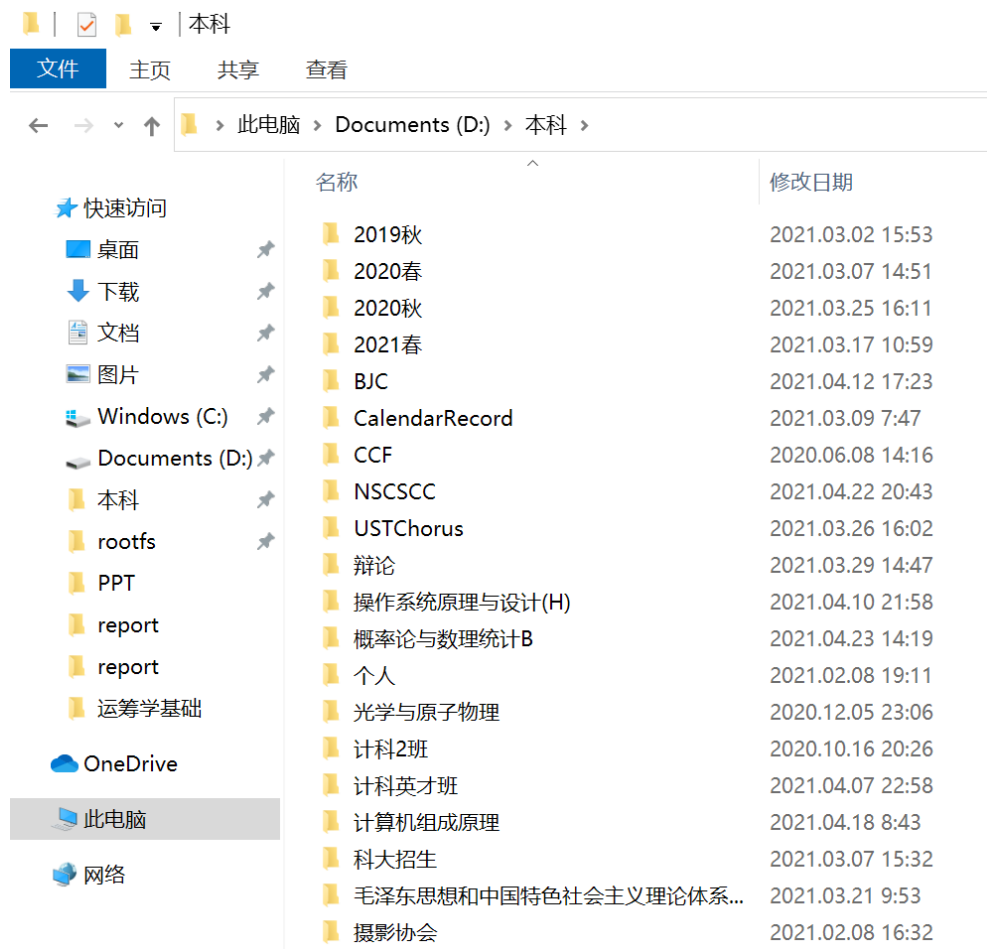
图文件系统更贴合人的思维习惯



- 树形结构方便机器管理
- 图状结构方便人类用户检索



个人使用场景无法充分发挥图文件系统的作用



- 用户会对自己整理的文件有印象
- 人的“印象”本身就是一种图索引



