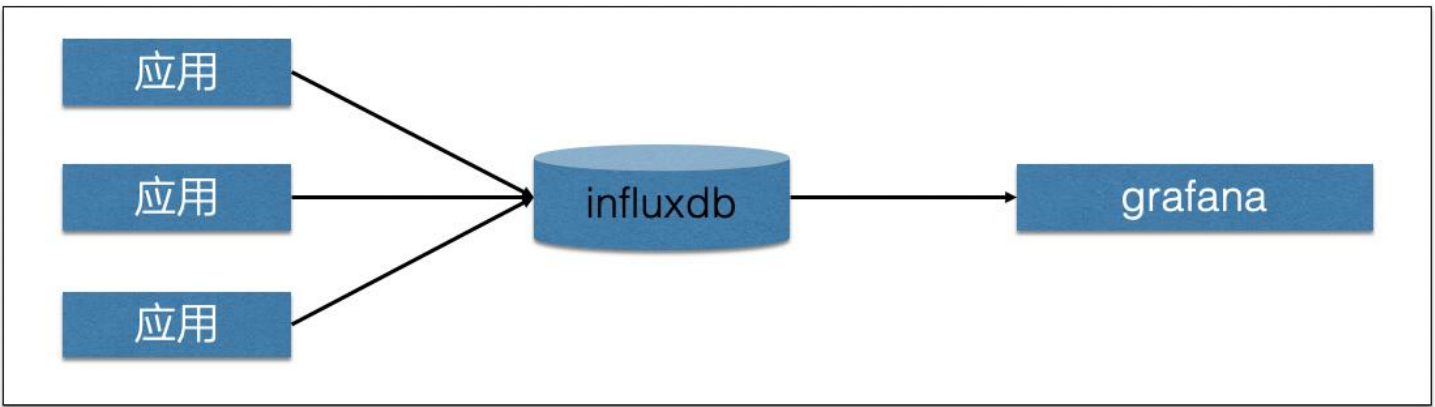


监控的实战体验

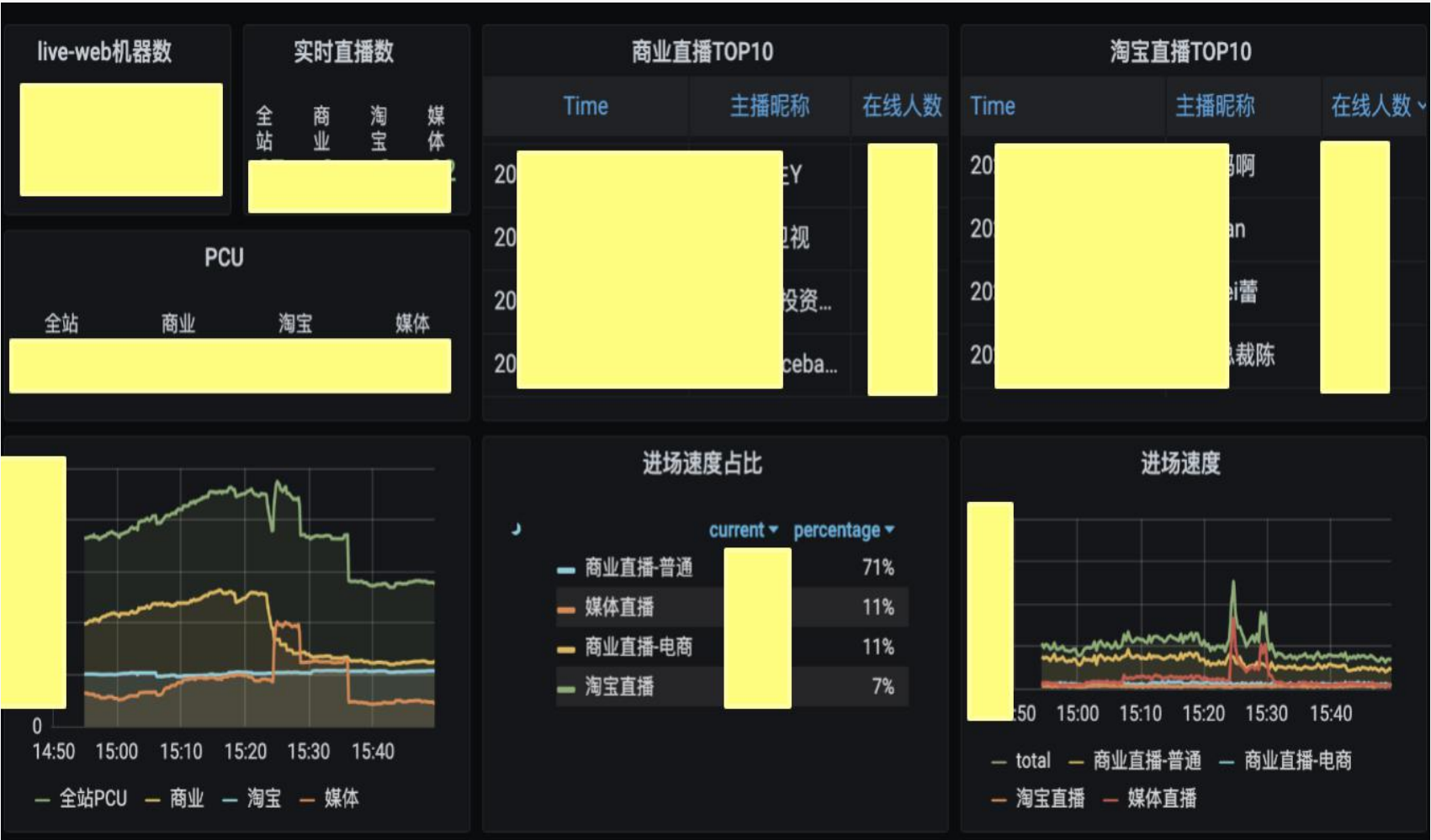
监控专题之监控实战体验

- ① 常用的监控（开源）
- ② 体验下各类监控（现有监控快速浏览）
- ③ 体验监控在各类活动的保障
