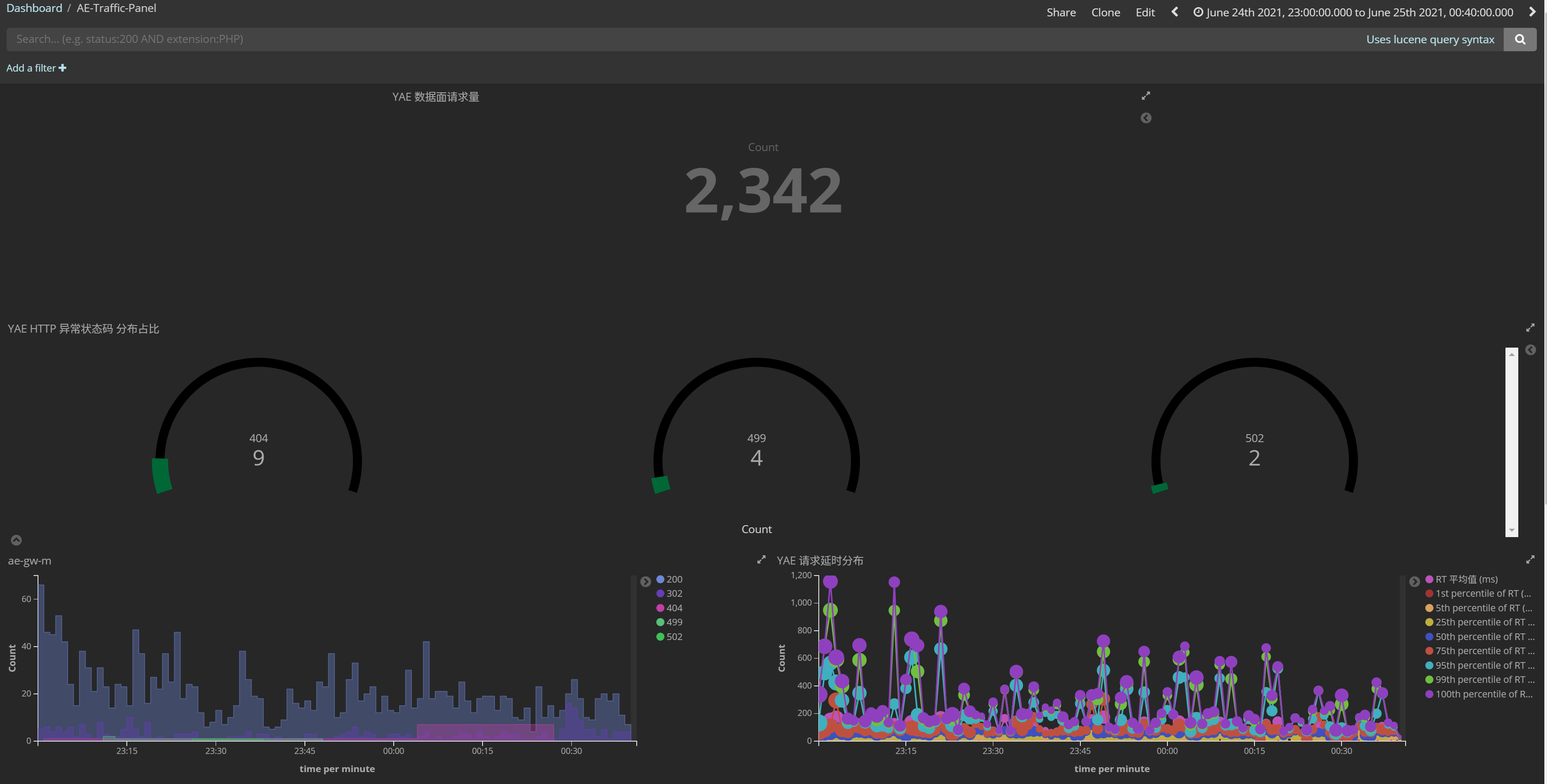
**20210624 YAE 生产环境故障演练记录**

* 参与人员
* YAE: 冉空,易大富,古几,苏枋
* 运维: 赵天冲 & 另一位不知名的大哥
* 时间
* 预期: 6.24 23:00 ~ 6.25 2:00
* 实际: 6.24 23:00 ~ 6.25 0:30