Master status: Ready

图文件系统需要走向大规模化、分布化



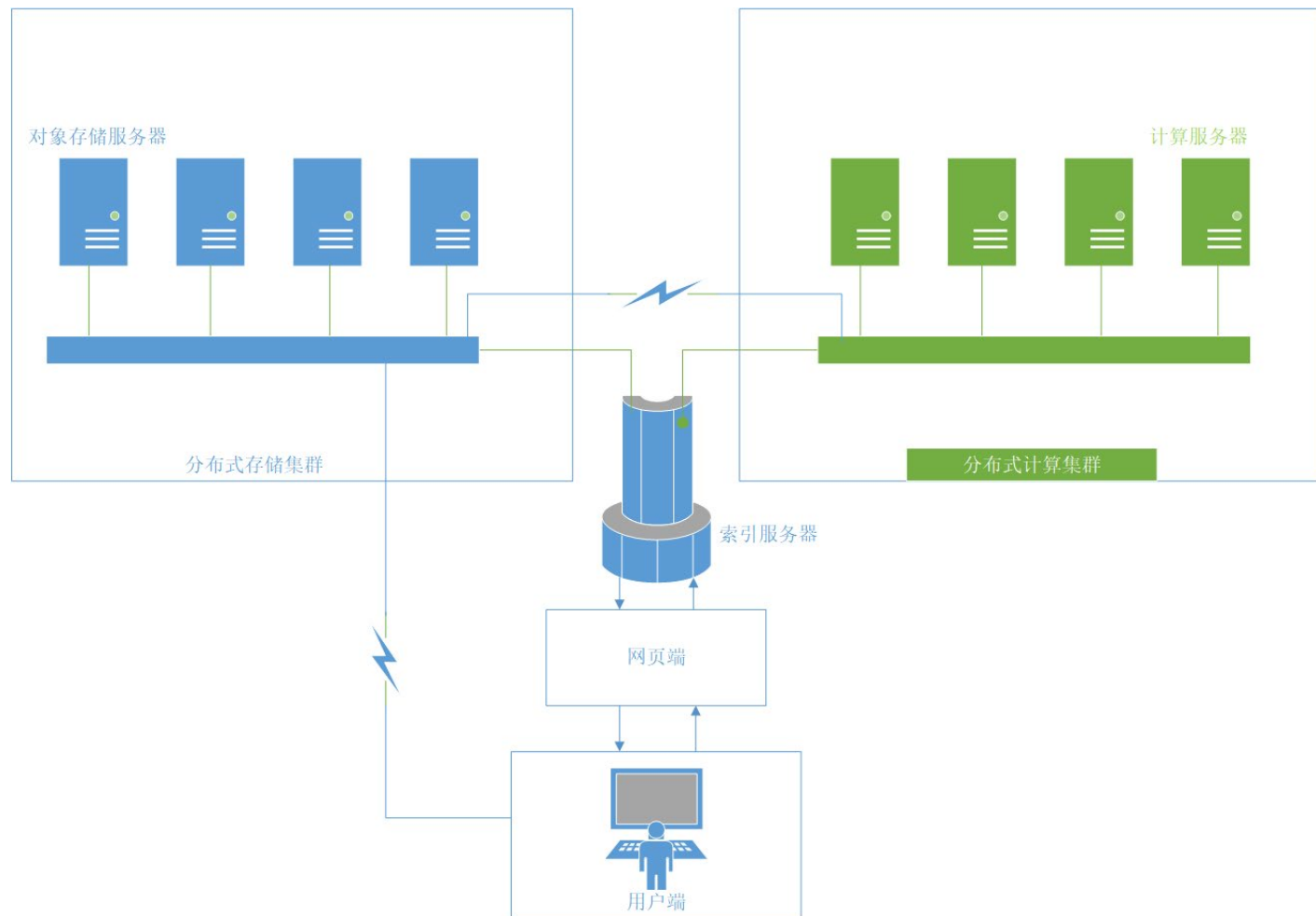
- 越大的文件系统，越需要语义的局部性
- 分布式存储提供高安全性
- 分布式计算提供高效率



