

# 双态IT下的 华泰证券稳定性工程探索实践之路

王帅

---

华泰证券 资深稳定性工程专家



2021 可信云大会  
2021 TRUSTED CLOUD SUMMIT  
数字裂变 可信发展



## 王帅

华泰证券 资深稳定性专家



扫一扫上面的二维码图案，加我微信

从事证券行业系统运维10+年，熟悉证券行业业务规则及系统运行特点，拥有丰富的交易业务系统运维、测试经验，目前专注系统稳定性运行研究、平台建设及产品设计领域，负责稳定性工程平台的产品研究和运营。

## 1 稳定性工程目标

---

## 2 稳定性工程的应用场景

---

## 3 稳定性工程建设实践

---

## 4 应用效果

---

## 5 未来规划

---



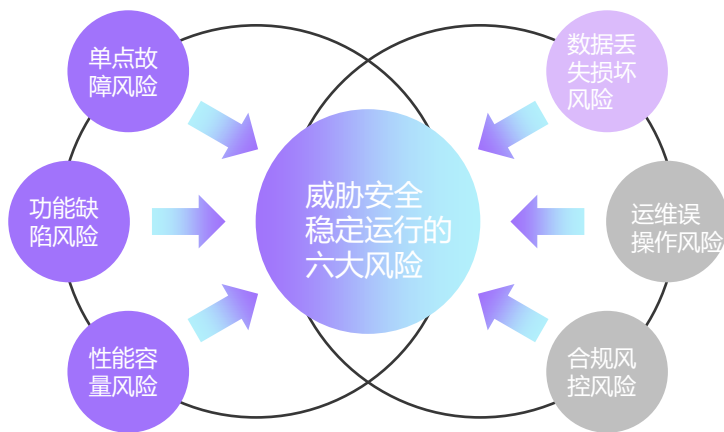
# 1、稳定性工程目标

---

# 稳定性工程的目标



通过大量的故障场景下系统脆弱性的检验，最大化的提前识别和消除技术风险，提升可靠性



## 技术风险识别由专家型向平台型演进

个人能力决定  