**流量情况**



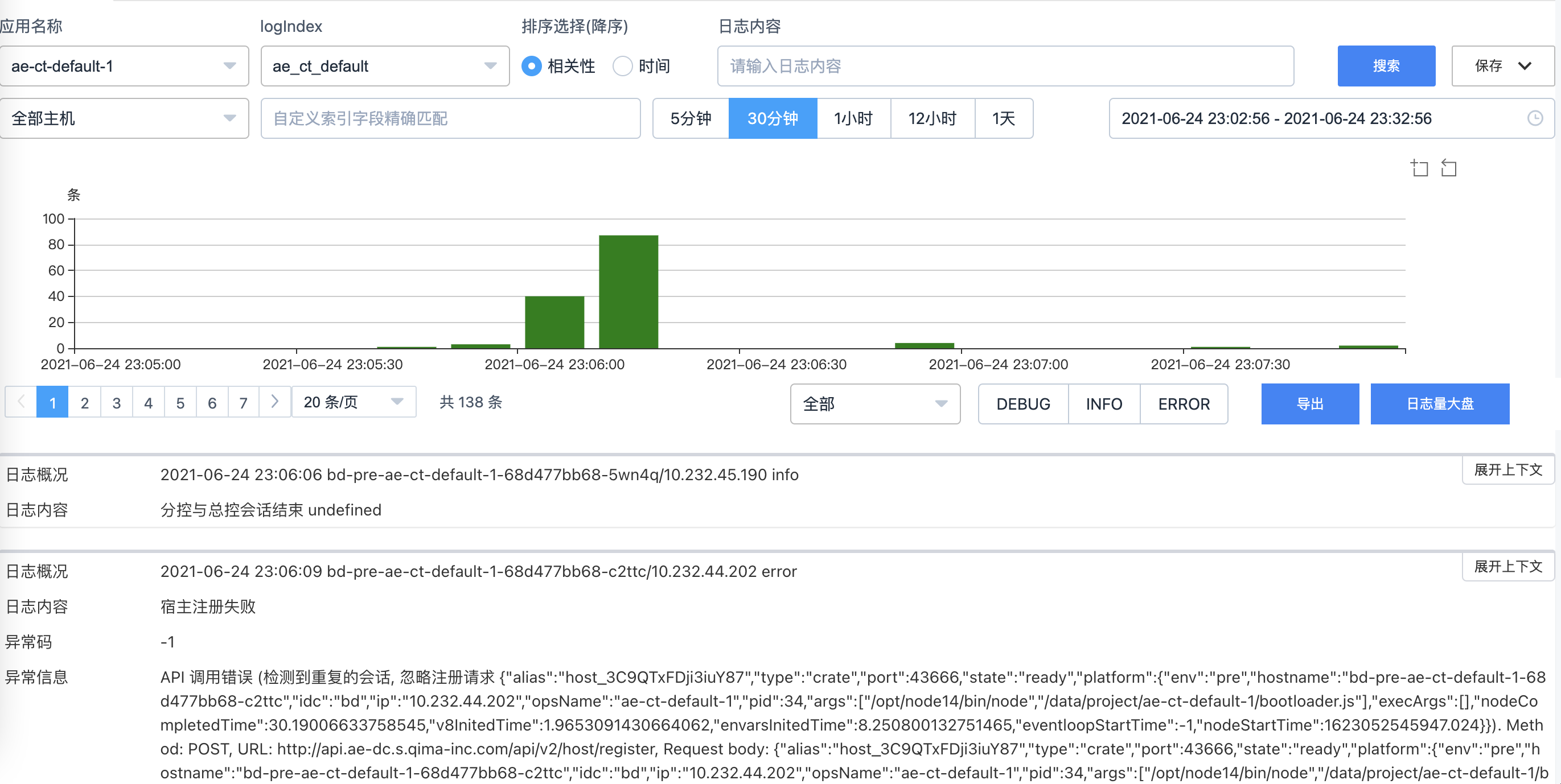
* 可见的影响是 2 个 502
* RT P99 分布个别时间段有突起, 属于调用后端服务有延迟