- ④ 体验监控在故障预告及发现的价值

微博电商业务监控系统



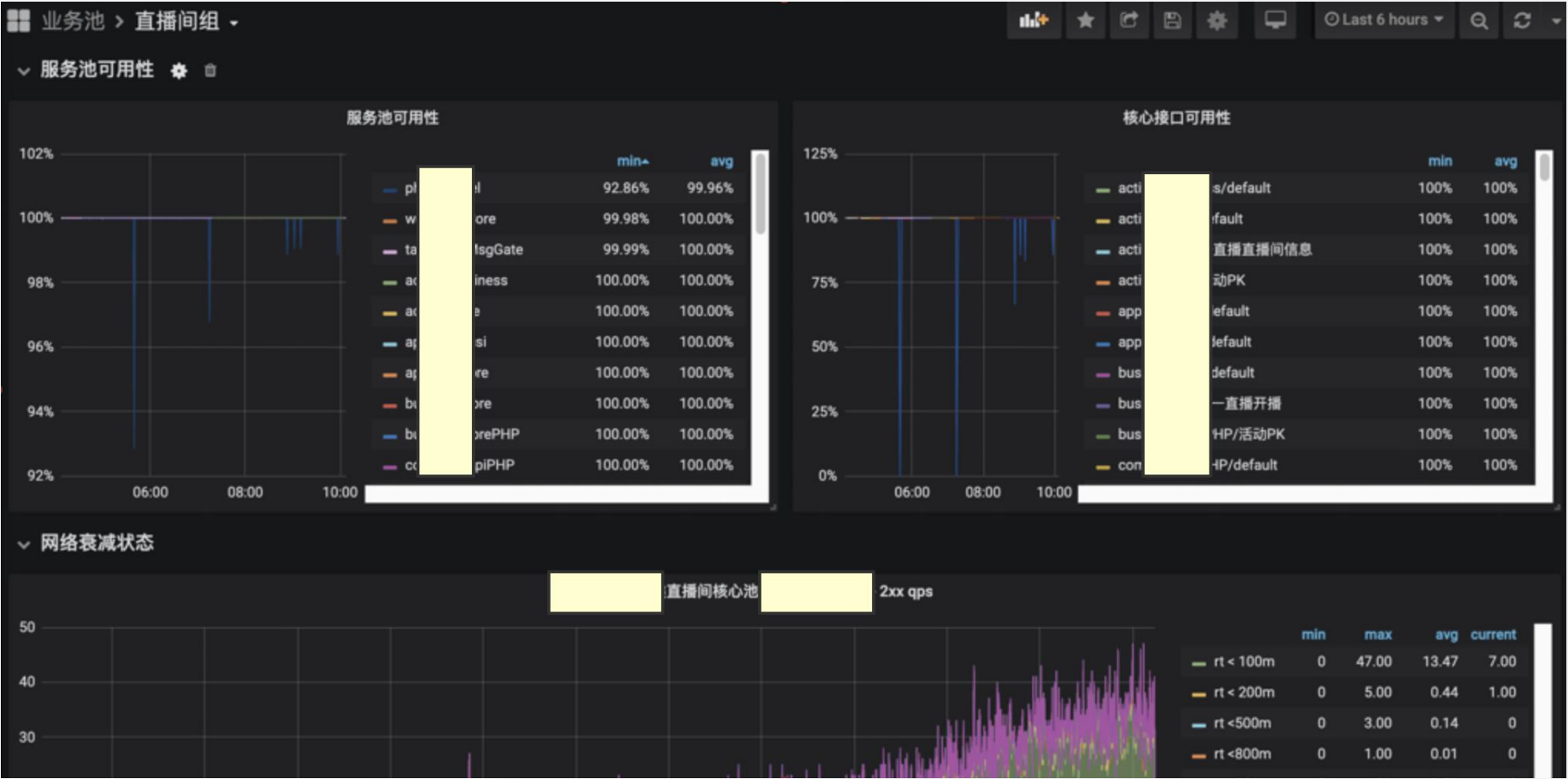
直播业务大盘



直播业务技术大盘



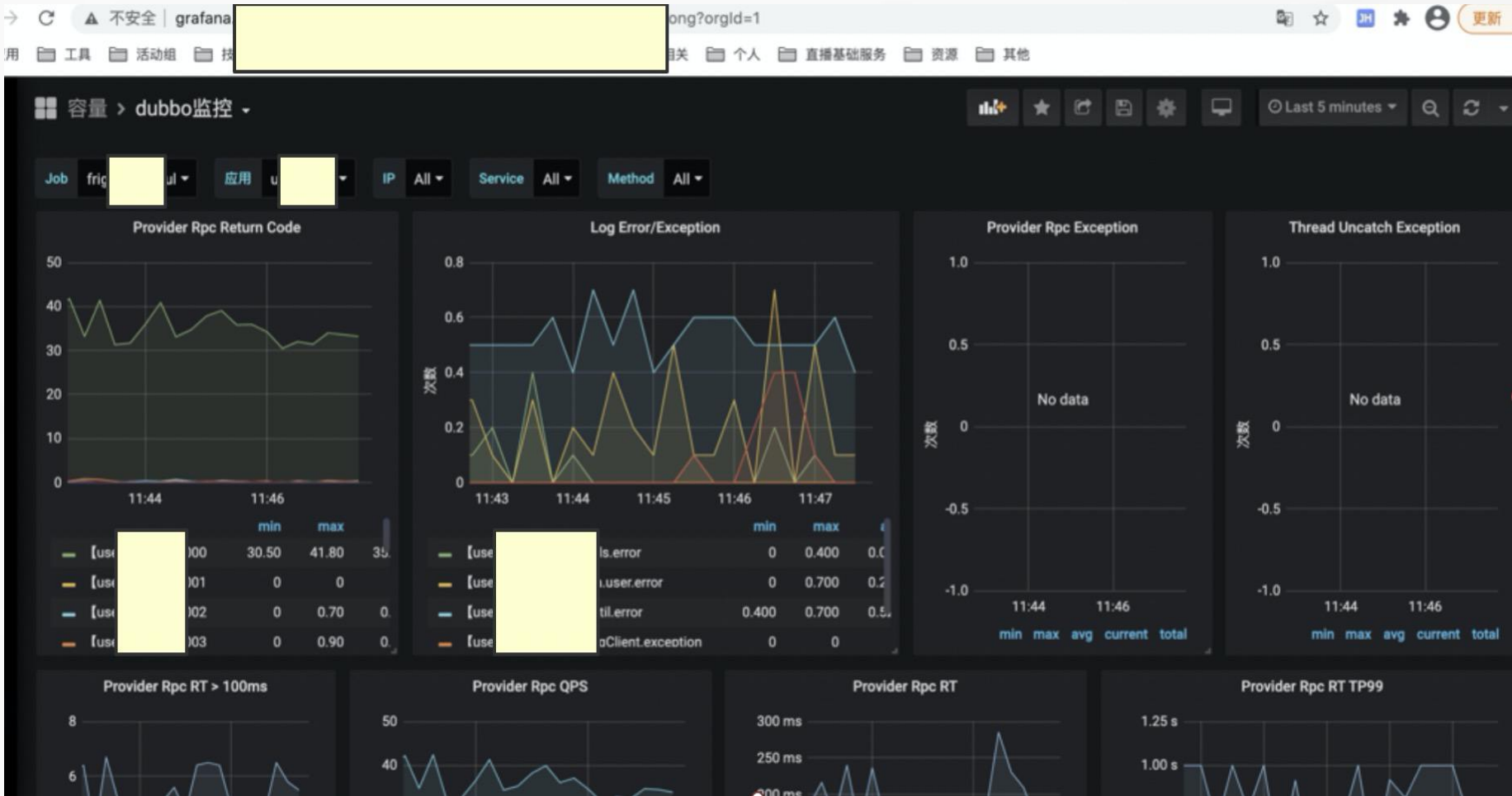
直播业务某个组整体监控



直播业务某个组的某个服务池基础监控