产品运行质量

技术手段化的  
风险识别

专家发展困境

专家知识的沉  
淀与复用

## 打破人力瓶颈，规模化风险识别及预防

组织的人力规  
模受控

人力规模与业务  
规模强依赖解耦

组织的业务贡  
献价值受限

实现规模化业务  
覆盖及质量稳定

## 技术风险管理科学化、系统化

自律、自发型  
风险管理

全面分析风险来源，  
建立科学化手段

风险管理的脱  
节

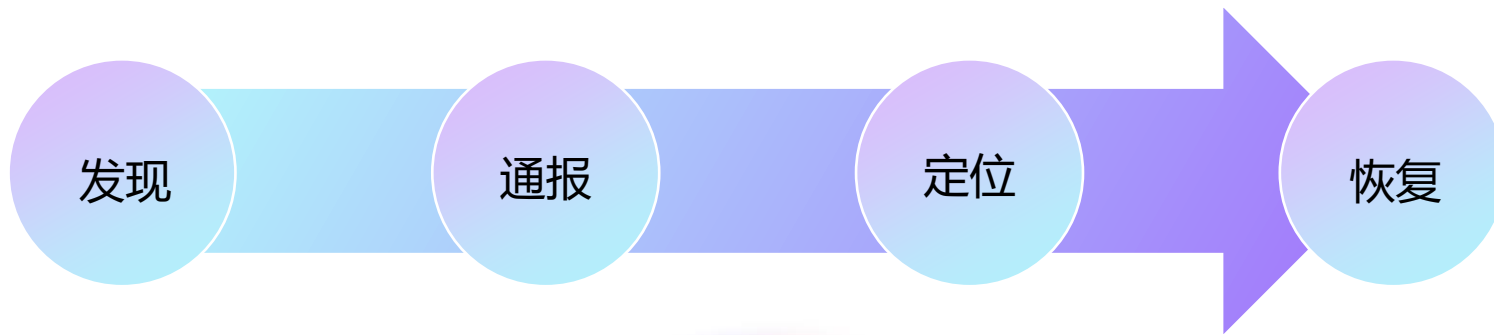
技术风险管理融入运  
行保障工作大体系

# 稳定性工程的目标



检验故障场景下的各环节协作配合效率，发现效能瓶颈，提升应急全流程的效率

## 四个第一



## 2、稳定性工程的应用场景

---

# 稳定性工程的应用场景



面向6大类应用运行风险，提供风险识别的技术化手段，实现风险识别全覆盖  
基于故障演练，建立风险应对度量、故障知识沉淀及应用一体化能力