2. HOW

如何构建分布式图文件系统

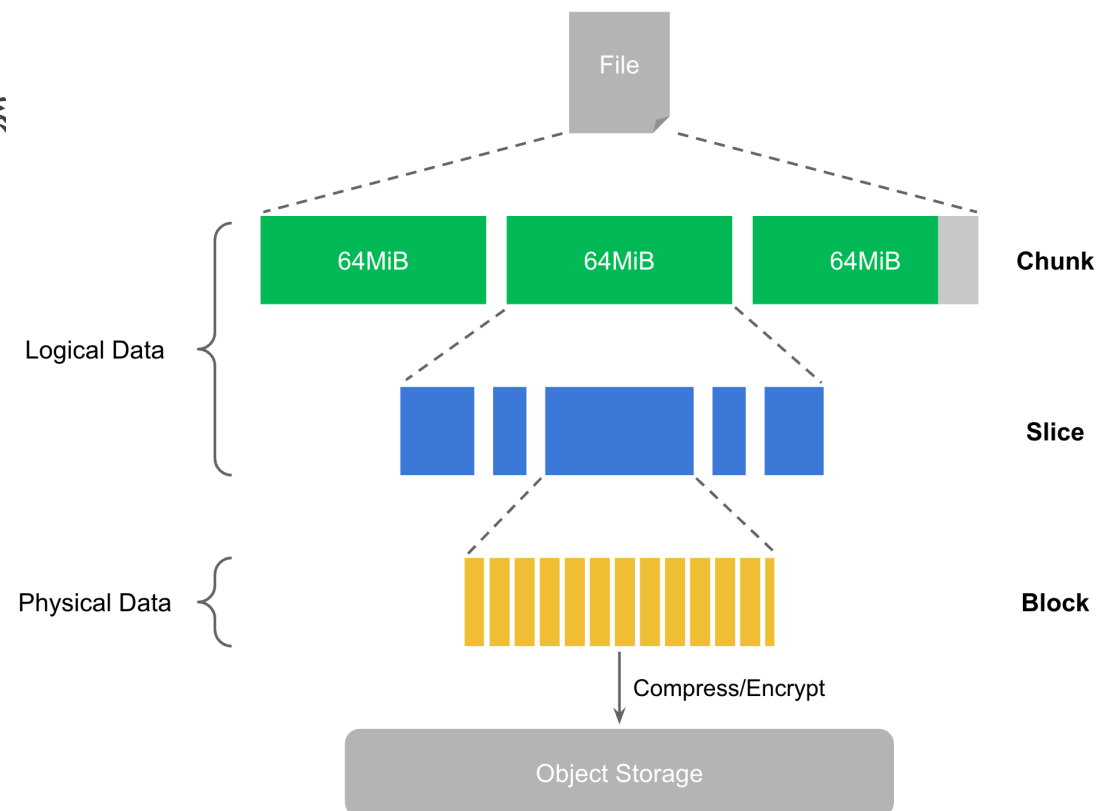
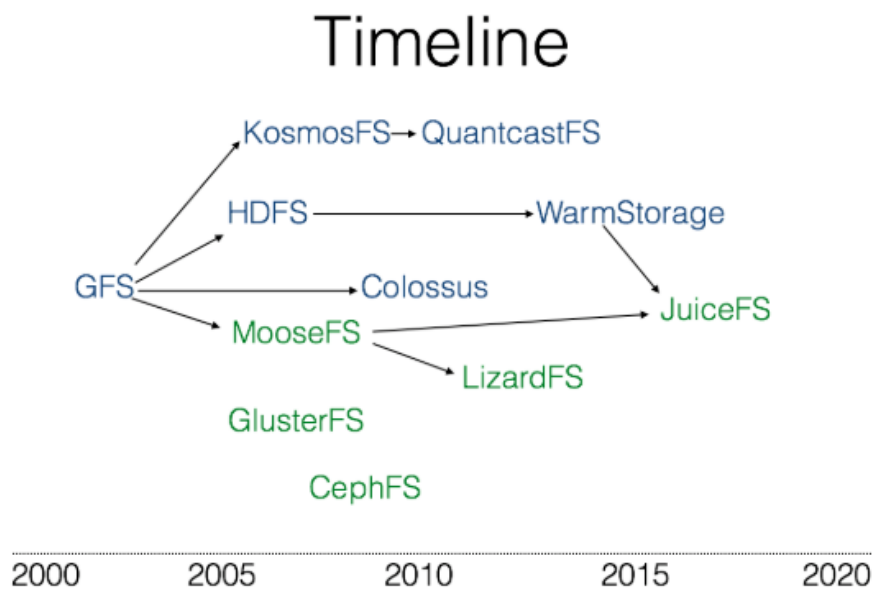
系统架构



