快手

Elastic Stack驱动下的直播体验优化

演讲人 罗喆

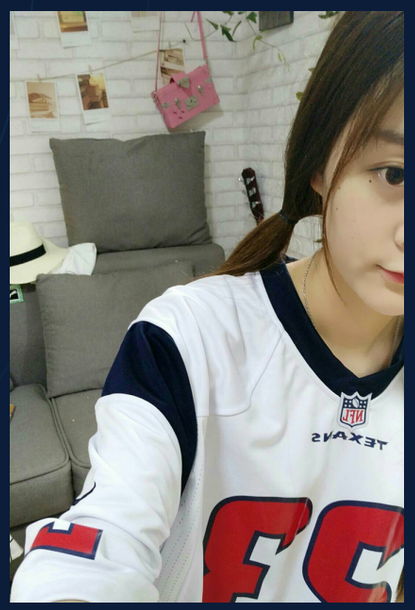
快手直播业务















快手视频团队简介

直播SDK团队

直播后端团队

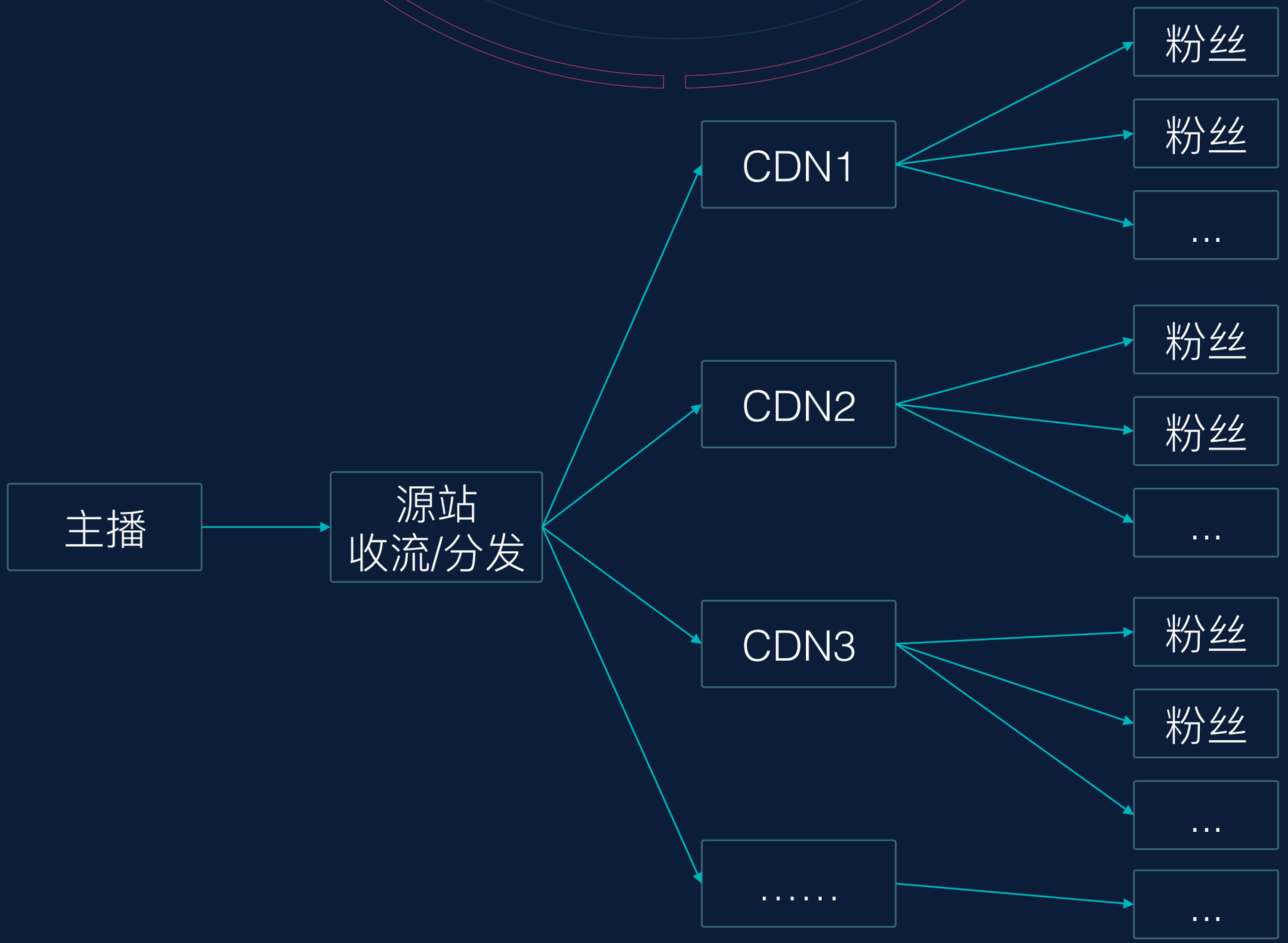
短视频技术团队

视音频算法团队

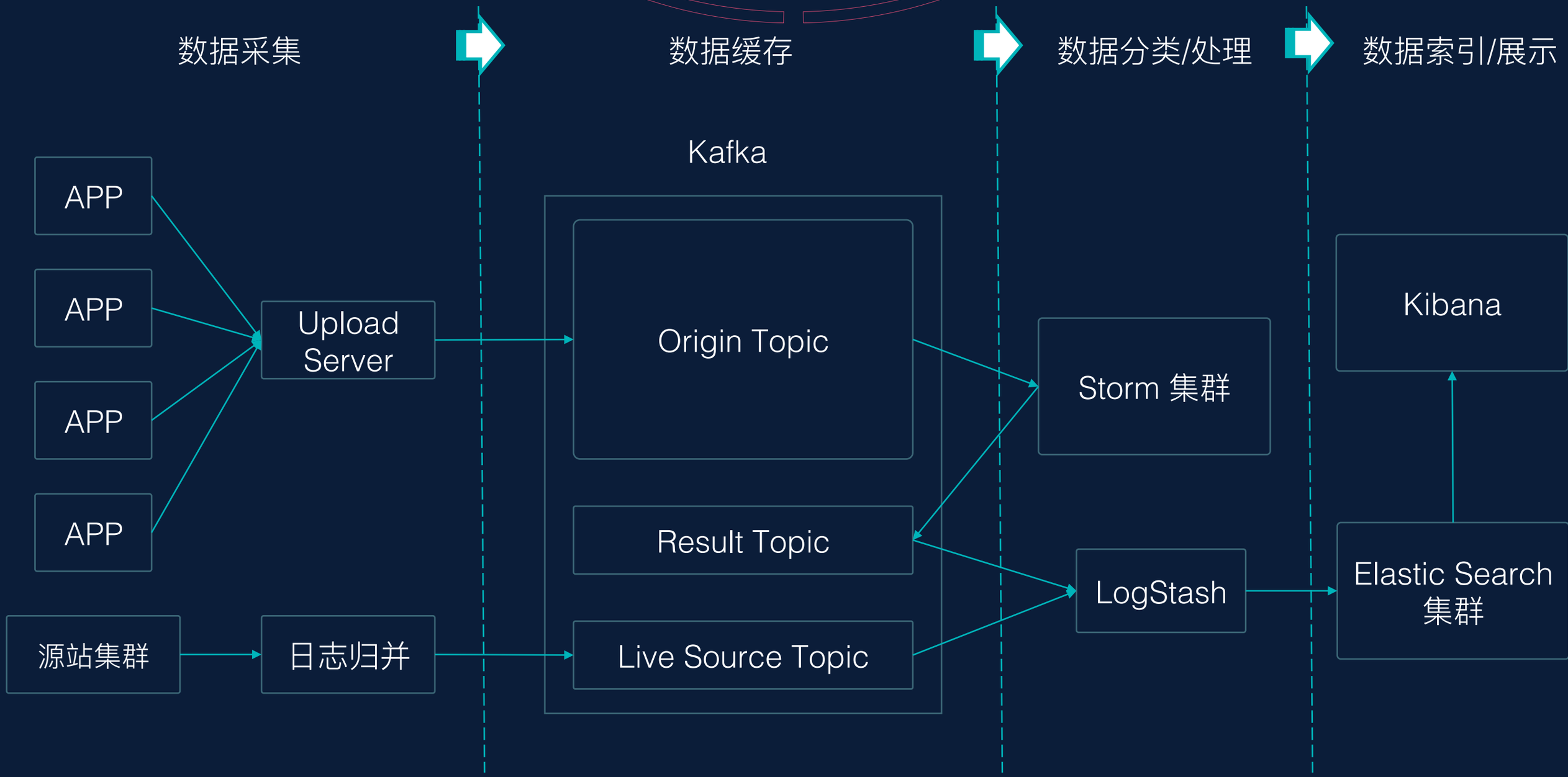
视频大数据团队



快手直播系统框架



快手直播质量数据处理Pipeline



快手直播质量数据处理Pipeline

- 保证数据可靠性

Kafka

Storm

- 强大的实时数据流处理能力

- 便利的接口连接Kafka ES

LogStash

快手直播使用ES的场景

实时监控

全地域监控

单例问题分析

实时监控-直播卡顿比

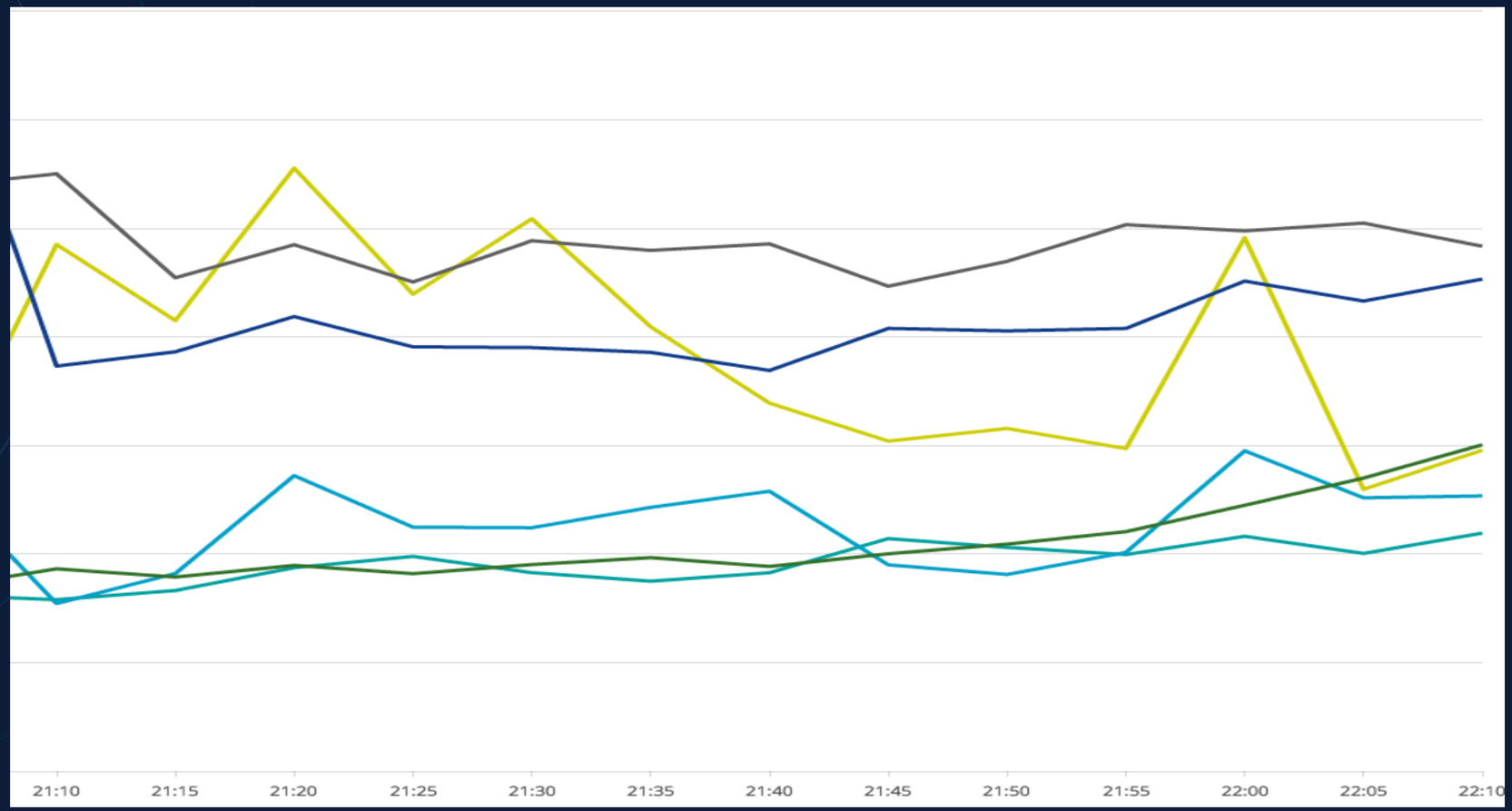