# 稳定性工程的应用场景



证券系统稳定性的要求越来越高，但是交付和变更更加频繁，双态IT特征下，通过稳定性工程手段检验典型业务场景下的系统稳定性表现，提升系统可靠性。



# 3、稳定性工程建设实践

---

# 华泰稳定性工程平台



面向证券行业特点，建设华泰“稳定性工程平台”，打通监控、变更、测试、应急等生产运维场景。

故障注入能力：AHAS + 证券特有能力的自研

1

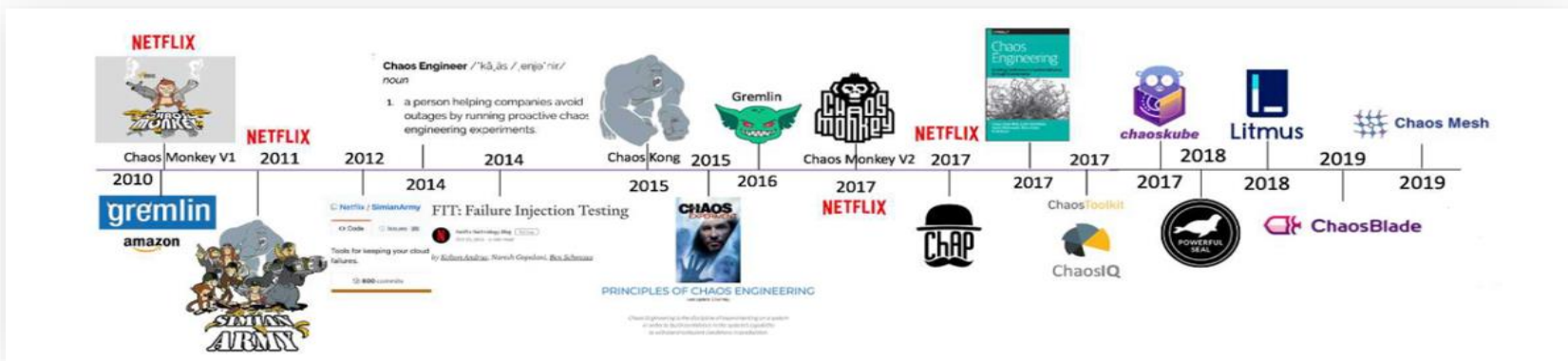
故障发现能力

2

故障定位能力

3

故障处置能力



# 华泰稳定性工程功能架构



- 1 演练管理
- 2 故障演练
- 3 演练自动化
- 4 演练评价



# 故障类型及故障构建能力规划



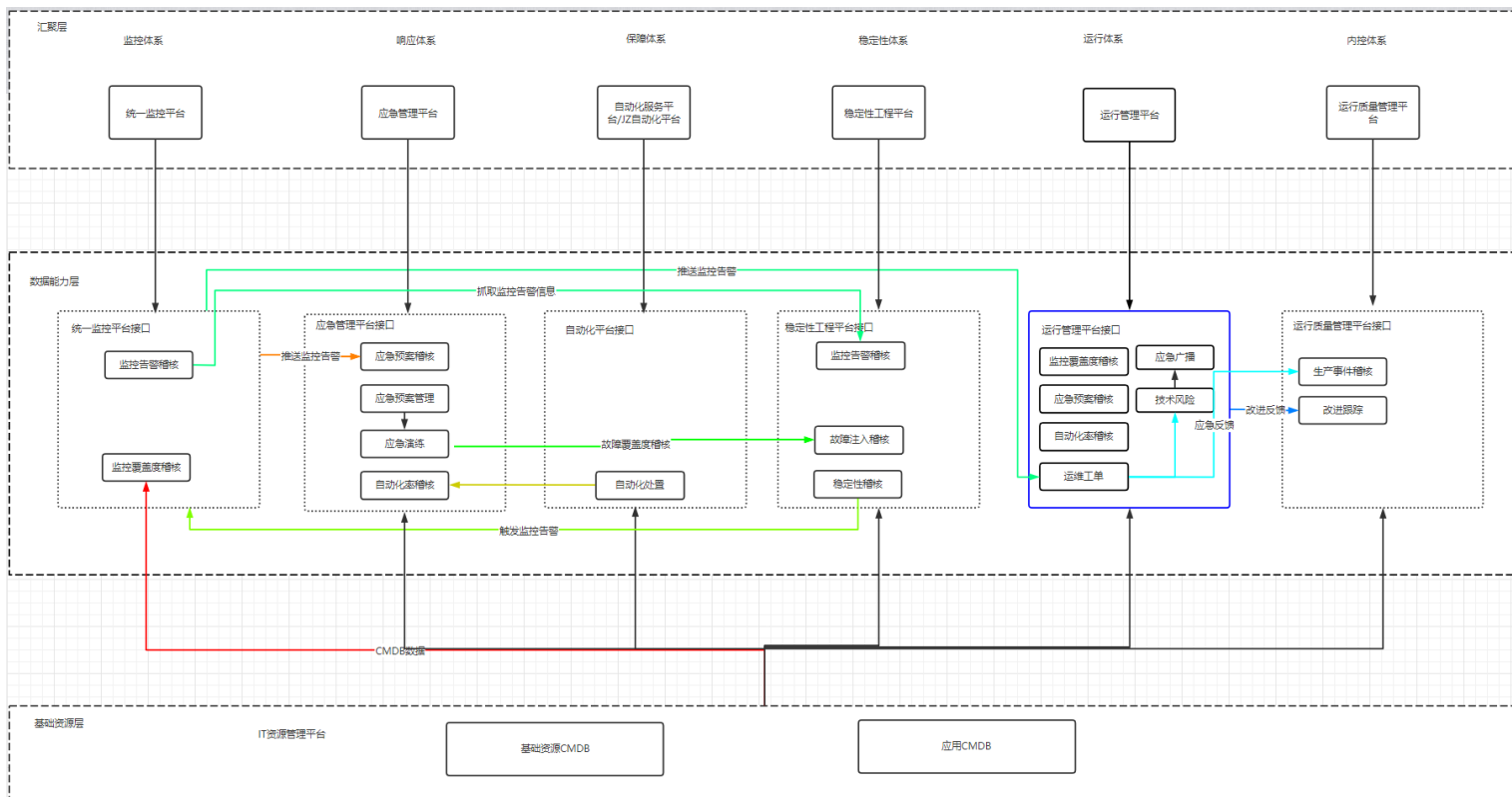
- 基础设施故障 (CPU、内存、网络、磁盘、进程等)
- 容器服务故障 (节点不可用等)
- 应用服务故障 (进程、内存、CPU、JVM、服务调用)
- 云资源故障 (宕机、重启)
- 数据损坏 (误删、丢失等)
- 故障编排 (故障节点组合)
- 业务“卡吊死”