四层监控

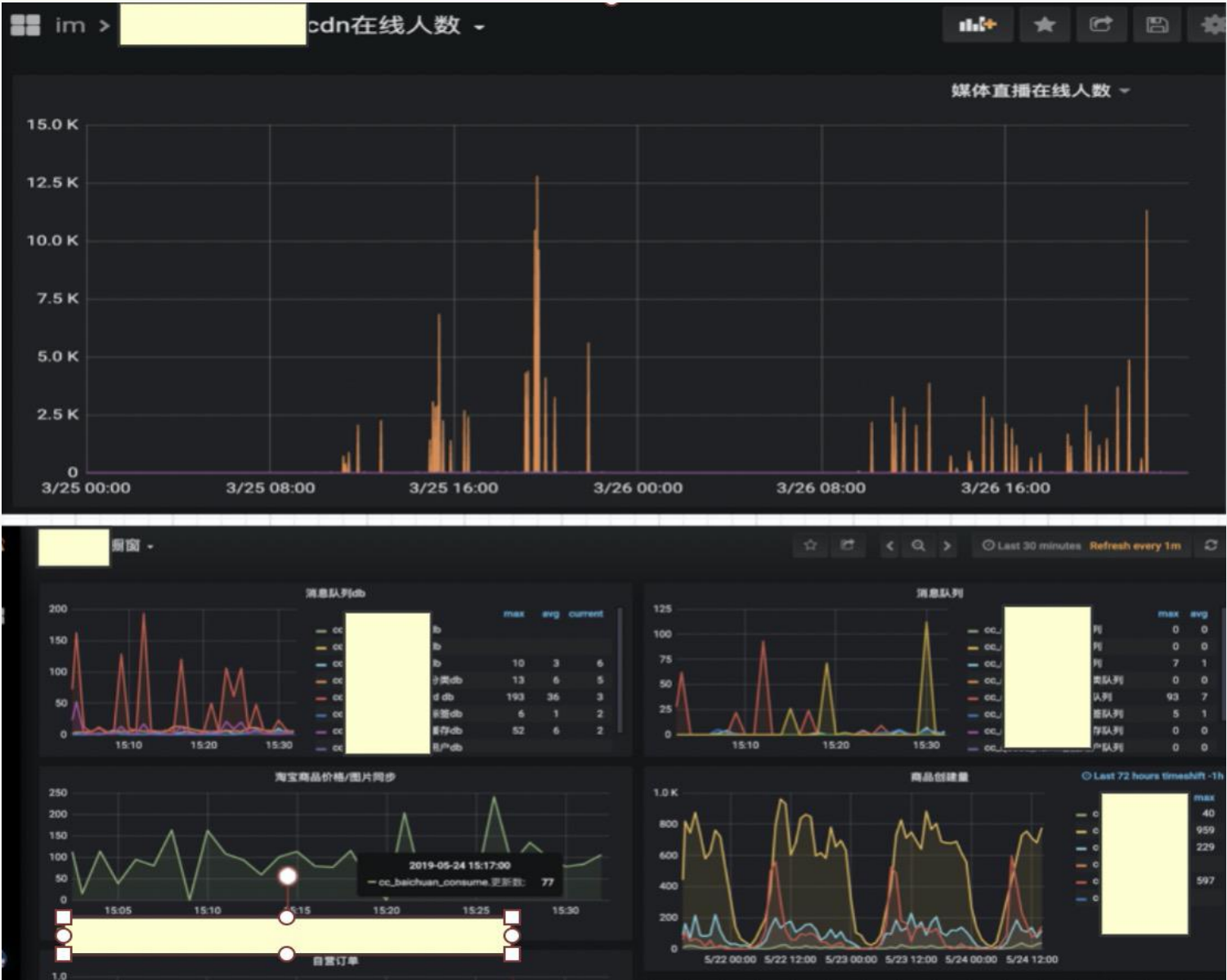


七层监控



增加了注解的任意方法监控

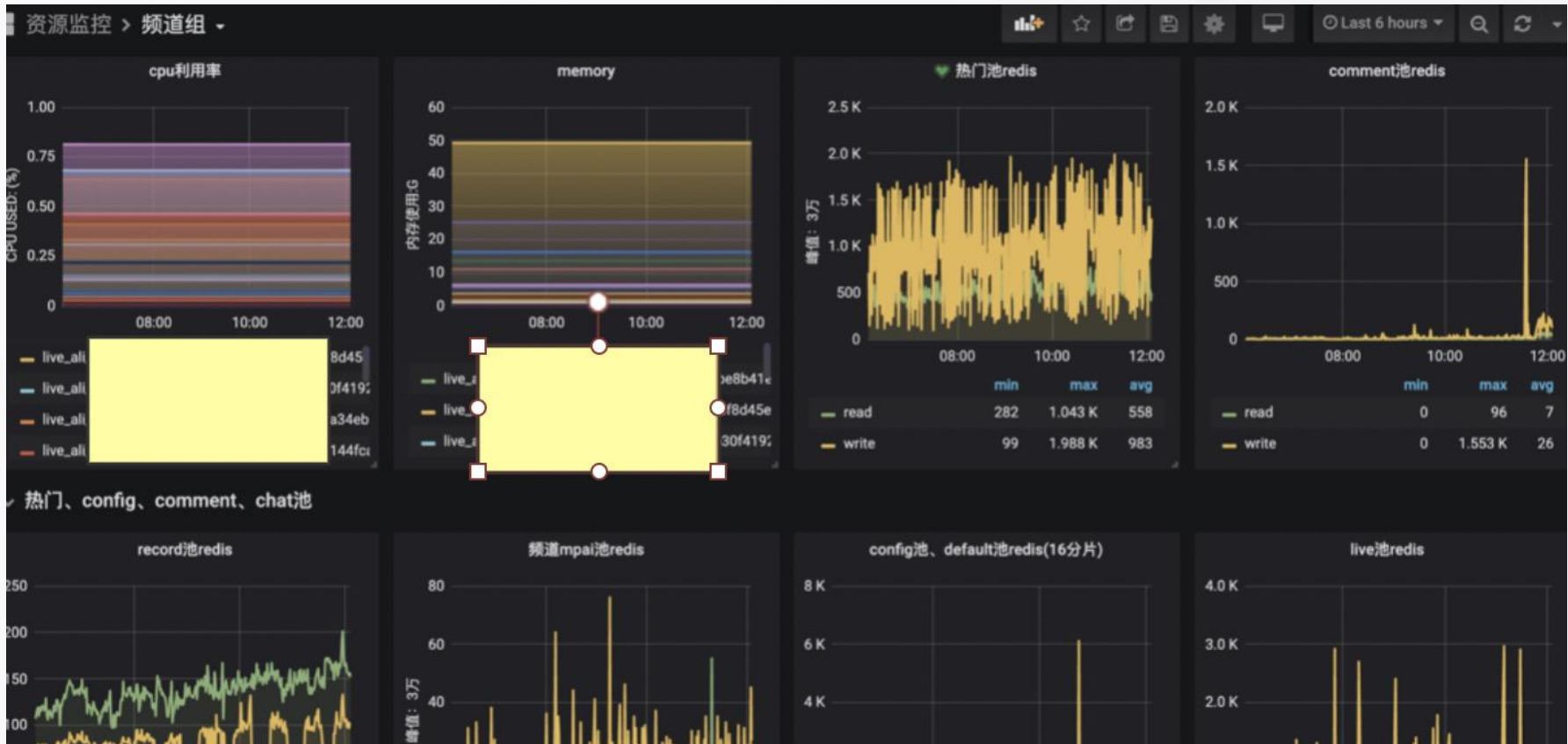




资源监控-mysql



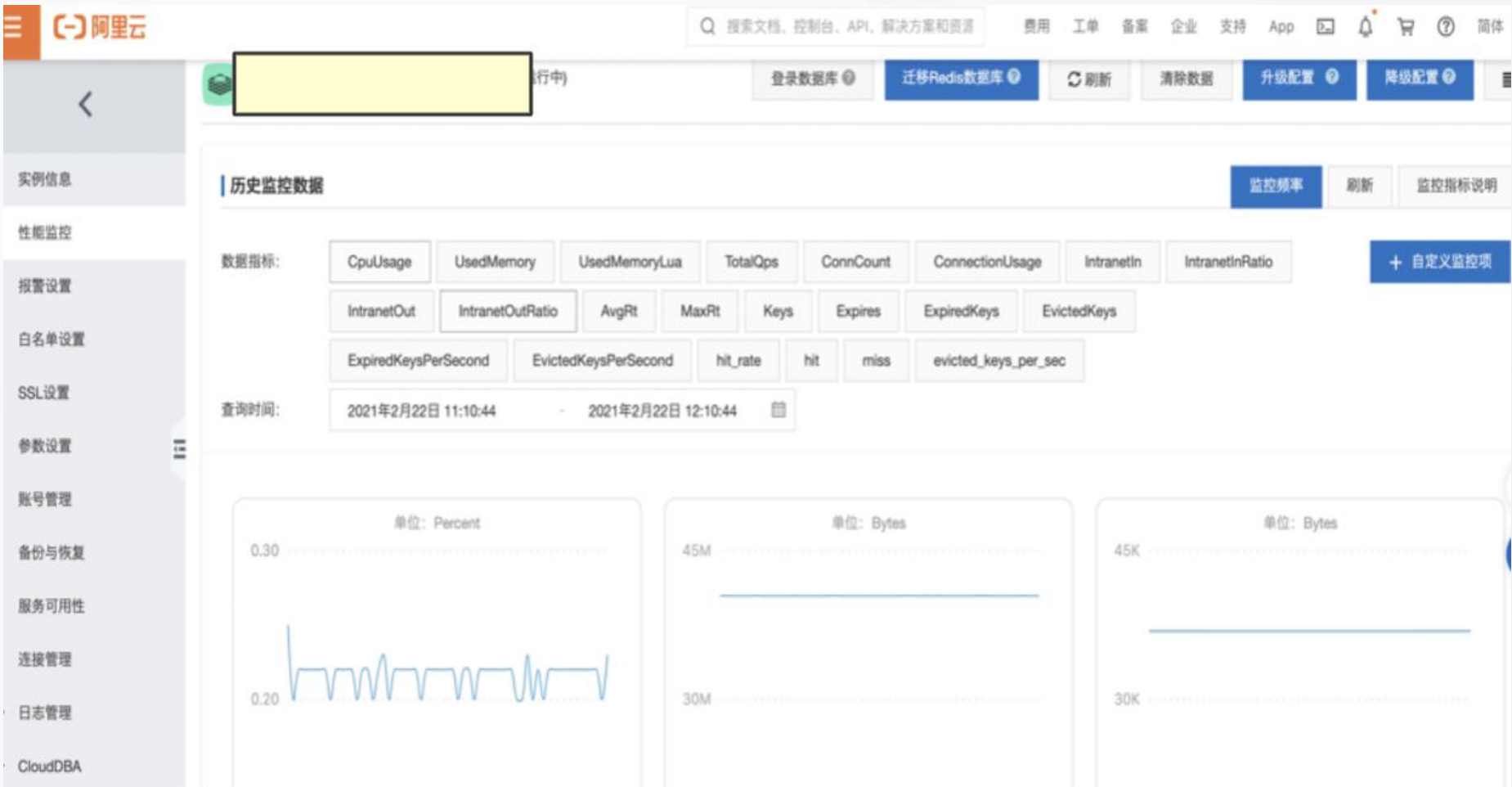
