

20220916 电子秤下秤故障

故障报障时间：2022-09-16 07:09:00

故障实际发生时间：2022-09-16 07:09:00

故障开始处理时间：2022-09-16 07:12:00

故障恢复时间：2022-09-16 07:44:00

故障影响范围： 五家门店传秤报障, 反应故障期间任务堆积, 无法快速下秤, 恢复后确认正常

【新增】故障恢复后数据修复时间：无需修复, 门店重新触发下秤即可

故障问题描述：20220916 电子秤下秤故障

故障根因：早高峰时期, 门店下秤任务量大, 导致新的传秤链路未能快速处理从而导致下秤任务堆积.

故障处理与分析过程：

2022-09-16 07:09:00 灭火群报障

2022-09-16 07:12:00 研发上线介入处理

2022-09-16 07:20:00 研发确认需要进行配置回滚

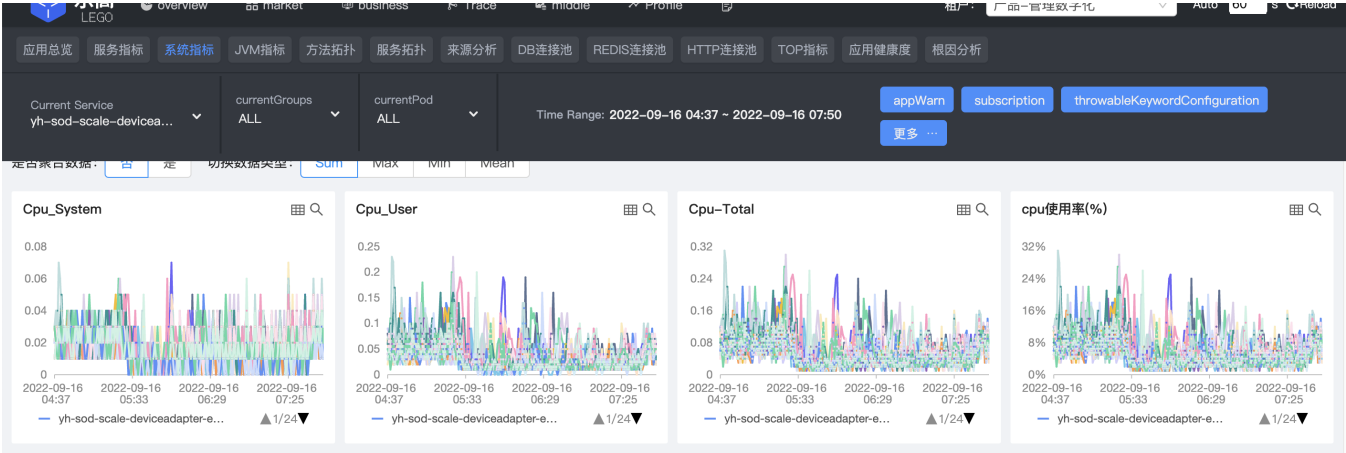
2022-09-16 07:21:00 研发发送回滚范围到灭火群, 运维进行回滚