# 稳定性工程融入运维一体化体系



- 应用CMDB
- 统一监控
- 统一操作
- 统一测试
- 运行质量管理
- 应急管理



# 面向改进，量化评价故障发现能力



1

稳定性工程平台是运维业务链的核心，通过故障注入发现上下游系统隐患，评价系统稳定性能力。



涉及统一监控平台、自动化服务平台/统一操作台、稳定工程平台、应急管理平台。

- 案例：故障注入后，产生监控告警，调用故障知识库的应急预案，进行应急处置，完成系统稳定性能力评价。

2

稳定性工程平台通过故障注入验证监控告警的有效性。



涉及统一监控平台、运维工单系统、稳定性工程平台。

- 案例：故障注入后产生监控告警信息与运维工单，形成技术风险或者生产事件，量化监控告警的覆盖度与一二线联动情况。



# 面向改进，量化评价应急预案有效性



3

通过稳定性工程平台注入生产故障，验证系统稳定性与应急预案的有效性。



涉及应急管理平台、稳定性工程平台。

- 案例：模拟真实发生故障后的应急处置能力，将故障注入与**应急预案覆盖度**同步给应急管理平台。

序号	分类	评估项	说明	分数
1	部署架构容灾能力	应用部署	跨机房灾部署记10分，单机房集群部署5分，中间件、模块、线路存在单点每项减1分，该项最多0分。	0
		数据库和文件存储	使用的数据库和文件存储具备跨机房自动备份和切换能力，记10分；只在单机房灾备，或需要人工备份和切换记5分；其他0分。有多种存储的取得分最高的分数。	0
2	变更风险控制能力	变更管控	通过部门内操作规范变更流程变更上线记10分，否则记0分。通过灾备墙抽取。	10
		版本测试	测试后上线10，无测试上线0分。	0
		版本部署	DEVOPS平台、自动化服务台、全证自动化、数据中台、EMAR发布定义为自动化，记10分，其他工具化或脚本化5分，手工部署0分。	0
		灰度/蓝绿	灰度/蓝绿发布10分，否则0分。	0
		回滚能力	具备脚本回滚能力（含数据库回滚）记10分，不具备0分。	0
3	运行态监控能力	变更监控	针对变更会调整相应监控记5分，对变更内容会做相应的监控进行观测8分。未调整或增加过监控记0分。	0
		基础监控	包括CPU、内存、磁盘，统一监控平台抽取。	10
		网络监控	网络监控，统一监控平台抽取。	10
		应用监控	进程、端口、接口监控全覆盖10分，部分覆盖6分，无监控0分。	0
		数据库监控	数据库监控，已监控10分，未监控0分。	0
		业务监控	具备业务监控模型，并对核心指标请求量、业务量进行监控记10分；有业务监控但没有业务监控模型记5分，无业务监控记0分。注：业务监控模型是指对业务范围、业务量等有整体监控，在技术故障时可以评估业务影响情况。	0
		体验监控	具备用户体验监控评估模型10分，否则0分。部分系统不涉及。	0
4	稳定性防护能力	主备切换	主备无需手工切换（比如分布式），或自动化切换10分，包括数据库，否则记0分。	0
		限流降级	针对业务突增、高并发等异常场景，具备限流、降级、熔断能力记10分，不具备0分。部分系统不涉及。	0
		弹性伸缩	具备基于业务流量的弹性扩缩能力记10分，不具备0分。	0
		流量调度	针对业务异常，具备流量自动调度能力记10分，不具备0分。部分系统不涉及。	0
		应急预案	针对线上业务异常或系统异常有有效应急预案，可快速止血记10分，无预案0分。	0
		压测	具有压测环境并定期开展压力测试记10分，无压力测试0分。	0
5	安全防护能力	SLA管理	具有SLA并例行检查SLA情况10分，否则0分。	0
		容量管理	具有容量管理模型，并例行开展评估10分；不具有模型但例行开展非系统评估5分，否则0分。	0
		安全事件	服务器上安全事件（补丁、防病毒、漏洞扫描软件、主机IDS）的全面覆盖得10分，90%覆盖得5分，否则0分。	5
		系统及组件	未使用超出生命周期操作系统和基础组件得10分，否则0分。	0
	风险评估	系统涉及的IT项目100%经过安全评估得10分，90%得5分，其余不得分。	0	