资源监控-redis



资源监控-rds

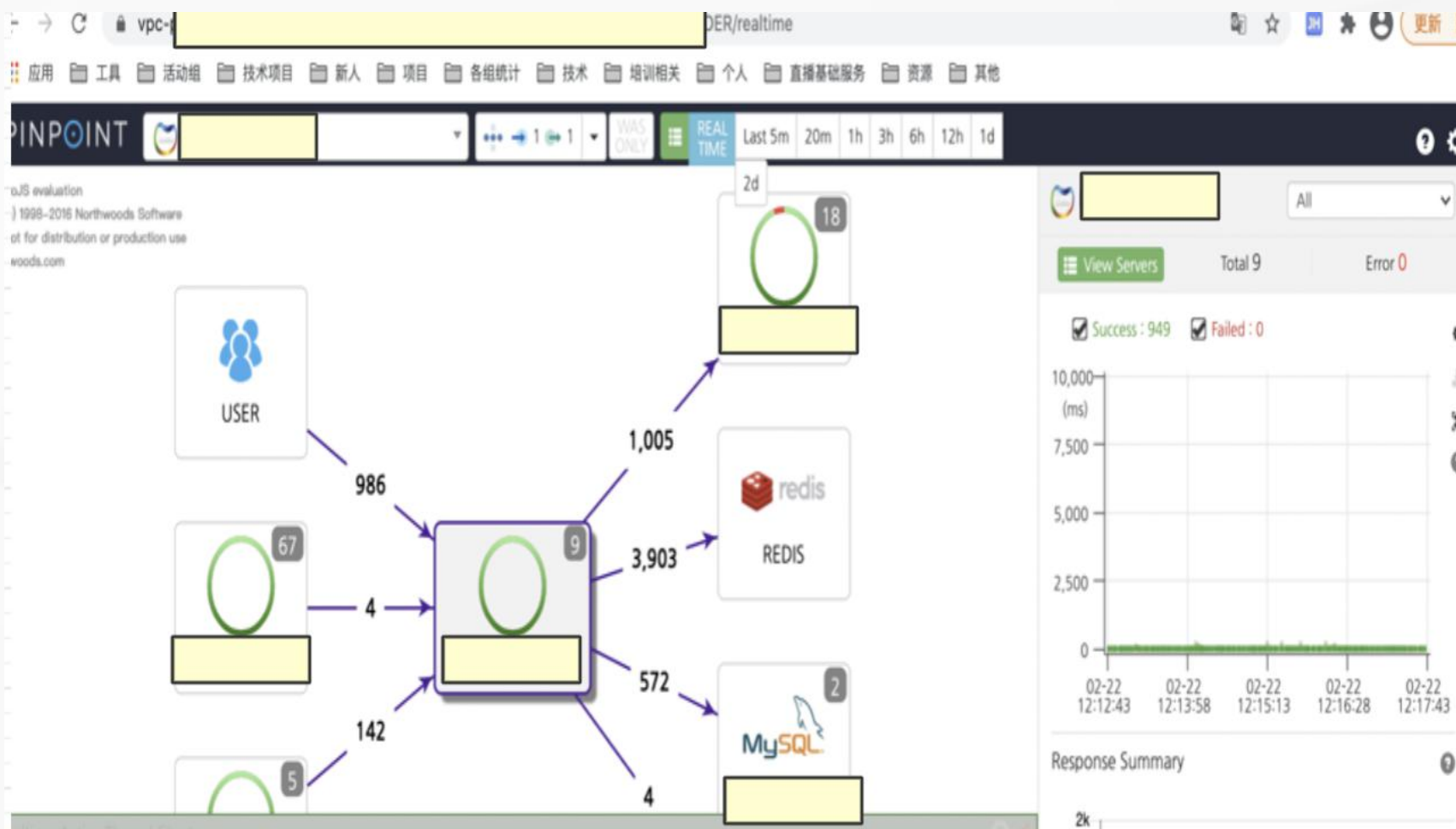


资源监控-redis

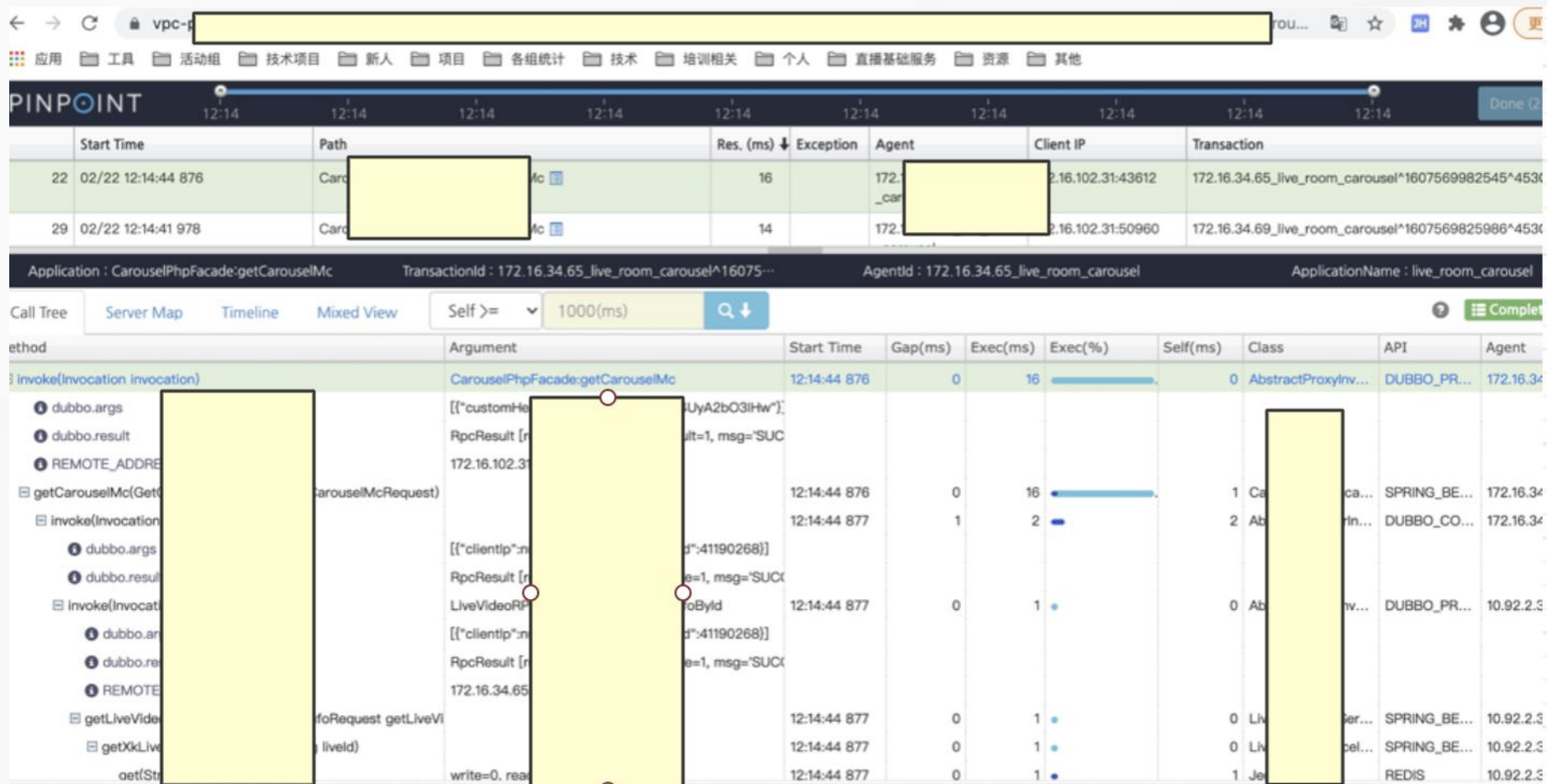


常用的监控（开源）

链路跟踪-pinpoint