定义：单位时间内发生过卡顿的直播占比

Timelion：sum(block_cnt)/total

难点：需要区分多家服务商
timelion split和divide不能同用

添加 divideseries方法

实时监控-直播卡顿比


21:10
21:15
21:20
21:25
21:30
21:35
21:40
21:45
21:50
21:55
22:00
22:05
22:10

地域监控-观看人数热度地图

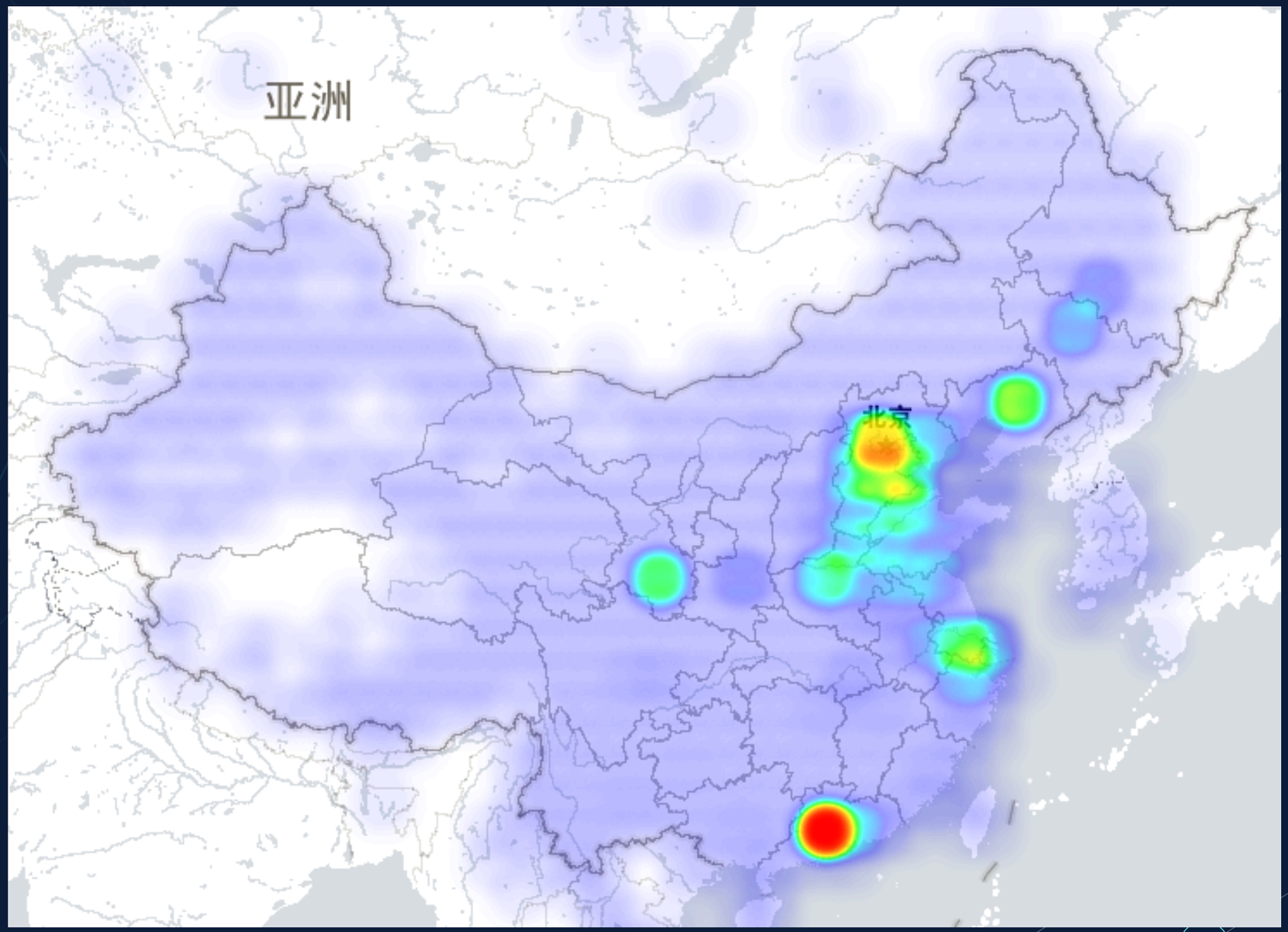
需求：了解当前全国直播观看热度

在日志信息中添加经纬度信息

Tile Map：地图缩放功能受限 英文地图

修改默认地图配置，使用高德地图
放开地图缩放限制

地域监控-观看人数热度地图



地域监控-观看人数热度地图



地域监控-分地区质量监控

需求：从省份+运营商维度监控CDN质量

↓

Tile Map：只能用做整体展示
无法独立区分各地区
区分运营商需要手动配置

↓

storm处理数据时添加地域，运营商信息

↓

Data Table：可以设置计算方法自动计算出省份+运营商组合信息

地域监控-分地区质量监控

Data Options

metrics

Metric

Add metrics

buckets

Split Ro... geoinfo.country.keyword: Descending

Split Ro... geoinfo.province.keyword: Descending

Add sub-buckets

geoinfo.country.keyword: Descending	geoinfo.province.keyword: Descending
中国	河北省
中国	山东省
中国	广东省
中国	辽宁省
中国	河南省
中国	江苏省
中国	黑龙江省
中国	吉林省
中国	山西省
中国	