**操作对象: ae-dc (控制平面)**

* 操作：销毁 1 个 dc 集群成员
* 时间: 23:00
* 现象: 短时间内 dc 集群有大量错误日志和重建日志, 触发一批告警
* 整体异常收敛用时：≤ 30s
* 控制面集群恢复稳态用时: ≤ 10s
* 数据面流量抖动: 无
* 业务影响: 无
* 告警: 电话 (有一定延迟) + 短信 + 企微

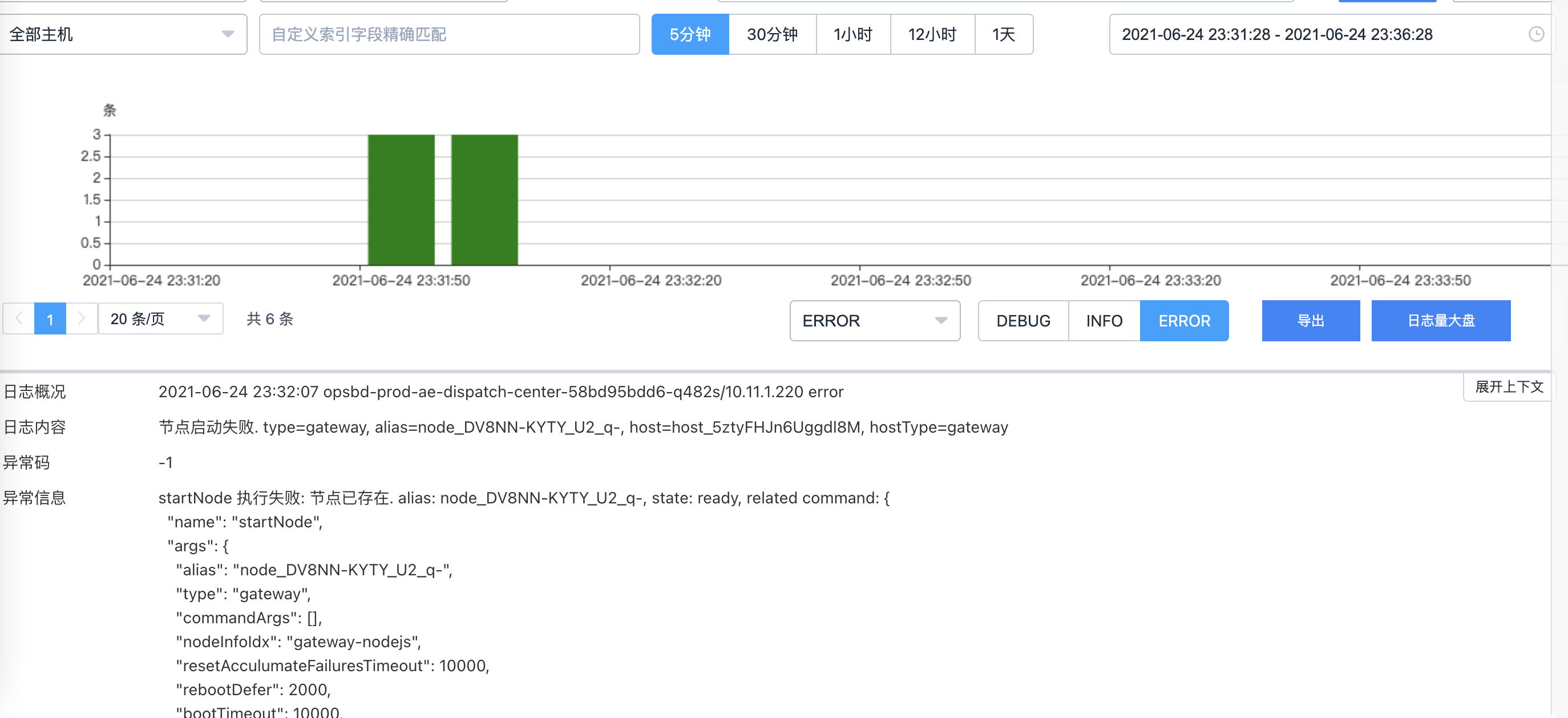


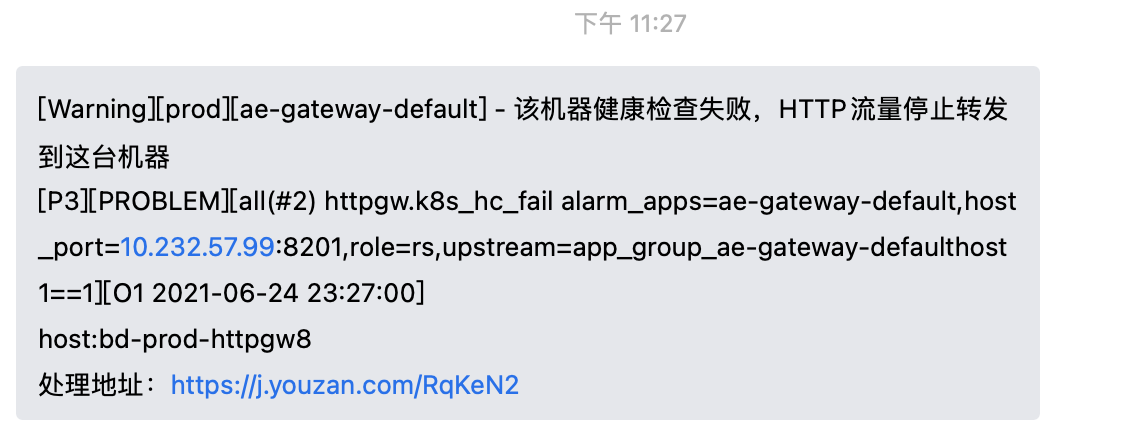
* 操作：销毁 1 个机房内的所有 dc 成员
* 时间：23:06
* 现象: 短时间内 dc 集群有大量错误日志和重建日志, 触发一批告警
* 整体异常收敛用时：≤ 30s
* 控制面集群恢复稳态用时: ≤ 10s
* 数据面流量抖动: 无