链路跟踪-pinpoint



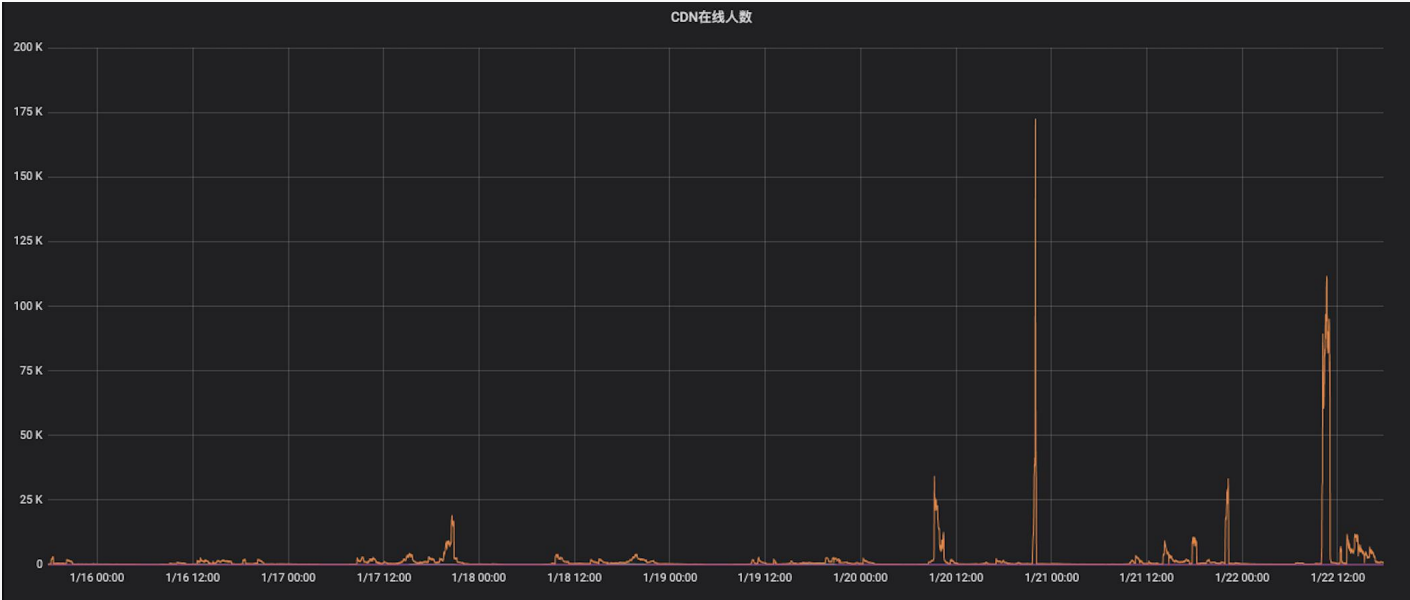
2	9范丞丞直播
2	8迪丽热巴直播
2	9直播最大V活动保障
2	8蔡徐坤直播
2	8黄明浩直播
2	9白宇直播
2	1 王俊凯
2	1王俊凯直播
2	8 心动餐厅
2	1 朱一龙直播
2	朱一龙直播
2	6吴亦凡
2	9淘宝双11晚会直播
2	6 央视带货直播
2	相信未来义演
2	范丞丞直播
2	秋收活动
1	邮件：2019-12-08 心动餐厅
1	保障一级模版
1	保障三级模版
1	保障二级模版
1	常规操作
1	直播间预案
1	演唱会
1	保障计划
1	活动保障计划邮件
1	活动媒体直播保障预案
1	