Export: Raw Formatted

1

地域监控-分地区质量监控

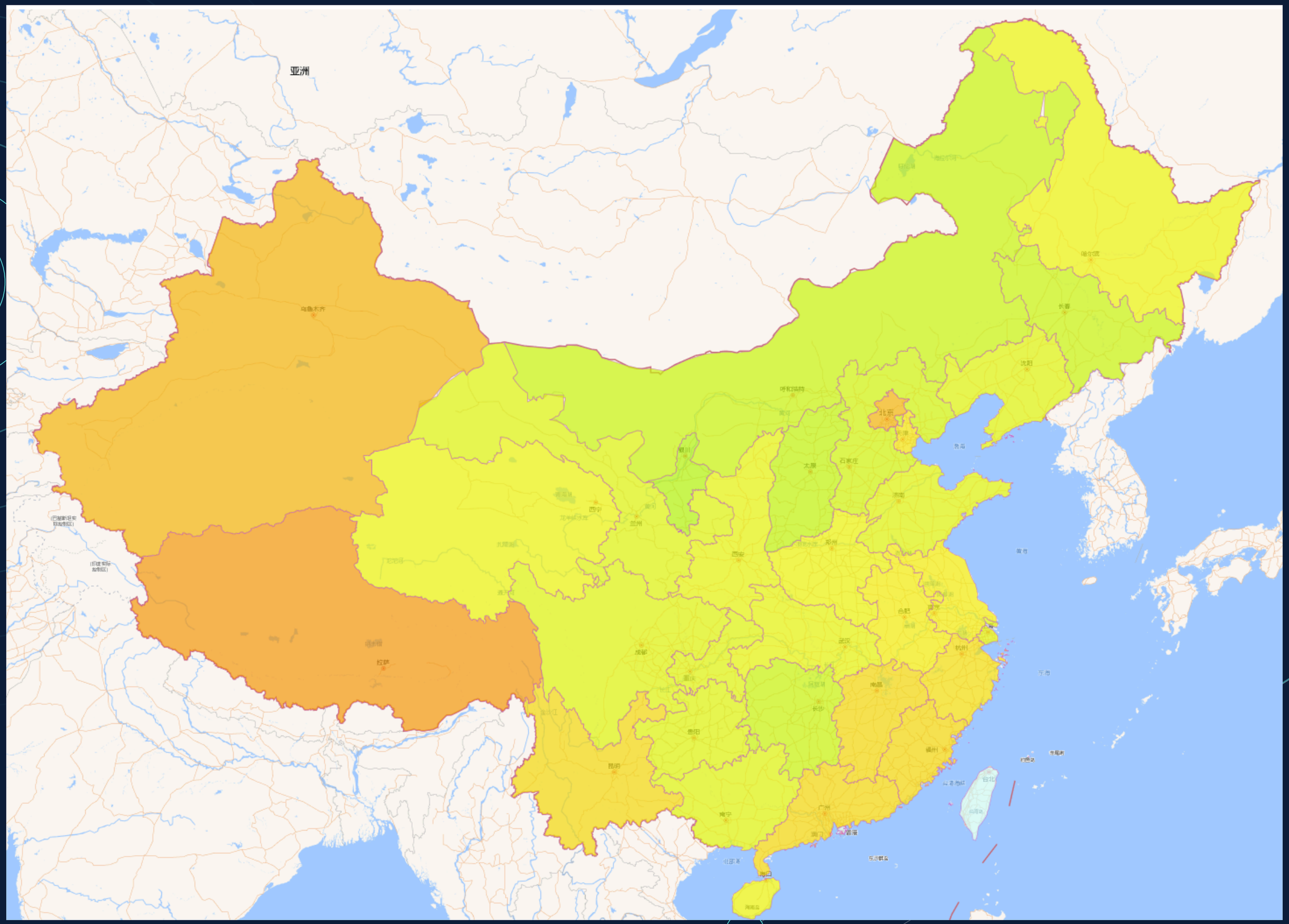
模仿Data Table调用ES API

获取根据地域+运营商组合的质量信息

调用百度地图API获取地区边界信息

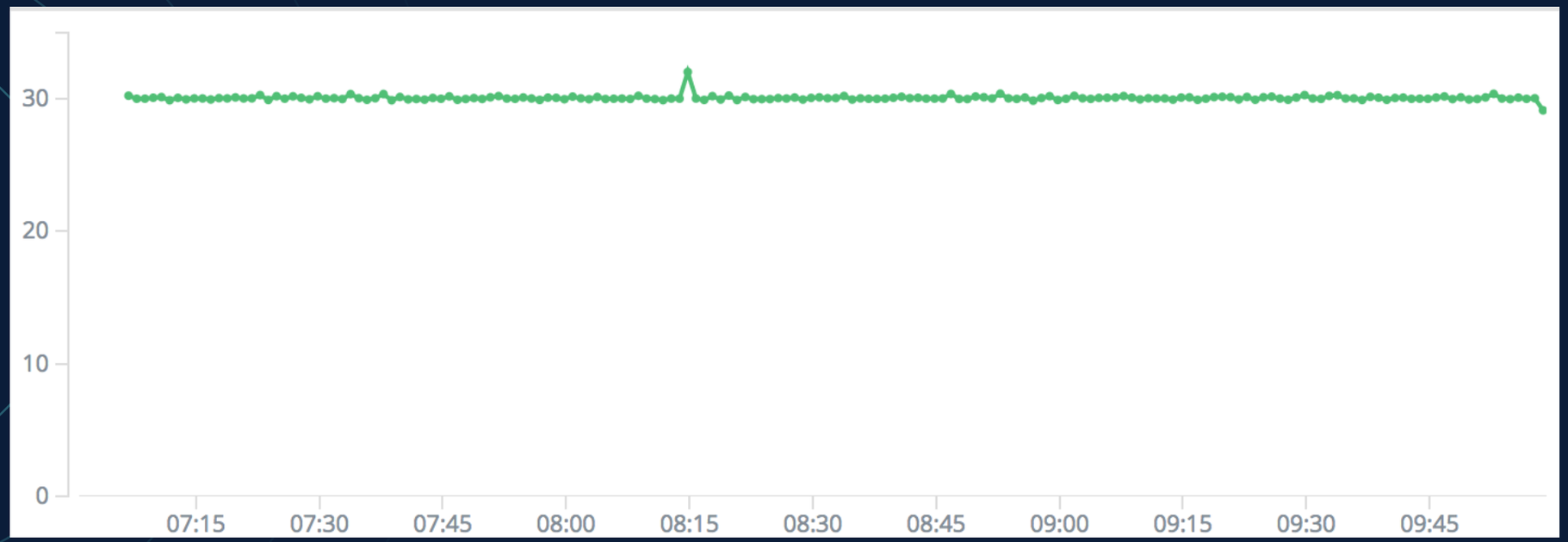
根据边界信息+质量信息渲染地图