# 面向应急，构建故障知识库



- 提供告警应急预案关系维护能力，应急预案包括处理方式和处理脚本，通过日常演练优化故障知识库。

- 演练结束后自动新增一条故障知识库记录，由用户确认入库。

序号	事件 ID	事件名称	关联事件	关联系统	关联模块	处置简称	处置详情	创建人	创建时间	操作
1	24127								2021-06-03 17:59:04	<a href="#">查看详情</a> <a href="#">编辑</a> <a href="#">删除</a>
2	12772								2021-05-28 15:27:47	<a href="#">查看详情</a> <a href="#">编辑</a> <a href="#">删除</a>

# 故障定制，构建数据丢失损坏



- 提供自定义故障构造与恢复能力。
- 支持数据丢失、权限丢失、数据库访问激增、数据库进程丢失、数据库主机宕机等常见故障注入，调用统一操作平台实现故障自愈。

稳定性工程平台  
STABILITY ENGINEERING PLATFORM

故障演练

演练列表

首页 演练列表 × 更多操作

演练名: [ ] 状态: [ ] 标签: [ ]

演练方式: [ ] 风险等级: [ ] 团队协作: [ ]

业务时间轴: [ ] 资源类型: [ ] 故障类型: [ ]

查询 新建演练

序号	id	演练任务名称	故障类型	风险等级	系统	模块	状态	标签	创建时间	更新时间	创建用户	操作
1	1059	自动化服务平台-后端模块-数据损坏	-	-	自动化服务平台	后端模块	检查成功	数据损坏	2021-07-06 14:35:46	2021-07-06 15:10:32		演练 查看详情 编辑 拷贝 删除
2	1057	故障演练平台-后端服务-数据损坏	-	-	-	后端服务	成功结束	数据损坏	2021-05-27 20:44:28	2021-05-28 11:34:31		演练 查看详情 编辑 拷贝 删除

共 2 条数据 < 1 > 10 条/页

# 降低门槛，预置演练场景库



- 实现系统与模板、机器的关系关联。
- 展示待演练系统的所有已安装探针的机器列表，选择对应的机器进行演练。
- 打通CMDB，一键批量生成演练任务。

The screenshot displays a 'New Scenario' dialog box with the following fields:

- 场景名称: 请输入演练场景名称 (例: 卡-Window)
- 演练标签: 请输入演练场景所属系统标签
- 故障类型: 请输入 (请从探针库中选择)
- 参数设置: [Empty field]

Buttons: 取消, 确定

Below the dialog is a table of scenarios:

序号	名称	类型	状态	创建时间	探针	操作	调用次数	操作
1								
2		卡		2021-04-30 15:46:47	CPU占用	进程自动获取	0	编辑 拷贝 删除
3		用		2021-04-30 14:45:00	网络延迟	进程自动获取	1	编辑 拷贝 删除
4		卡		2021-04-30 14:36:31	CPU占用	进程自动获取	1	编辑 拷贝 删除
5		卡		2021-04-30 13:05:50	CPU占用	进程自动获取	1	编辑 拷贝 删除
6	卡-...ws	卡		2021-04-27 16:30:24	网络延迟	5000ms	20	编辑 拷贝 删除
7	卡-...3			2021-04-26 13:39:21	杀进程	进程自动获取	2	编辑 拷贝 删除
8	卡-...03			2021-04-26 13:39:21	杀进程	进程自动获取	2	编辑 拷贝 删除
9		卡		2021-04-26 11:17:54	CPU占用	进程自动获取	1	编辑 拷贝 删除
10	卡-...ows	卡		2021-04-27 16:30:24	网络延迟	5000ms	19	编辑 拷贝 删除