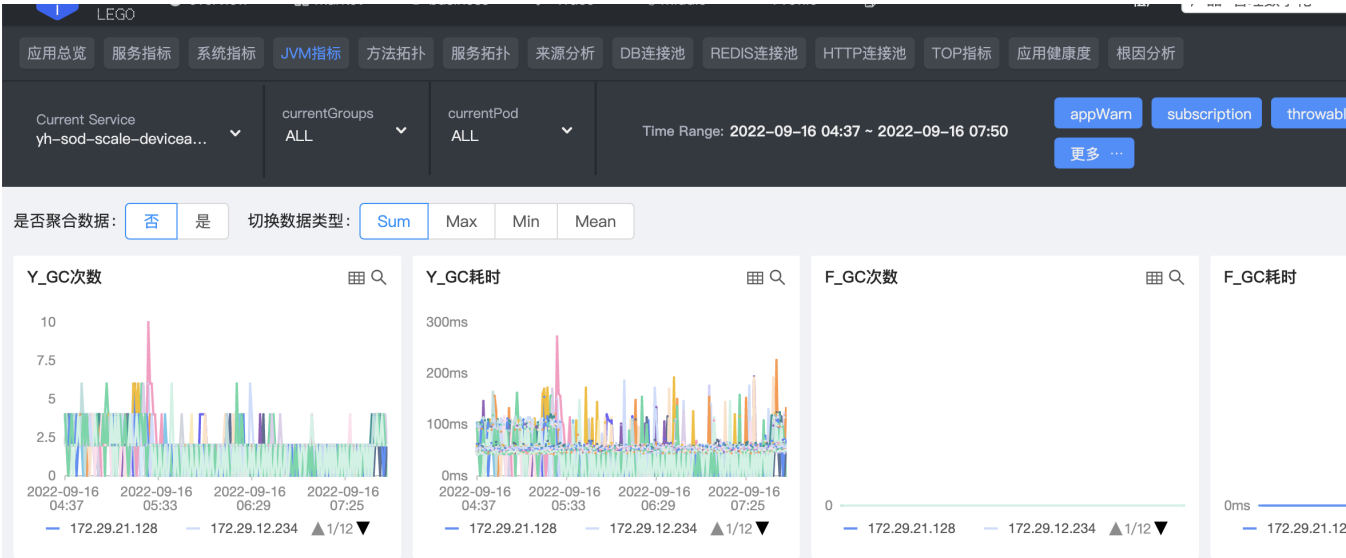
2022-09-16 07:35:00 发现回滚未生效

2022-09-16 07:43:00 研发重新提供配置修改, 运维操作进行修改

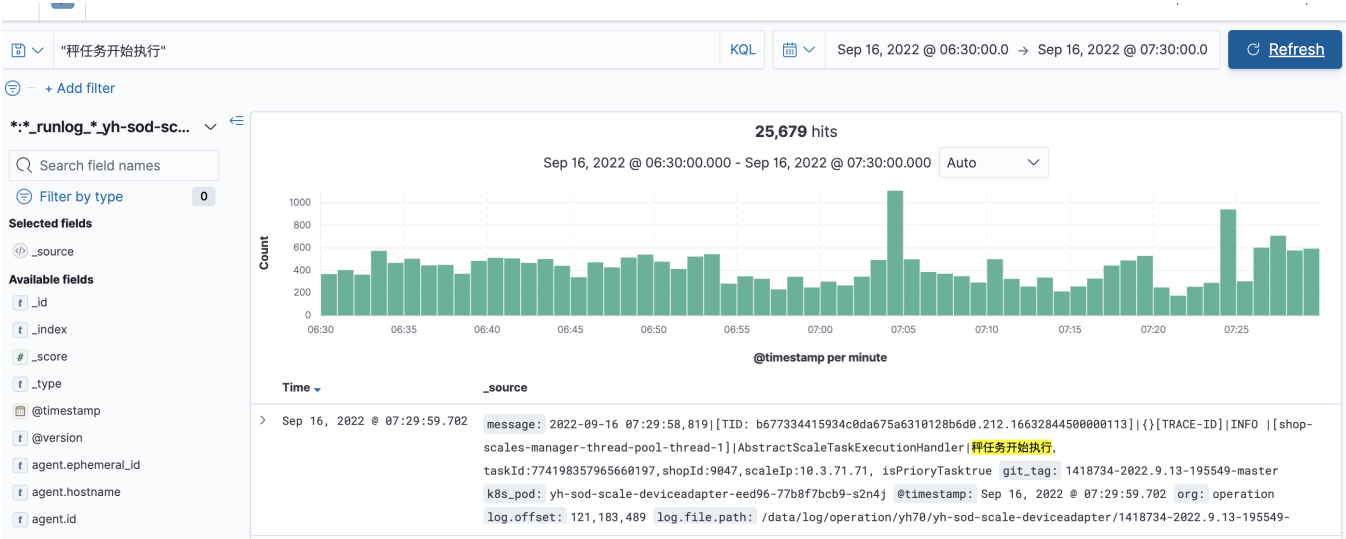
2022-09-16 07:44:00 配置修改完成, 故障恢复

事情发生的时间点, 服务器运行较稳定, 未超过机器负载:





下秤任务正在执行:



根因分析:

新的传秤链路上线后,灰度了20多家门店,上线几周都运行稳定,经过压测(90家门店)之后,判断系统资源能够支持目前的所有门店,于是09.15跟各方沟通进行了全量切到新链路,当晚10:30之后切换,切换后,运行稳定,任务执行正常.

但是没有考虑到早高峰的任务大量下发的场景(一个是任务量大,一个是频繁变更),新链路每次拉取一个门店列表,进行顺序任务查找门店是否有任务,导致处在队列前面的门店,因为其一直有任务,一直占用着系统资源,导致队列后面的门店,被拉起的频率过低.

因此,问题主要分成以下几个主要原因:

1. 每次定时任务拉起后,按照顺序查找门店任务的方式,导致在高峰期前部门店占用过多资源,后面的门店拉起频率过低,需要进行逻辑优化
2. 门店的秤如果ping不通,并没有将任务记录失败次数,导致这种无效秤的任务被频繁拉起,从而导致前面的门店老被重复拉起

【新增】故障恢复后数据修复过程:

无需修复,门店重新触发下秤即可

改进措施action:

- 1、优化传秤门店任务扫描逻辑,使系统的资源对更加公平的分配给各个门店刘钦涛#81116320开发完成10.1之前,上线10.15之前
- 2、门店秤如果ping不通,直接将该秤任务直接设置成失败,不再重试,防止ping不通的秤一直被拉起,占用系统资源刘钦涛#81116320 开发完成10.1之前,上线10.15之前
- 3、优化对传秤数据的监控逻辑,更有效的发现数据堆积情况刘钦涛#81116320 开发完成10.1之前,上线10.15之前
- 4、灰度切换每次量少一些,降低切换引起的影响范围刘钦涛#81116320 开发完成10.1之前,上线10.15之前
- 5、阿波罗回滚逻辑优化——@张海

【新增】4、故障异常数据修复action(下次怎么可以快速)

故障责任人: 刘钦涛

故障等级: 线上BUG