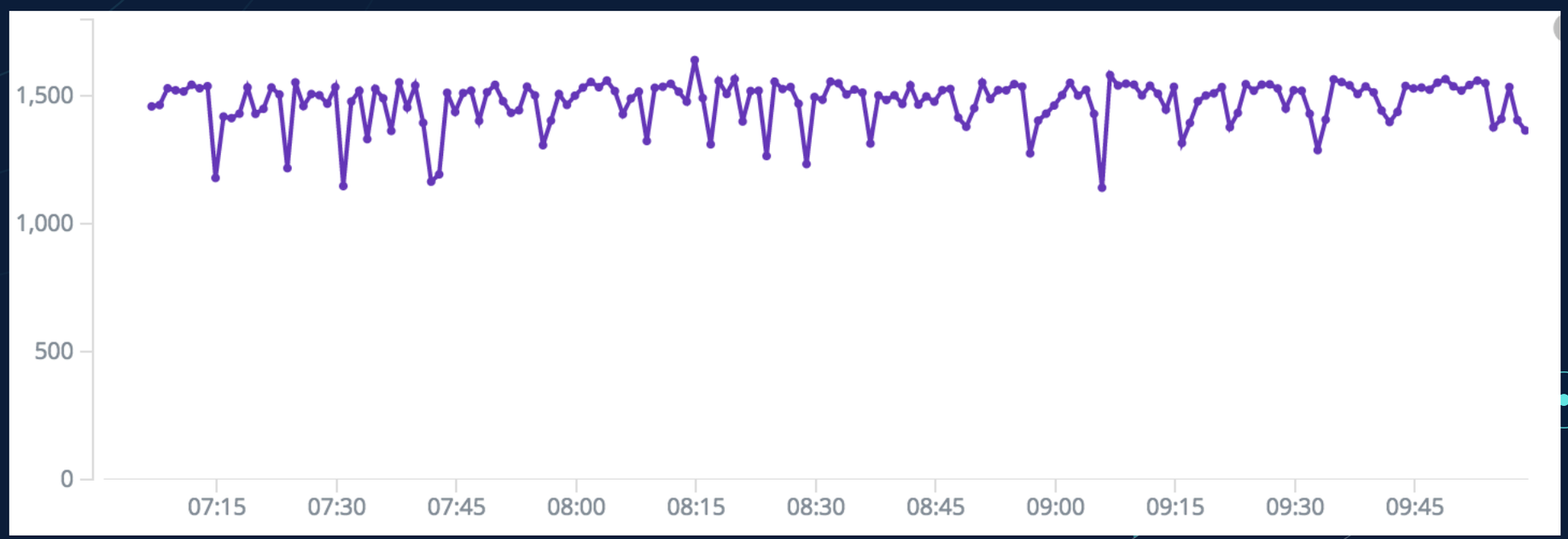
地域监控-分地区质量监控



单例问题跟进-主播质量曲线

源站信息定时打点 → ES 集群 → 根据时间+特点流信息展示曲线





单例问题跟进-直播观众质量数据

实时打点上报数据不方便看全局质量

观看一次直播汇总质量数据上报

根据特点信息索引所有观看某个直播的汇总上报

