

data

config

系统初始化+调用接口

data buffer

传统块级重复数据删除模块

文件切块

指纹生成

重复块查找

数据管理

unique
block

数据类型鉴别模块

数据抽样

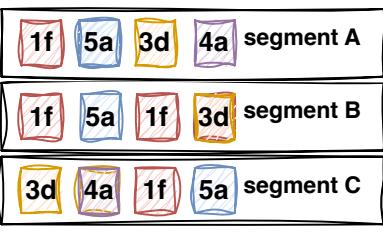
数据拆分

熵值分析

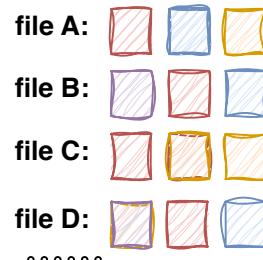
类型判别

数据管理模块

段管理器



文件结构缓存



索引缓存

块指纹缓存



细粒度特征缓存



模型数据

冗余识别

非模型数据

特征提取

冗余集中

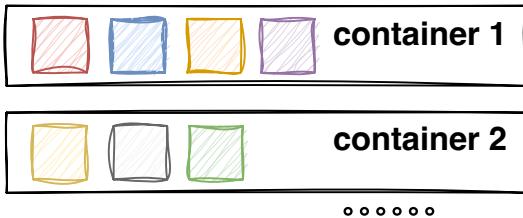
局部冗余去除

压缩冗余区

压缩差异区

持久化存储模块

容器持久化存储



文件结构存储



KV存储

块指纹索引

细粒度特征索引

compressed
data