值班安排

团队	值班人员	电话
礼物	吴	13187
活动	常	85213
活动	陆	74860
活动	王	11191
频道	孙	10992
频道	万	63554
频道	佟	11672

保障计划				
模块	扩容时间	缩容时间	扩容资源数量	扩容后容量
直播前端机	12.30 10:00	12.30 18:00	250	可支持13WQPS
直播前端机	12.30 22:00	12.30 23:00	175	可支持22WQPS

终南山谈疫情



房间名称	主播Uid	播放数	在线数	操作
【1月26日20:00中国消防 知识专场】	我是答题王 (uid=6541899632)	2,015,882 (基本值:145,225)	2,015,882 (基本值:30,280, 连接数:30,796)	详情 消息 直播
欧巡赛-迪拜沙漠精英赛决赛轮	新浪高尔夫 (uid=149799611)	215,751 (基本值:19,269)	215,751 (基本值:997, 连接数:1,040)	详情 消息 直播

直播间在线人数

	2019.02.08 19:30 ~ 20:30(峰值)
单个直播间	26w
直播间总量	33.4w

普通直播基础保障

- 系统自动降级

根据qps或者cpu负载等进行系统自动降级

- 人工直播间降级

当直播正在观看人数到达一定量时，就只能观看直播，而不展现直播间内功能比如评论送礼等

- 直播前扩容

大明星到来之前进行扩容

- 自动扩缩容

当系统承受不住（目前使用cpu利用率），进行自动扩容，但是需要时间。

- 保障预案

当出现其他故障，提前做好预案保护。

体验监控在各类活动的保障

预案

模块	操作时间	操作恢复时间	功能调整内容
Profile页资料卡、成就系统拆分接口	开播前10分钟	活动结束后人工恢复	降级后profile页资料卡礼物贡献榜、真爱粉丝团、主播成就模块不展示
搜索下拉提示接口	开播前10分钟	活动结束后人工恢复	降级后搜索下拉不显示
关注页为你推荐及feed列表	开播前10分钟	活动结束后人工恢复	降级后feed不显示短视频,为你推荐不展示短视频封面
直播间资料卡	在线人数达到6w	活动结束后人工恢复	降级后直播间资料卡展示只用户昵称、头像、id
热门降级	开播前10分钟	活动结束后人工恢复	降级后热门不显示兴趣召回、关注主播不降权

第三方厂商预案——主机位阿里云流转推服务异常

流监控

执行预案【主机位阿里云流转推服务异常】

运营后台手动操作

观看出现一段时间黑屏；用户需重新进入直播间

张健、陈俊龙、佟堃、梁忠文、邓宇

第三方厂商预案——阿里云CDN异常

流监控

执行预案【阿里云CDN异常】

运营后台手动操作

无法观看，用户需重新进入直播间

张健、陈俊龙、佟堃、梁忠文、邓宇

第三方厂商预案——服务器全部故障

流监控

执行【服务器全部故障】

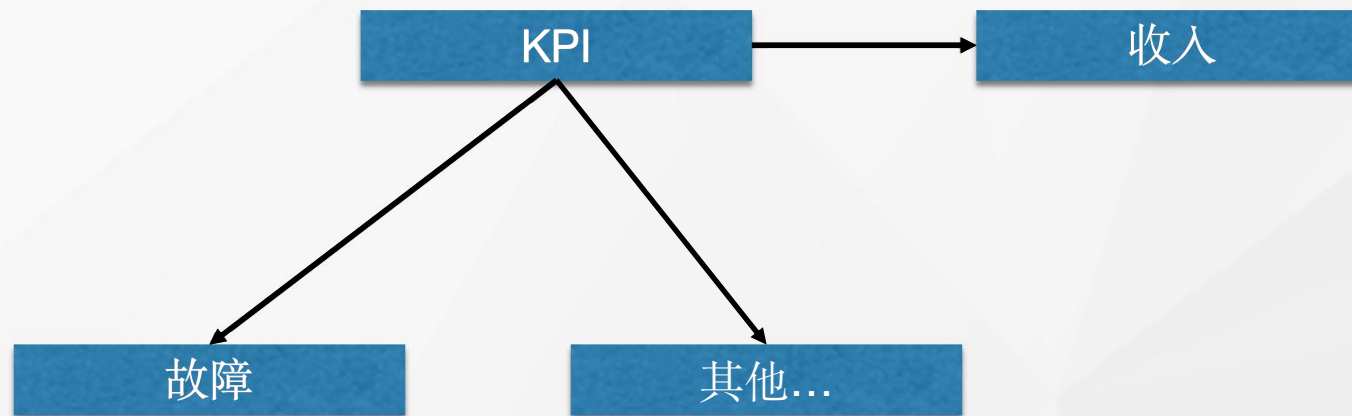
运营后台手动操作

用户进入降级h5页面，出现一段时间黑屏

张健、陈俊龙、佟堃、梁忠文、邓宇

体验监控在故障预告及发现的价值

故障-大公司质量管控的核心手段



故障决定收入

故障-故障等级

A级：85分以上		
B级：71-84分		
C级：61-70分		
D级：41-60分		
E级：40分及以下		
权重值		
30%	1、影响功能	
20%	2、影响服务时长	
15%	3、影响用户范围	
15%	4、用户投诉级别	
10%	5、影响服务程度	
10%	6、影响用户时段	
影响服务时长	权重20%	级别分值
1级	91分钟及以上	100
2级	51分钟-90分钟	75
3级	21-50分钟	50
4级	20分钟之内	25
影响用户范围	权重15%	级别分值
1级	影响10%及以上的用户	100
2级	影响5%以上-10%的用户	80
3级	影响1%以上-5%的用户	60
4级	影响0.1%-1%的用户	30
5级	影响0.1%以下的用户	20

级别	功能	说明
一级	首页	热门列表
	开播	
	观看	
二级	评论	
	礼物弹幕	
	直播间二级	pk匹配功能
		头等舱（信息卡展示头像昵称）
		主播信息（信息卡展示头像昵称）