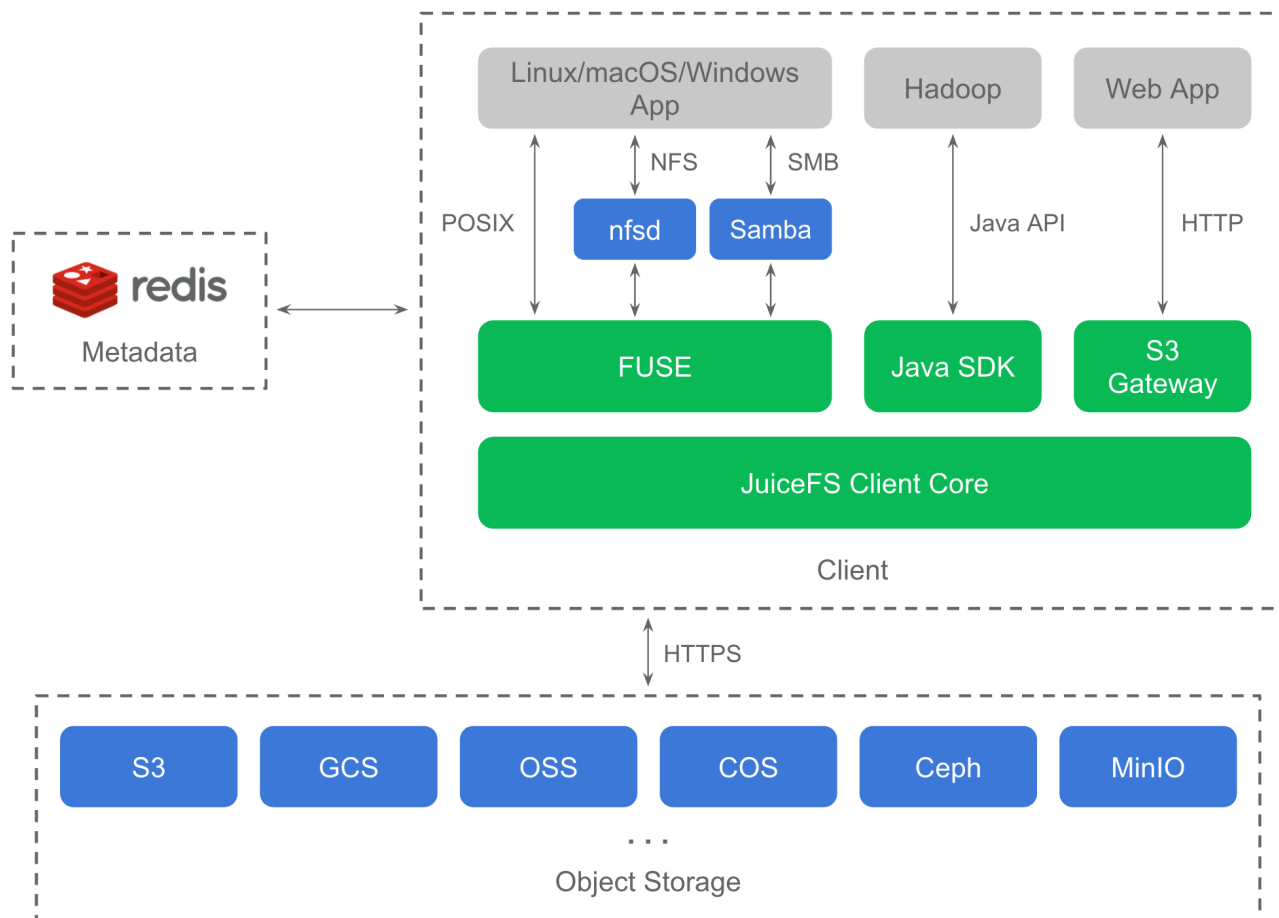
- 中央索引服务器负责调度、维护图数据库、提供用户接口
- 分布式存储集群负责底层存储
- 分布式计算集群负责语义识别任务