# 4、应用效果

---

# 建立稳定性工程模型



演练模型化、体系化

首次通过体系化、规模化开展地毯式故障演练，建立起故障演练模型，构建了故障场景矩阵，建立了一线监控与二线处置联动，实现了历史故障回放机制。



故障演练  
模型



故障场景  
矩阵



一二线运  
维联动



历史故障  
回放

业务卡顿

业务无响应

业务完全故障

全业务链

## 常见业务故障场景



### 卡

客户的所有请求都可以被连续处理，但响应的的时间标准差比较大，但基本上没有超时出现。



### 吊

系统进程、端口正常，系统响应部分超时或全部超时，长时间无应答或超时重试。



### 死

系统业务进程已不存在，客户的服务请求很快得到失败应答。

# 保卫波特姆行动：故障演练地毯式覆盖



- 应用资源
- 通讯链路
- 数据库
- 负载均衡
- 域名

以史为镜 故障回放

业务线	应用系统	演练日期	模块名称	模块	TP
报盘	已业务运营平台				
场内交易业务线-柜台中间件	已业务运营平台				
场内交易业务线-涨乐委托	接入网关				
	财富通委托系统				
行情业务线	行情资讯系统-通达信				
	行情资讯系统-同花顺				
	n期权策略交易系统				
	衍生品交易系统				
清算业务线	功行情系统				
全账户	已业务运营平台				
	互联网服务平台				
账户业务线	客户系统				
	互联网线上开户系统				
	已业务运营平台				
	高级客户信息整合系统				
	中心				
证券接入	柜员业务平台				
	业务处理平台				
	报盘接口程序				
总计	报盘接口程序				



# 可视化稳定性工程活动

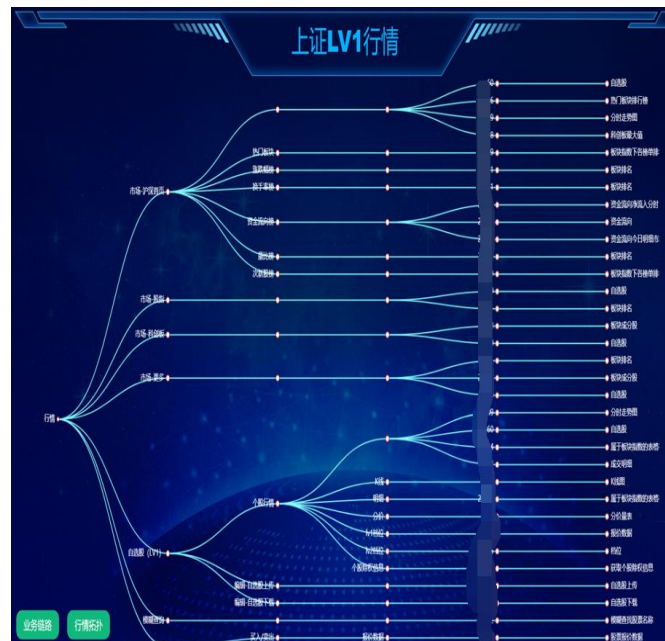
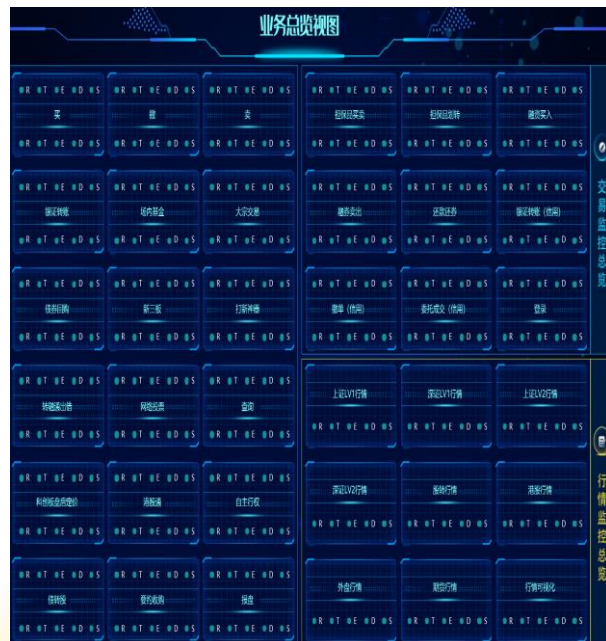




