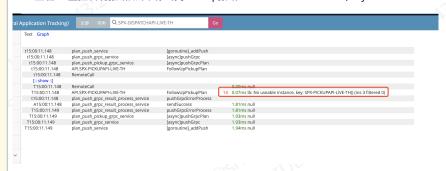
2022.04.02-SPX- TH地区Pickup plan下发报错"lb: No useable instance, key: SPX-PICKUPAPI-LIVE-TH() (ins 3 filtered 0)"

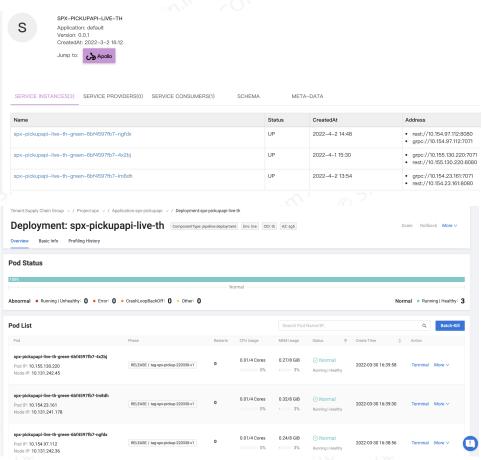
故障 标题	TH地区Pickup plan下发报错"lb: No useable instance, key: SPX-PICKUPAPI-LIVE-TH() (ins 3 filtered 0)"			
故障 等级				
故障时间	发生时间	2022.04.01 下午2点11分		
	发现时间	2022.04.02 下午2点52分		
	恢复时间	2022.04.02 下午3点48分		
	故障历时	25小时37分		
故障 处理 人	xingdong.yin、shaohua.xia、lingfei.wang、yanlu.li、zhiming.hu、yuanjing.lin			
责任 主体	13:126			
故障 报告 来源	业务反馈	17:10		
故障 类别	5/10/6cc			
故障 描述	调度中心下发Pickup plan存在1/3概率失败,报错"lb: No useable instance, key: SPX-PICKUPAPI-LIVE-TH() (ins 3 filtered 0)"			
影响评估				

处理 过程

- 1. 4月2号早上11点dongyang找yuanjing反馈分拣机分拣报错
- 2. 下午2点左右yuanjing排查确定是查询pickup报错没有pickup order,找shaohua确定,反馈至hongtao
- 3. 下午2点52分hongtao反馈有订单下发plan失败, xingdong和lingfei 介入排查
- 4. 查看cat监控发现报错如下图,请求Pickup报错"lb: No useable instance, key: SPX-PICKUPAPI-LIVE-TH() (ins 3 filtered 0)"



3. 查看Pickup注册中心和容器情况,发现容器没有重启,但是注册中心的服务注册时间有变化



- 4. 排查确定Pickup新重构的代码导致容器内应用进程挂掉后被重启,注册中心有注册成功的实例 , 按理可以下发,不会有报错
- 5. Pickup关闭新重构内容开关,逻辑走老链路
- 5. 与中间件同学确定此种情况会触发注册中心客户端的一个bug: https://git.garena.com/shopee/bg-logistics/go/chassis/-/issues/23
- 6. 与中间件同学确定临时处理方案,重启调度中心容器,触发重新获取注册中心信息,使用plan控制台重推机制重推下发失败的plan

原因 分析

- 1. Pickup服务崩溃进程重启但容器正常触发注册中心的一个bug,导致调度中心下发plan时存在1/3概率失败
- 2. 调度中心告警配置存在考虑不足的情况,导致Plan下发失败没有第一时间发现
- 3. 调度中心告警配置有误,导致没有第一时间发现问题

改进 方案

- Pickup针对重构服务做一轮修复之后再上线打开开关,切流过程需要灰度,避免一次全量引发其他问题
- 调度中心与中间件团队对齐chassis依赖组件的版本做更新,此项需要SPX全部golang应用自查-- zhuohuang.yu@shopee.com
- 新增应用监控体系完善,针对重要接口及时告警-- xingdong.yin@shopee.com lingfei.wang@shopee.com 待同步进度和deadline,后续辛苦

贴一下Jira task

⚠ SPXFM-29018 - Jira 问题不存在或者您没有权限查

, deadline: 4月13号go-live

后续待跟进事项

Pickup plan下发失败后重试推演

调度中心针对下发失败的plan在36小时之内会不断进行推送,推送机制:前四次按配置推送([4]uint32 $\{60, 2*60, 4*60, 10*60\}$),后面每45分钟推送一次,直至到达36小时

根据cat监控,每次推送有1/3下发失败概率进行统计

SQL: select count(DISTINCT(target_number)) FROM spx_plan_dashboard_tab WHERE plan_create_time >= 1648793880 and plan_create_time <= 1648861200 and push_status = 3 and plan_update_time >= 1648861200 and plan_status = 1 LIMIT 1000 ###: 341 10

No errors, Total: 1

count(DISTINCT(target_number))

28119

推送次数	plan下发延迟时间	失败概率	备注
1 0	1min	1/3	om,
2	1min + 2min = 3min	1/9	766. LIA
3	3min + 4min = 7min	1/27	HOL DEL
4	7min + 10min = 17min	1/81	- NETT
5	17min + 45min = 62min = 1h2min	1/243	至此开始,每45分钟推送一次
6	1h47min	1/729	
7	2h32min	1/2187	
8	3h17min	1/6561	
9	4h2min	1/19683	
10	4h47min	1/59049	
11	5h32min	1/177147	

看。

Chassis框架依赖版本监控

针对chassis框架使用的依赖组件做一个监控统计,识别使用危险版本的应用告警更改

新增golang应用监控告警完善

针对新增golang应用监控和告警做一个统一的规范,查缺补漏

监控演练