分布式存储集群：JUICEFS+对象存储

- JuiceFS：2021年开源，基于对象存储的分布式文件系统
- 相当于一个无状态的中间件，帮助各种应用通过标准的文件系统接口来共享数据

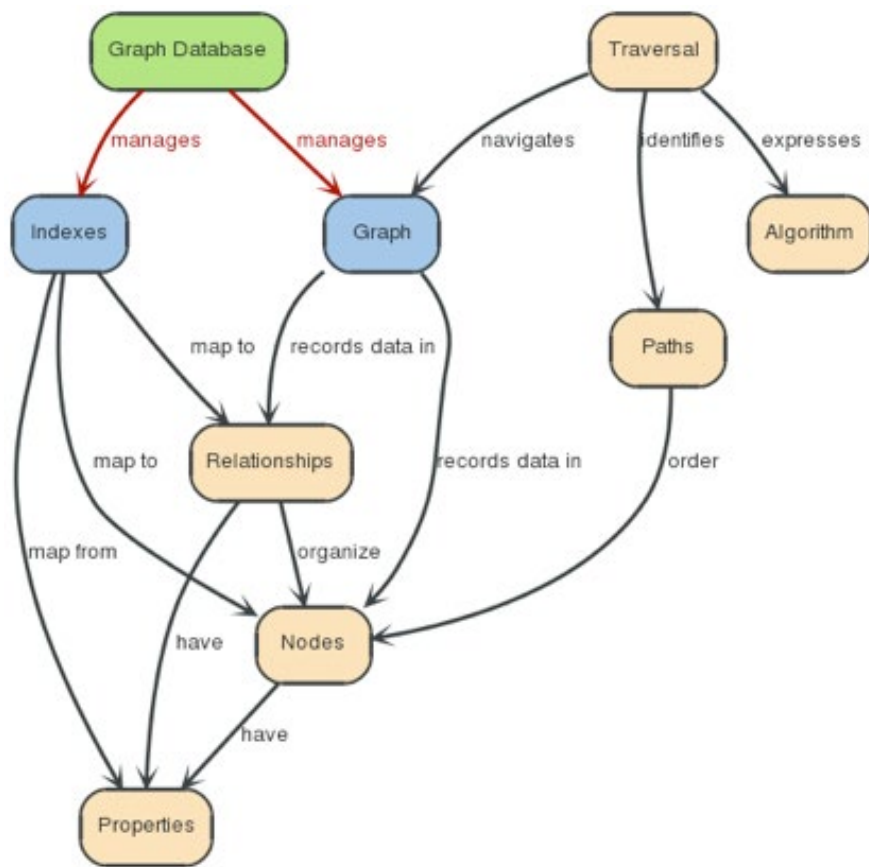


分布式存储集群：JUICEFS+对象存储



- 完整POSIX、S3、Hadoop兼容，服务端、客户端全平台兼容
- 支持阿里云、华为云、Ceph等各种对象存储
- 毫秒级的延迟，近乎无限的吞吐量（取决于对象存储规模）
- 支持数据加密与数据压缩