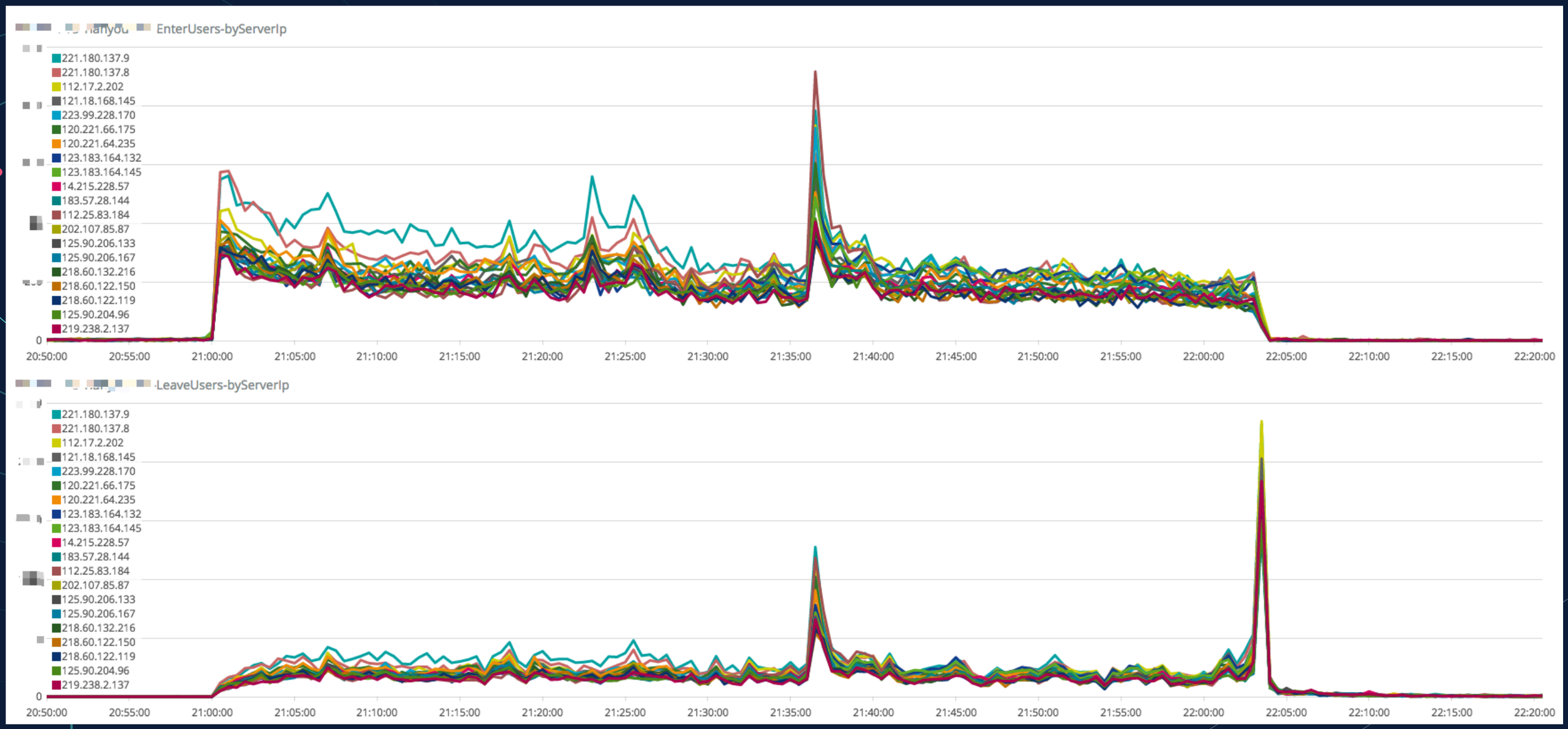
展示质量信息

单例问题跟进-直播观众质量数据

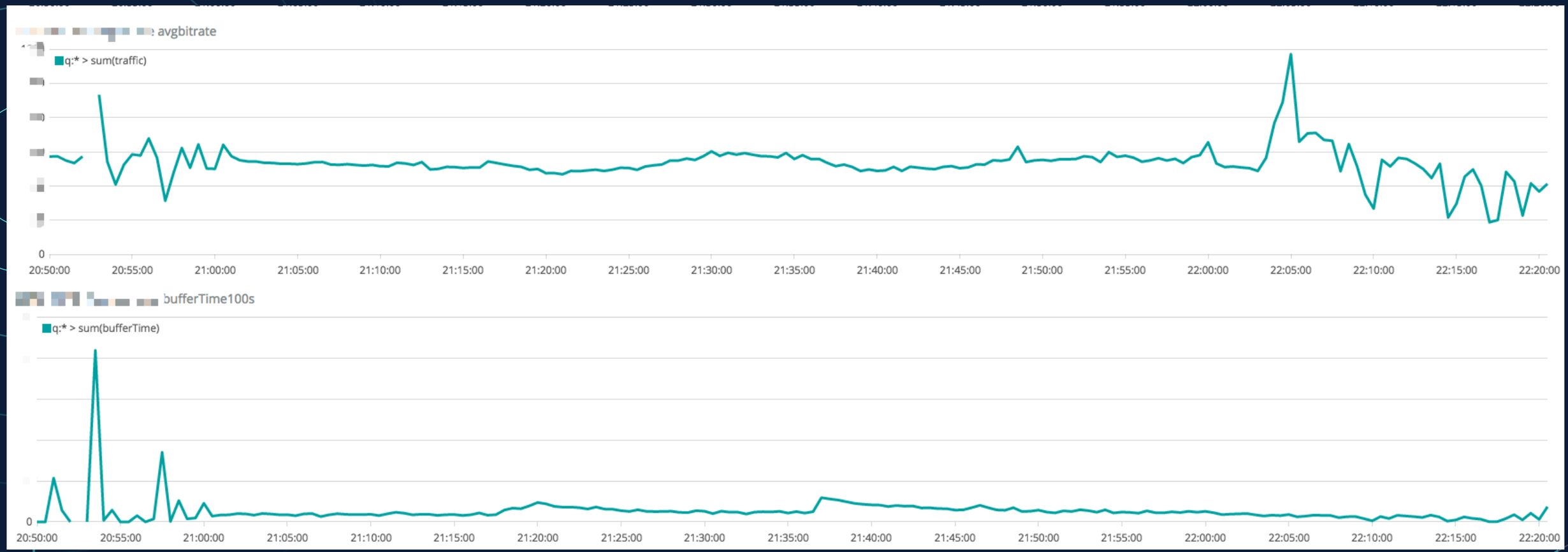




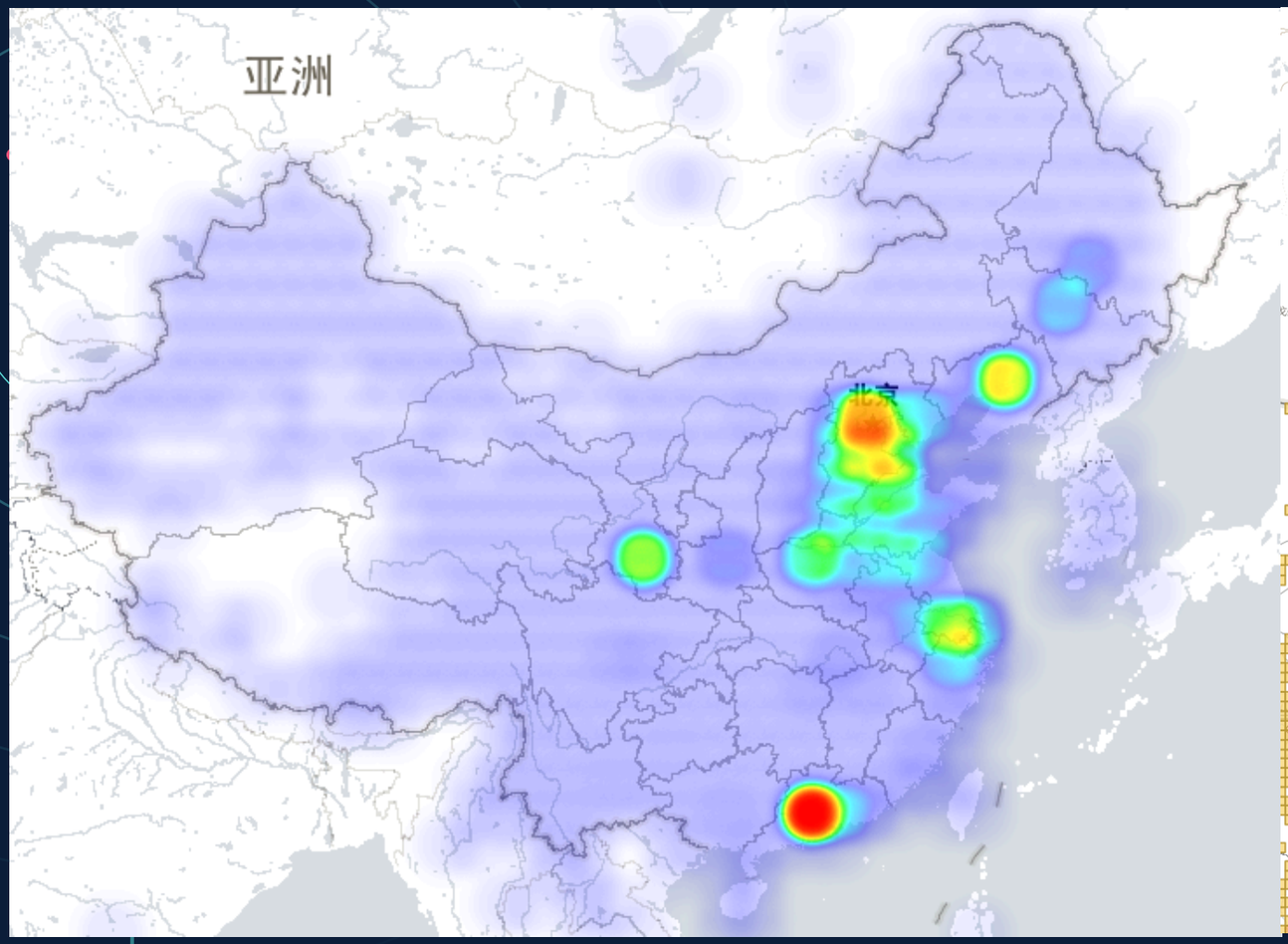
单例问题跟进-直播观众质量数据

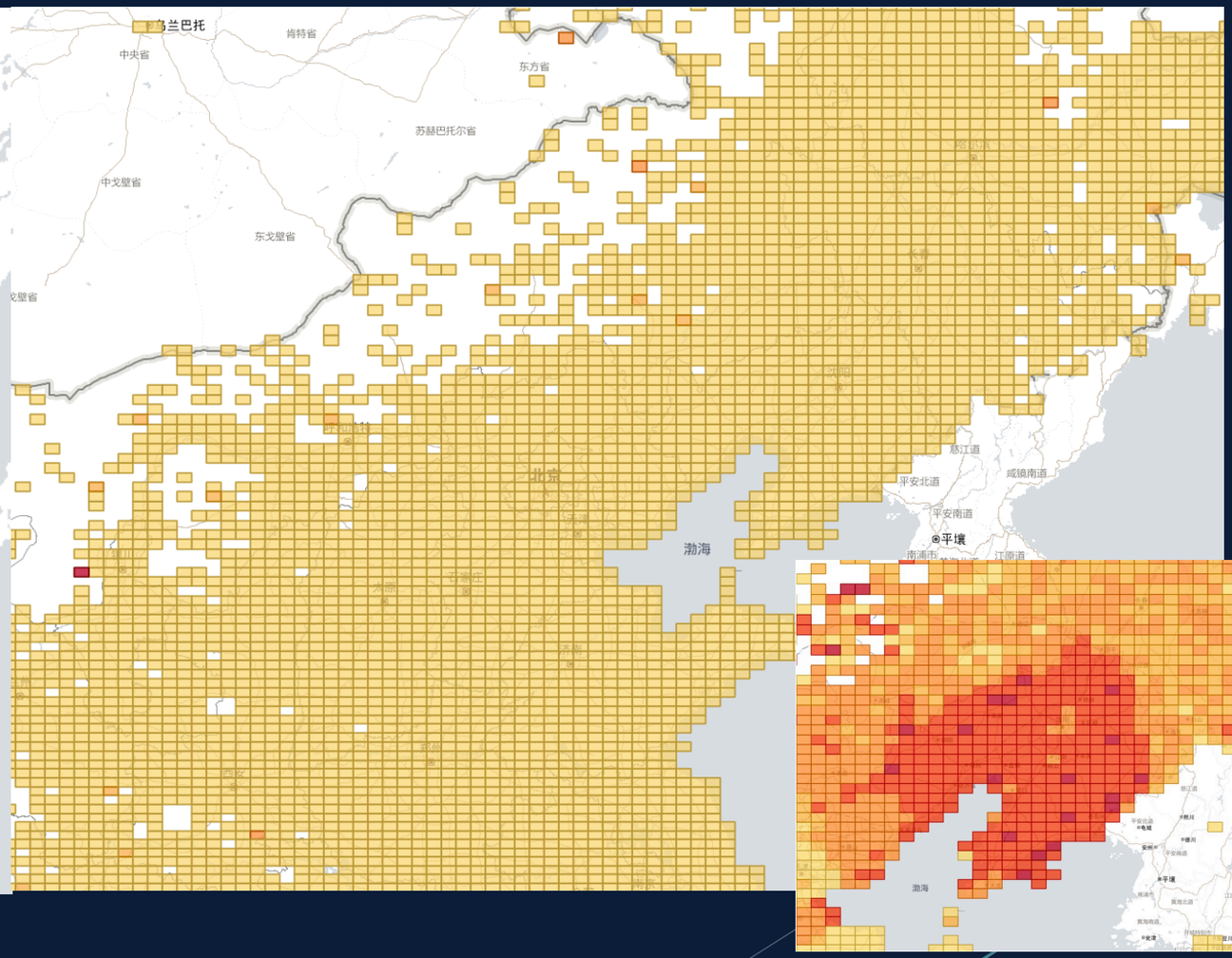


单例问题跟进-直播观众质量数据



单例问题跟进-直播观众质量数据





碰到的一些问题

无法处理时间跨度较长的数据(一天) → 使用SSD，冷热分开

不支持对单条日志内的多项数据之间进行运算 → 预处理之后存入ES集群

kibana 不支持展示原始value值 → 调整intervel 配合avg/sum

timelion不支持split之后再做复杂运算 → 自己编写计算方法

碰到的一些问题

timelion编写过程中，没有item自动补全

timelion不能预先指定index，而必须写在timelion语句内部

dashboard和visualize的时间interval不方便自适应匹配

下一步计划

智能报警机制

智能CDN调度

Advanced DashBoard

Thanks