- 故障等级，A-E级
- 故障分数组成
- 功能等级划分

体验监控在故障预告及发现的价值

历史故障汇总（2018、2019、2020年）

2018	自六月份开始共42个故障（自2019年开始正视故障，进行故障评级并列入kpi）			
2019 上半年	级别： 原因：容量	级别： 原因：资源评估不足	级别： 原因：代码bug	级别： 原因：人为操作qconf有误
	032 朱一龙 法加载	032 IM评论	032 小福袋被	022 点送礼
2019 下半年	级别： 原因：容量	级别： 原因：突发流量	级别： 原因：代码bug	级别： 原因：容量不足
	090 媒体直	101 媒体直播	101 随机PK	110 热门
2020 全年	级别： 原因：自身架构+第 三方	级别： 原因：DCP扩容 请求超时	级别： 原因：代码	级别： 原因：运营误操作+历史 问题数据不一致
	040 央视带货全	04 媒体直	041 榜单	06 福袋无法
				0803 PK



代码-故障策略-质量

评审

重要项目技术评审：重大项目启动后，在，合理排期中务必增加设计评审，降低设计不合理造成的故障。提前三天通知

重要项目CR：上线2天前进行CR。提前三天通知

KPI绑定：架构师绩效绑定，评审保证质量。近期全部评审，根据目前水平进行策略调整。

目前状态：已经执行半个月，由于新人较多暂时所有项目走评审，之后根据人员水平进行调整。有邮件CR目前通过率是80%

Bug统计及可用性统计

bug统计：每个项目增加千行代码bug数统计。进而提高开发人员对代码质量的重视。

每周周报会有一级二级服务可用性统计
环19年bug统计：

直播间：150

应用服务：111

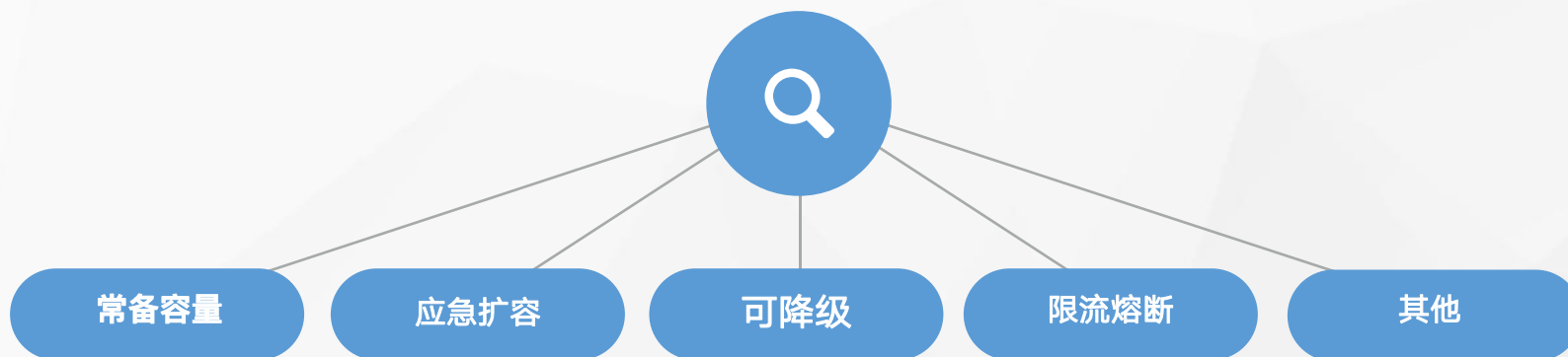
频道：588

礼物：277

活动：506

不足：评审节奏不合理，评审质量有待提升（需要增加小组内评审提高评审质量）；
代码质量需启动自动扫描调研并在本年度落地；

突发流量-故障策略-质量



目前应对突发流量已有策略：

一直播常备容量：5万

一直播应急扩容：大V来临可应急提前扩容

一直播可降级：可降级到白板

一直播限流熔断：核心业务限流熔断，待健全

不足：核心服务的限流降级熔断需要持续关注，给出技术使用标准。

对标微博，自动扩容缩容和大部门一起来推进；

1.业务场景

监控的意义-故障的发现与解决-两千三百万资损案例

活动红包，用户在活动期间可以抢红包（活动跨度可长可短，几个小时或者几分钟），红包有大小，大包分布也配置（一个小时出现几个大包，多少个普通红包出完再出一个大包，或者高峰时段内分布红包多一些）。所有配置由运营人员来操作。

2.故障触发原因

运营人员配置失误，配置很多大额红包，在一天多的时间内发放了很多大额红包。

3.系统监控缺失

缺乏业务监控，没有红包最大额度的报警，红包不断发送，直到一天后才发现，此时开始进行追回，但是多数用户拿到大额红包后就进行核销，无法追回。

4.最后统计

最后统计已核销的红包大约是2300万。

5.结果

Vp走人，整个部门重组。

体验监控在故障预告及发现的价值

监控的意义-故障的发现与解决-基础监控发现故障案例

故障标题	直播期间异常关播故障
所属产品线	直播
影响功能	观看
故障定级	
服务影响时长	5分钟
故障开始时间	
故障发现时间	
故障定位到原因时间	
服务恢复时间	
根本原因	阿里云数据库主从同步异常停止； 当前对未主动关播的直播进行5分钟自动关播处理，且无法针对所有极端异常场景进行兼容。
触发原因	从库异常停止主从数据同步，业务上因主从同步停止，触发异常断流
影响用户数	影响用户数： 21:13~21:14峰值PCU总数：60032 : 48439 11593； 影响主播数： 出现异常的主播：1174，
投诉情况	运营反馈13例
故障原因分类	
故障处理过程	2020-4-6 21:14 发现服务异常告警 2020-4-6 21:18 定位到连接池超限 2020-4-6 21:20 重新开播恢复 2020-4-6 21:27 联系运维进行扩容后，主从同步未恢复 2020-4-6 21:50 停止断流任务 2020-4-6 22:04 通知阿里云后台临时调高cpu，主从同步逐步恢复
改进措施	1、迁移数据库宿主机，并升级从库配置。 3、增加自动关播容错策略：接入平台断流检测接口（阿里云流质量数据） 接入阿里云的流质量数据，接入平台依赖平台排期 4、完善数据库监控告警。 5、长期方案：纳入容量提升项目多云方案。

体验监控在故障预告及发现的价值

监控的意义-故障的发现与解决-基础监控发现故障案例