* 业务影响: 无
* 告警: 电话 (有一定延迟) + 短信 + 企微



**操作对象: ae-gw (数据平面)**

* 操作：隔离 bd 机房的所有网关
* 时间：23:27
* 现象: 半数网关网络隔离后, 业务访问瞬间出现 50% 的不可用
* 流量异常收敛用时：≤ 15s
* 数据面流量抖动: 有, youzan-security 出现一个 502
* 业务影响: 无
* 告警：收敛速度过快, 未能触发电话告. 企微告警正常



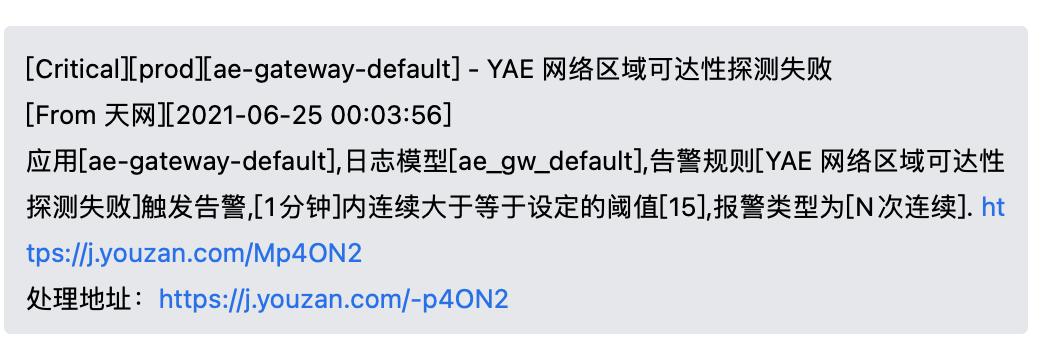


* 操作：销毁 bd 机房的所有网关
* 时间：23:37
* 现象: dc / gw 出现部分 warn 日志, 未发现明显异样
* 流量异常收敛用时：≈ 0s (平滑切流)
* 数据面流量抖动: 无
* 业务影响: 无
* 告警：无