图数据库



- 节点
- 关系
- 属性
- 图操作：创建、删除、查找、遍历……

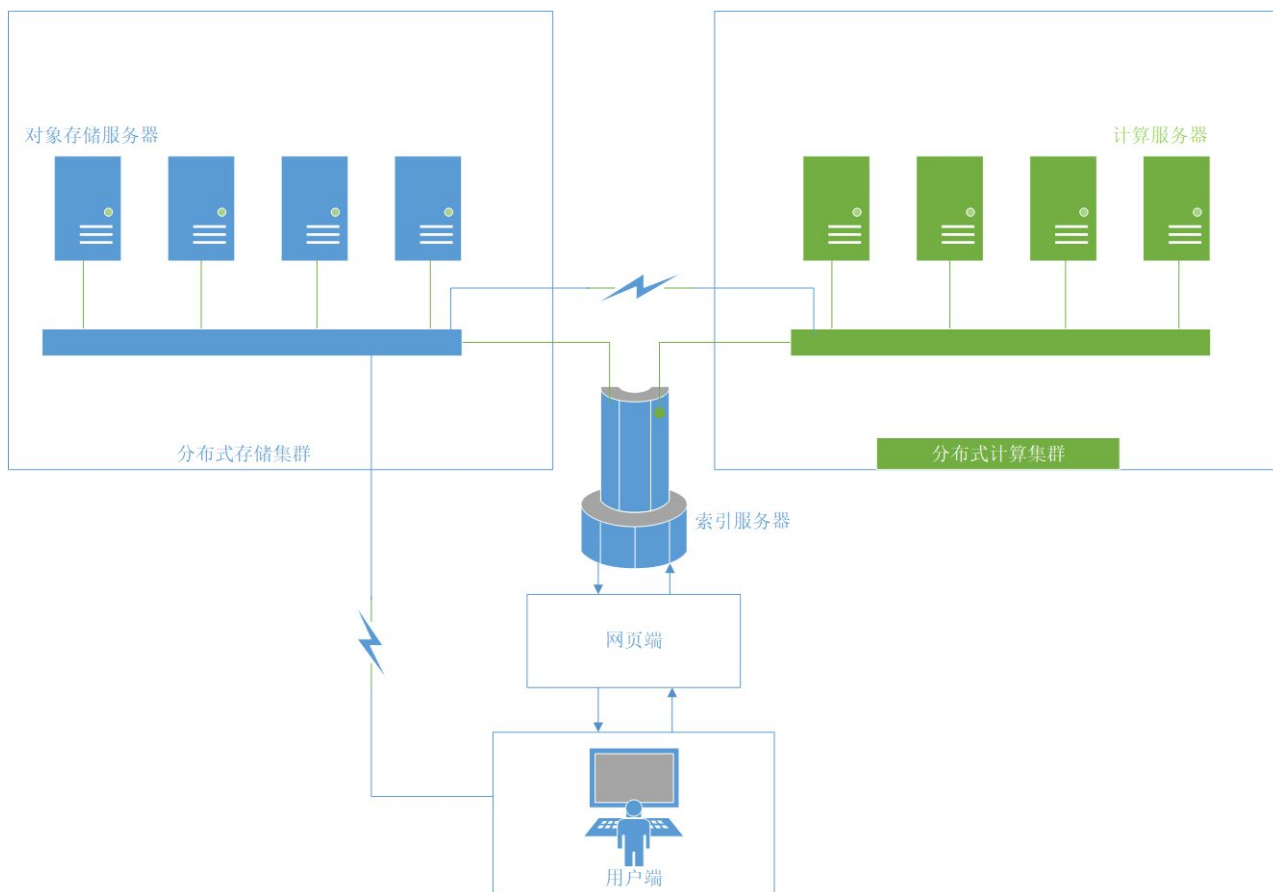
图数据库：NEO4J



- 极高的市场占有率
- Nosql-图数据库
- Cypher语言
- 强大的可视化工具
- 丰富的API接口

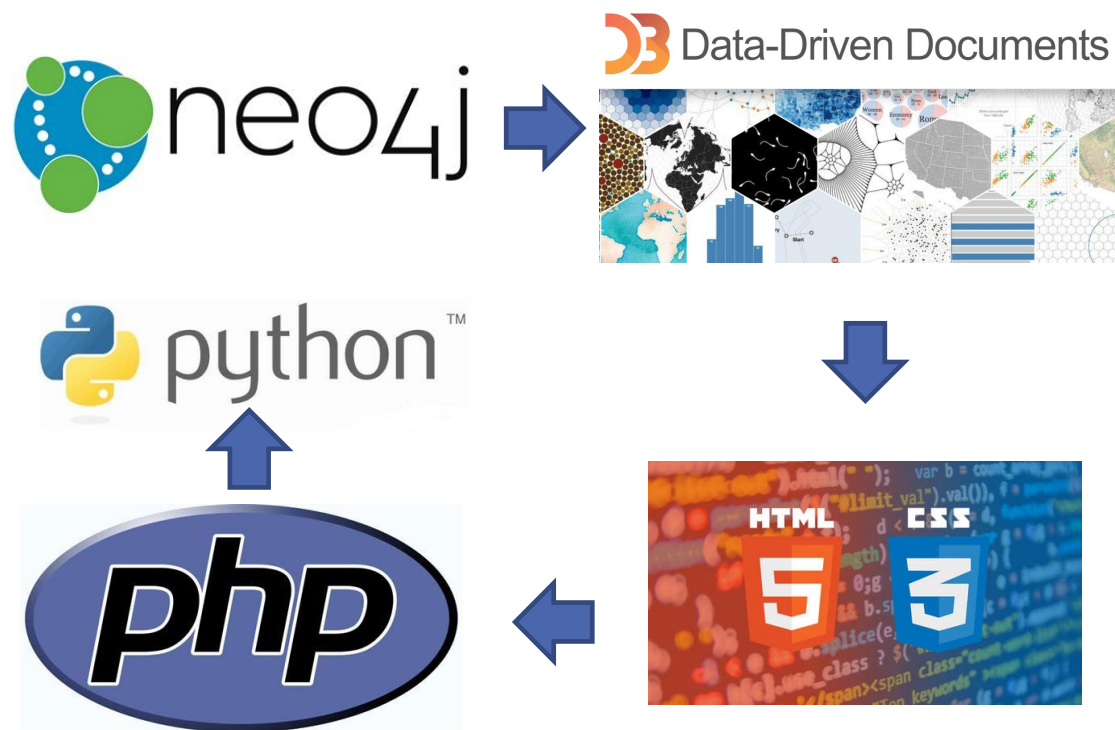
```
match (TomH:Person) where TomH.name='Tom Hanks'
WITH TomH as a
MATCH (a)-[:ACTED_IN]->(m)<-[:DIRECTED]-(d) RETURN a,m,d ;
```

网页前端

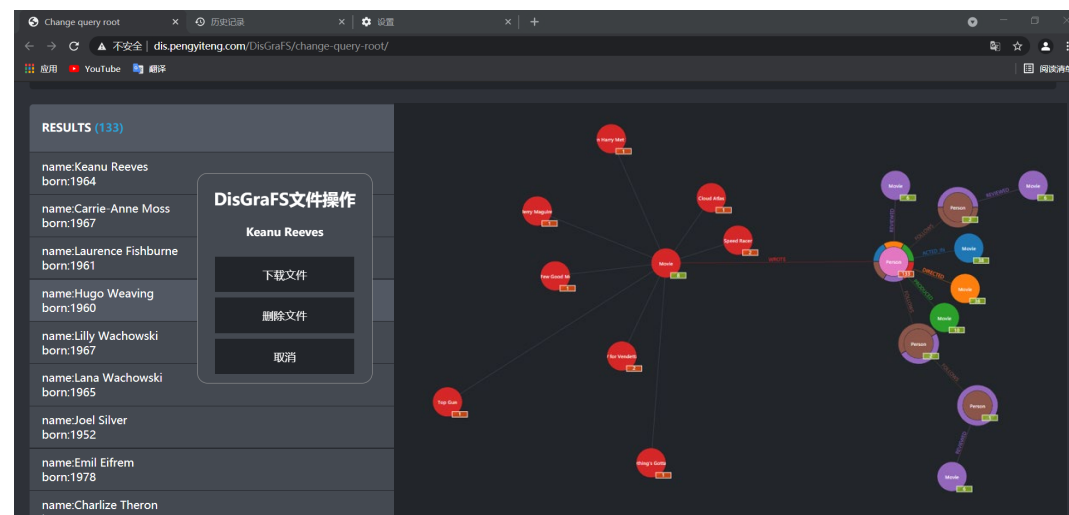


- 尽量减少客户端需求，全平台通用
- 更加用户友好
- 图形化显示图数据库
- 下载文件
- 删除文件
- 搜索文件
- 上传文件
- 它本身并不做任何实质性的工作，它的作用是将用户的信息发送到对应的服务程序，再由服务程序进行处理。

网页前端



- 使用neo4j官方提供的d3.js和pototo.js框架
- 结点实时显示
- 使用php与后台服务端进行通信
- 调用服务端提供的应用



文件语义识别



体育	57.26%
海滩	55.48%
滑冰	54.25%
海	48.54%
海洋	45.05%
沙子	41.38%
活动	39.78%
夏季	37.35%
跳跃	35.82%
生活方式	35.47%

LC-3 Assembler Report

Brief Introduction

An LC-3 Assembler implemented with C++. Converts `.asm` file to `.bin` opened with *LC-3 Simulator*.

