LogCUA 设计文档

Owner： MOV Group

1 目标

1. 解决日志文件采集，采集文件特殊日志，以及文件夹日志，socks接口等。
2. 支持自定义的采集组件接口（采集特别的内容，比如块日志）

2 实现

1. 采集数据语言选择

采用Rust语言，原因无GC,并且效率高，语言本身是静态语言，跨平台等特性，占用内存极少。

1. 采集数据组件选择
   1. 采集数据包含文件夹采集，文件采集。
2. 更新版本
   1. 添加正则表达式处理数据
   2. 根据时间一次性加载数据

3 分析

1. 采用python + Deeplog进行数据的分析 ，生成日志模型，以及日志的清理，和特殊单词的保存，来计算词频，以及文本相似度，来确认服务是否存在异常。