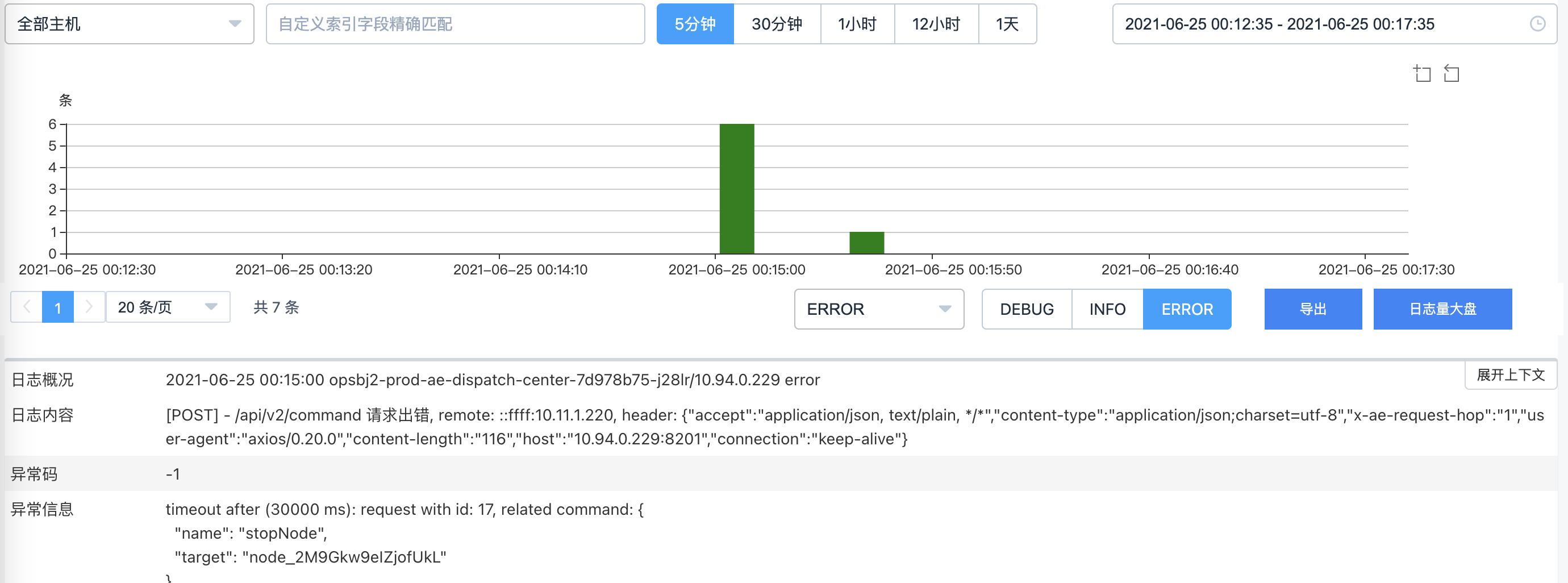
**操作对象: ae-ct (数据平面)**

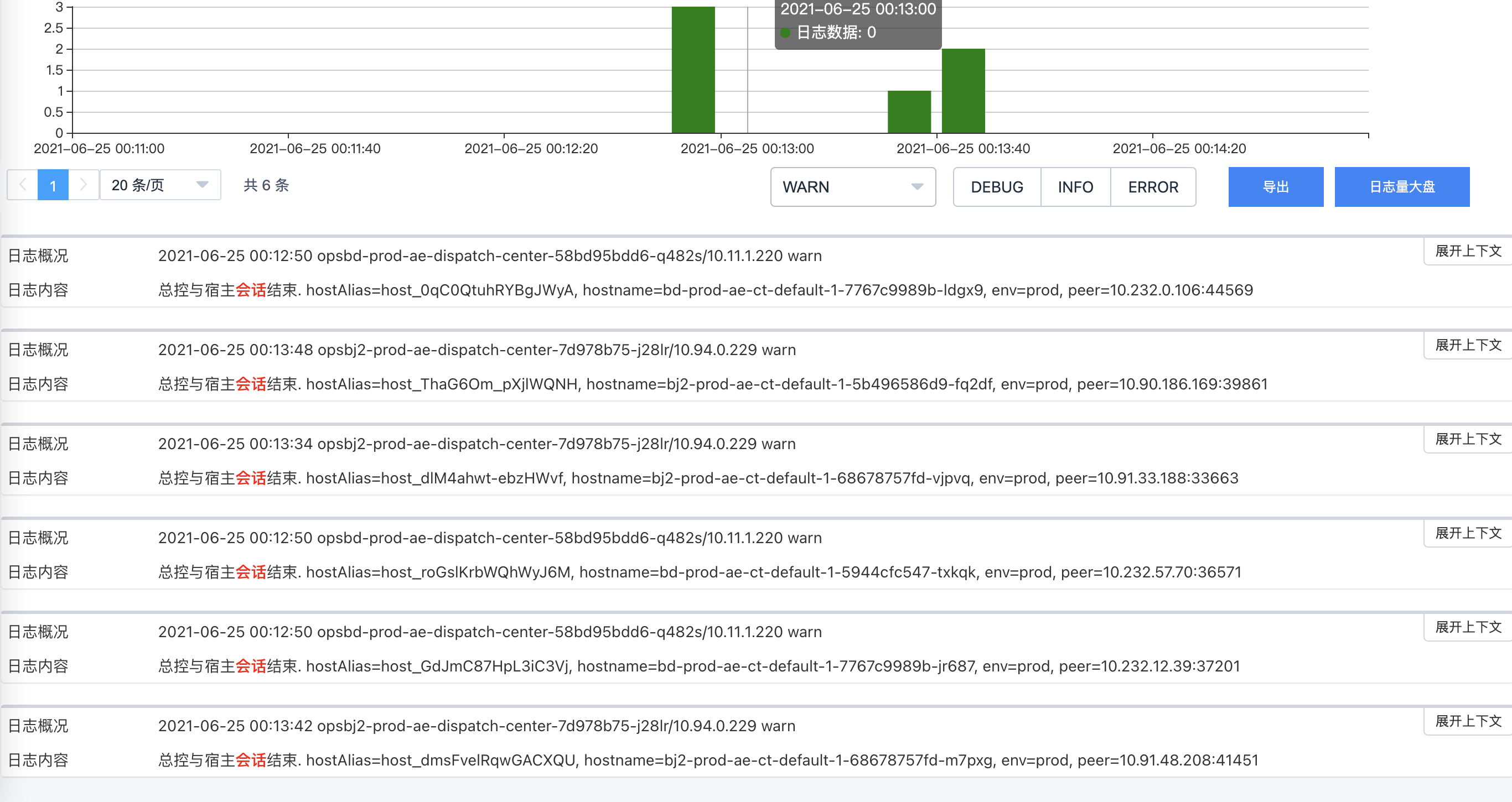
* 操作：隔离 bd 机房所有装箱
* 时间：00:02
* 现象: 业务访问 50% 不可用 & 超时, dc / ct 出现部分错误日志, 被隔离的主机上的业务应用被驱逐到其余未被隔离的主机上
* 流量异常收敛用时: ≤ 13s
* 数据面流量抖动: 有
* 实例故障转移用时: ≤ 3min
* 业务影响：无
* 告警: 电话 + 短信 + 企微





* 操作：缩容半数装箱
* 时间： 00:13
* 现象: 无, 因为缩容走了graceful shutdown
* 流量异常收敛用时: ≤ 0s
* 数据面流量抖动: 无
* 实例故障转移用时: ≤ 0s
* 业务影响：无
* 告警: 无





* 操作：下线一整个 ct
* 时间：00:28
* 现象: 出现告警, 两个区域其中一个不可用, 无法补充副本. 在 ct 重新上线之前每个业务实例的健康度保持在 50%
* 流量异常收敛用时: ≤ 0s
* 数据面流量抖动: 无
* 实例故障转移用时: N/A (另一半在 ct 重新上线之前没办法恢复)
* 业务影响：无
* 告警: 电话+企微