Features

- Shows the corresponding lines in `.asm` file of every 16-bit word.
- Shows the address of every 16-bit word.
- Shows the symbol table.
- Show information of errors if the `.asm` file is wrongly written.
 - The assembler now supports the following kinds of errors:
 1. The out-of-range label/offset.
 2. The overflowed immediate.
 3. Non-existed label.
 4. Same-named label indicating multiple addresses.
 5. Missing `.ORIG/ .END`

('lc-3 assembler', 0.036)

('arguments', 0.034)

('errors', 0.032)

- 文本：Python库pke提取关键词
- 图像：使用imgga提供的tagging接口
- 语音：speech_recognition: CMUSphinx语音识别 + pke关键词提取
- 视频：使用tinyTag提取丰富的元数据



分布式计算集群：RAY

ray-project / ray

<> Code Issues 1.4k Pull requests 156 Actions Projects Security Insights

master 53 branches 43 tags Go to file Code

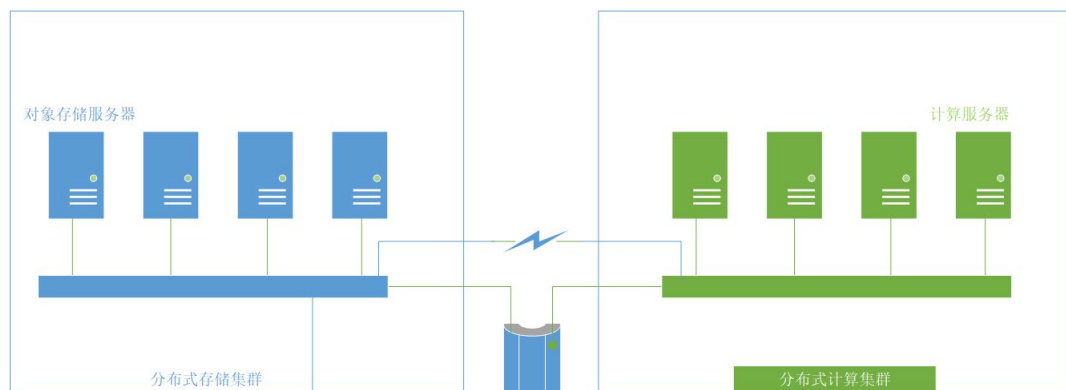
sven1977 [RLlib] Fix test_dependency_torch and fix custom logger support for R... 354c960 yesterday 7,895 commits

.buildkite	[CI] Run Flaky tests in macOS (#15487)	yesterday
.github	[Tune] Revert Pinning Tune Dependencies (#14059)	2 months ago
.gitpod	[dev] Enable gitpod (#15420)	4 days ago
bazel	Upgrading Redis to 6.0.10 in order to be functional on Apple silicon ...	2 months ago
benchmarks	Move scalability envelope back down to 250 nodes (#15381)	8 days ago
ci	[Buildkite] Use the build link for Travis Tracker (#15317)	10 days ago
c++	[C++ worker] Ray actor task for RAY_REMOTE (#15039)	13 days ago
dashboard	[Kubernetes] [Dashboard] Remove disk data from dashboard when run...	19 days ago
doc	[Core] Add `concurrent.futures.Future` wrapper for ObjectRef (#15425)	2 days ago

- 2017年12月开源
- 高性能分布式计算框架
- 架构完善，社区有强大生命力
- 性能、安全性、可扩展性有保障
- 接口简单易用

```
1 import ray
2 ray.init(redis_address="192.168.0.1:6379")
3
4 @ray.remote
5 def tagging_task():
6     pass #code here
7
8 return ray.get(tagging_task.remote())
```

开发路线



- 各个部分并行开发，各自完成模块测试
- 增量式集成测试





DisGraFS

Q&A