代码量：2w万行

异常操作汇聚系统

|  |  |
| --- | --- |
| 模块 | 功能 |
| n个客户端 | 客户端后台运行，每个客户端给部署在一个节点上，各个节点给我的客户端程序上报异常操作； |
| 数据处理 | 客户端上报的异常操作信息加密之后发送给服务器，客户端有保活机制和掉线重连机制 |
| 服务器网络模块 | 使用muduo，epoll + reactor +多线程，处理网络连接和网络数据 |
| 服务器任务处理 | 使用线程池来处理数据包括：数据解密，数据json的解析 |
| 服务器任务队列 | 多线程处理完成数据后，将数据放到任务队列中，将数据存入不同的表中 |
| ftp服务 | 在网上github找的开源项目集成自己项目，密码登录，权限分配，指定开放的文件。与第三方交互，获取第三方需求后定时将数据以csv文件的格式放入指定文件目录中 |
| Qt管理前台 | 1. 服务器上有一个qt的控制界面，可以看每个模块上报的异常信息，管理ftp服务器上的文件，可以手动操作数据库（查询，去重操作），对子节点的监控（后期还会控制子节点的发送速度） |
|  |  |
|  |  |
|  |  |

亮点：

1. 多线程同步处理完善，软件运行很高效稳定
2. 测试操作清晰简洁，测试数据存储完整，有完善的测试流程

技术瓶颈：

1. 架构没有采用先进的框架，比如网络模块都是自己从头开始写的，导致并发没有很高，不能同时处理非常多的数据，后面全部换成了muduo，让代码简洁，用了很多muduo的组件，比如日志，定时器等等
2. 界面和数据运行了大量相同数据，一旦数据处理流程卡住，便无法解锁这部分数据，导致界面卡死，后来改成异步显示，不让界面出现假死
3. 数据库是国产数据库，没有好的封装，全部自己按照开发文档手动封装
4. 数据接收和存入数据库有延时，设计了消息缓存队列，可以防止丢弃消息

开发流程：

1. 客户提出需求，产品在需求网站上写需求
2. 组长根据需求分配任务给组员
3. 组员拿到需求后找设计出图
4. 组员在开发途中有不明确的需求点，找产品沟通补需求，如果需要补图找设计补
5. 完成需求后，提测
6. 组员改BUG，让测试重测（发版本不取决于是否全部BUG改完，取决表面的恶性BUG是否修复）

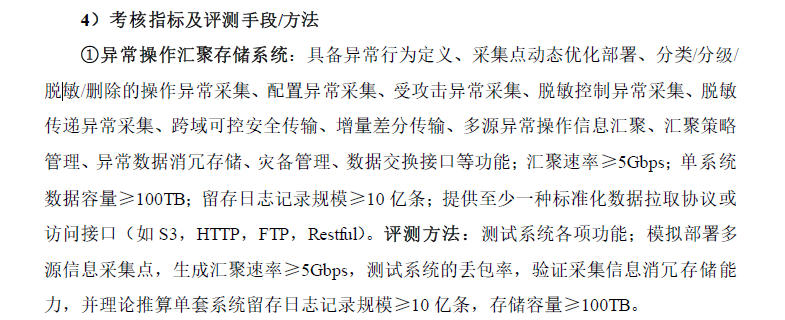
优化：

目前我做的工作中，优化的工作有：

1. 用muduo网络库提高代码稳定性和简洁性
2. 排查并优化数据处理多线程中缓慢的阶段，使临界区或锁快速解锁，防止数据持久化和网络io之间不匹配导致数据丢失

## 项目背景

1. **名称：**隐私数据的个人权益保障研究，国家重点研发计划，项目编号2021YFB0300
2. **牵头单位：**中国科学院信息工程研究所，项目负责人：李凤华（我导师）
3. **参与单位**：中国科学信息工程研究所，电子科技大学，西安电子科技大学，华中科技大学，中国电信，国家计算机与信息管理中心，中国网络空间研究院，中国电子技术标准研究院，京东城市数字科技
4. **项目预算**：2285万全部来自中央财政专项资金
5. **项目预期**:研发 16 种软件、搭建 1 种平台。软件预期达到的关键指标有：支持类别大于等于15 个、 场景≥50个、5000 万人以上个人敏感信息分类分级，准确率≥95%；5000 万人以上的信息拆分脱敏存储和信息重构，在线并发≥10万、在线用户≥500 万，单条个人敏感信息拆分和重构毫秒级；综合服务平台提供权益保障监管等服务，在线 并发≥10万、用户≥1000 万、信息记录≥10 亿条。
6. 我负责的子课题：异常操作汇聚存储系统：



## 1 应用场景

## 

## 2 产品形态

1. 有n个客户端，每个客户端给部署在一个节点上，各个节点给我的客户端程序上报异常操作；
2. 我将客户端上报的异常操作信息加密之后发送给，服务器
3. 服务器对各个节点的信息解密后，解析各个不同模块的数据信息，存入数据库（国产达梦数据库）
4. 和个人信息分析系统交互，按照他们的要求，定时将他们需要的异常信息，以csv文件的格式放置到ftp服务器上
5. 服务器上有一个qt的控制界面，可以看每个模块上报的异常信息，管理ftp服务器上的文件，可以手动操作数据库（查询，去重操作），对子节点的监控（保活机制，子节点定时重连）
6. 数据库定时备份功能

## 3 主要模块

就是上面的几个功能。

网络通信模块用的muduo网络库

任务处理用了线程池

ftp服务器是我自己找的开源项目集成进去的

数据库模块用的国产的达梦数据库（插入，去重，备份，查找，删除）

日志模块用的muduo里面的日志模块

加密模块用的第三方的库

网络数据基本用的json传递，用的jsoncpp解析网络数据。

## 4 主要技术

Muduo网络库的源码基本都看了一遍：

计时器的实现，

Epoll+reactor模型+多线程网络io

多线程日志模块设计，

多线程编程

Qt使用，利用信号槽的界面异步显示

数据库的使用，利用给的接口实现了，数据定时备份和基础的增删改查