# 提升易用性，爆炸半径可视化



根据业务特点，建立端到端的实时感知体系，监控故障演练期间业务表现



# 稳定性工程管理视图



过滤条件

演练人员工号

系统名称

演练日期范围

所属团队

查询

可演练主机数量: 14248

今日新增可演练主机数量: 0

用户数量: 246

演练场景总数: 2643

## 演练系统统计

### 中心系统统计

中心名称	系统数	已演练系统数	完成比例
委员会			100%
中心			100%
中心			100%
技发中心	3		100%
技发中心	2		100%
技发中心	1		100%
技发中心	1		100%
技发中心	1	1	100%
技发中心	2	2	100%
中心	6	6	100%

### 团队系统统计

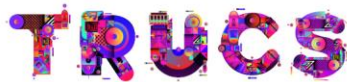
团队名称	系统数	已演练系统数	完成比例
委员会			100%
运营与攻防团队			100%
运营团队			100%
运营团队			100%
中间件团队			100%
平台开发团队			100%
云计算平台团队			100%
技术研究团队			100%
基础研发团队			100%
平台团队			100%

# 稳定性工程文化运营

- 安全生产专项行动
- 建党100周年专项业务保障
- 保卫波特姆行动



# 效果-保卫波特姆行动



全

实现300+核心业务  
系统全覆盖

快

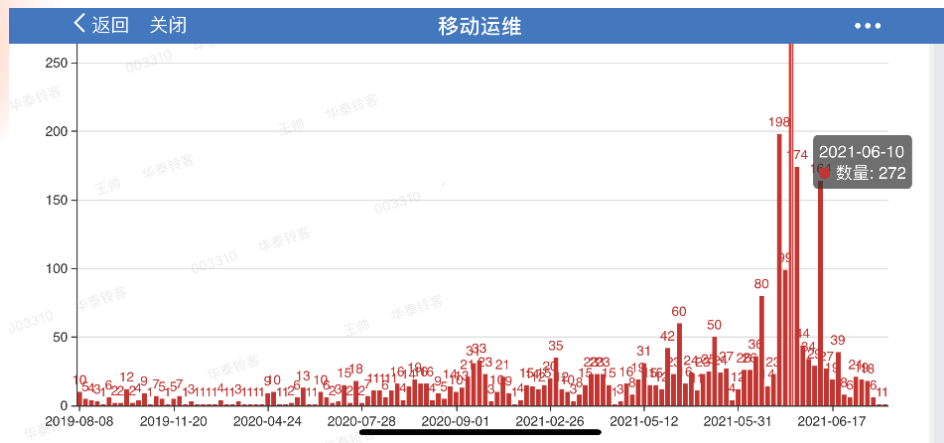
单日最高演练272次

多

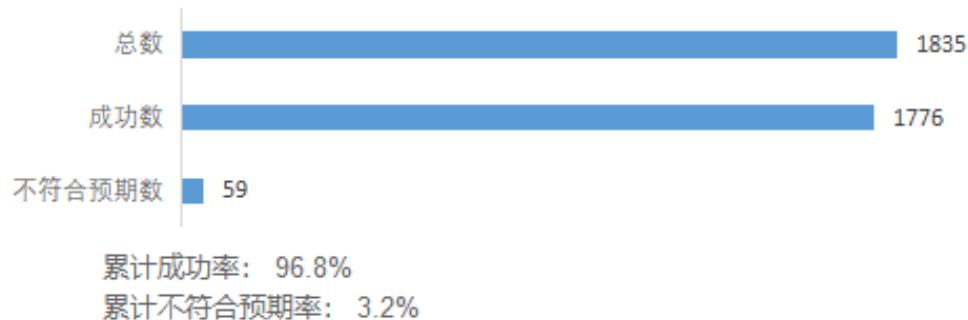
完成平台化故障演练  
2087次

好

发现23类技术风险、落实  
300+个改进项



## 累计演练数据统计



# 5、未来规划

---

# 展望



01

双随机演练 (含安全防控)

04

六大类风险补全

02

红蓝对抗

05

架构风险自动感知

03

系统稳定性能力可视化

06

建立故障知识库, 实现故障处置手段智能推荐



# THANKS!

2021  
TRUSTED CLOUD  
SUMMIT

