

Gdevops

Global DevOps Summit

全球敏捷运维峰会

腾讯游戏混沌工程实践

吴召军



目录

什么是混沌工程

混沌工程平台建设

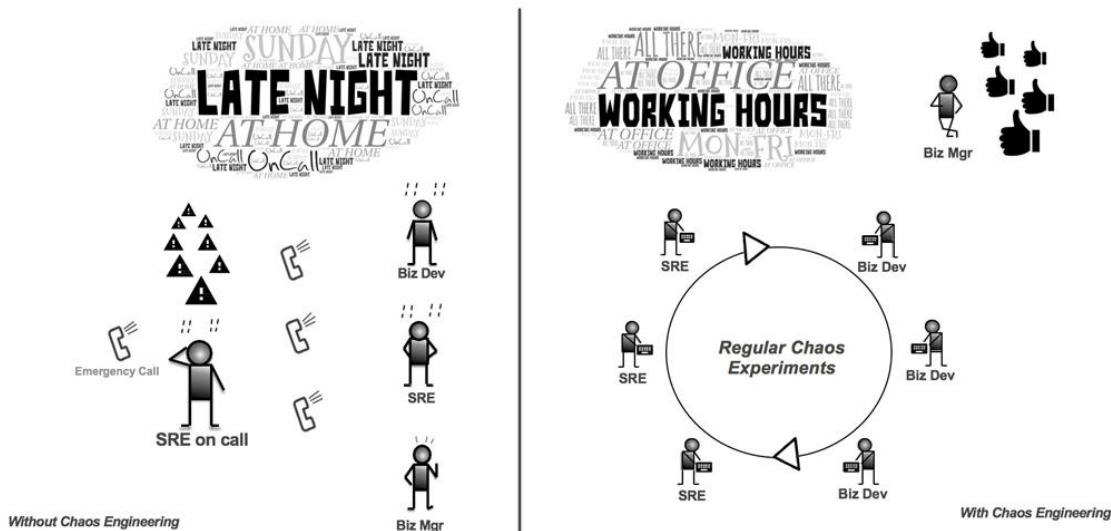
混沌工程实践案例



什么是混沌工程

主动在分布式系统上进行异常实验，观察系统行为，发现系统弱点与新知识，并持续优化和实验，不断提高系统容错能力，让人们建立复杂分布式系统能够抵御突发事件的信心。

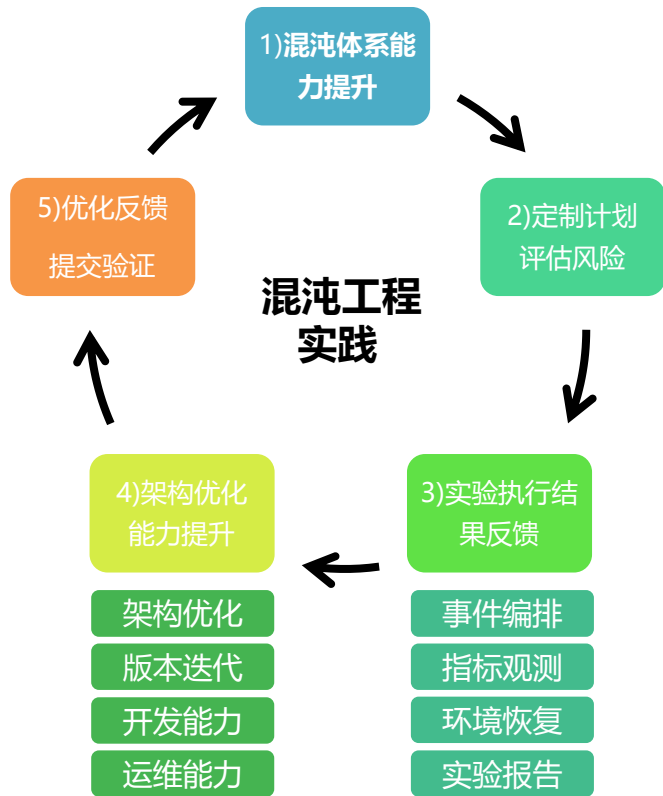
——混沌工程的定义



什么是混沌工程



混沌工程平台建设



1) 混沌体系能力提升

根据业务场景给混沌能力的反馈与要求，不断迭代建设Chaos体系。

2) 定制计划评估风险

从业务实际场景出发，设计具体实验计划，包括实验目标、范围、故障，选取稳态观察指标，限定爆炸半径，控制风险。

3) 实验执行并反馈结果

执行前检查事件编排，查看当前观测指标状态，确认无误后下发实验，实验过程中观测稳态指标表现，据此判断实验是否符合预期，实验结束后恢复环境，同时输出实验报告。

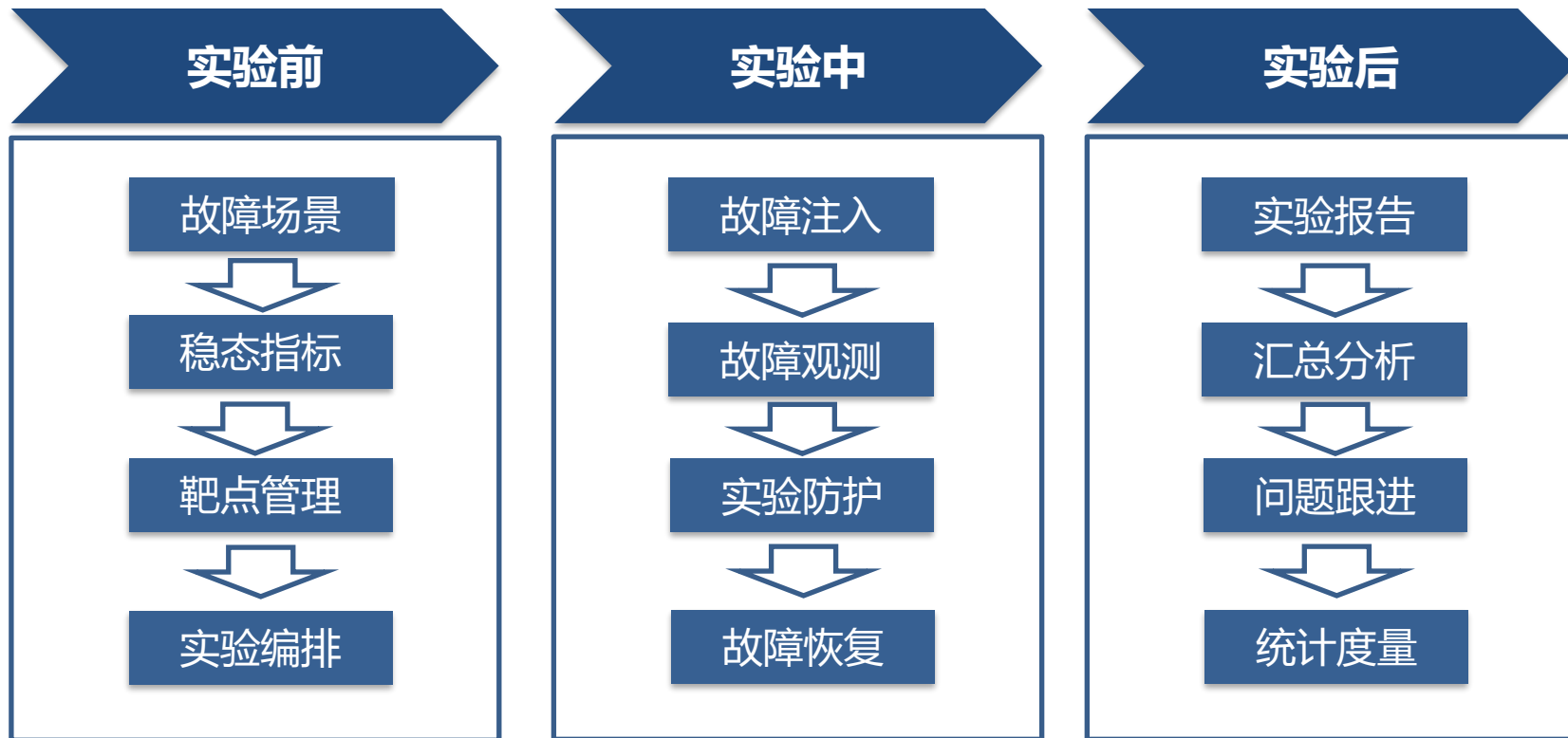
4) 架构优化与能力提升

业务相关干系人（运维、开发、测试等人员）收到结果反馈后需对已存问题进行review、评估整改方案、修复计划并检查同类问题，最后进行系统升级。

5) 优化反馈并提交验证

根据业务的优化反馈，再次提交实验请求，验证改进是否生效，进入下一轮混沌实验环节。

混沌工程平台建设：流程设计



混沌工程平台建设：故障原子

存储	计算资源	网络	节点/容器	应用	自定义
<ul style="list-style-type: none">• 磁盘空间满• IO高负载• IO延迟• IO错误• 文件句柄耗尽• 文件删除	<ul style="list-style-type: none">• CPU负载高• 单核CPU负载高• 内存满• 应用内存满• 虚拟内存满• GPU负载高	<ul style="list-style-type: none">• 丢包、延迟• 乱序、重复• 带宽满• 限速• 端口耗尽• DNS篡改	<ul style="list-style-type: none">• 关机X秒• 开机• 重启/销毁重建• 删pod• 杀容器• 杀pod	<ul style="list-style-type: none">• 杀进程• 进程僵死• 状态码错误• Body篡改• 请求延迟• 访问量激增	<ul style="list-style-type: none">• shell脚本• Python脚本• 二进制包• API插件• 开源工具接入

自研CHAOS TOOL

开源CHAOS TOOL

网关MESH

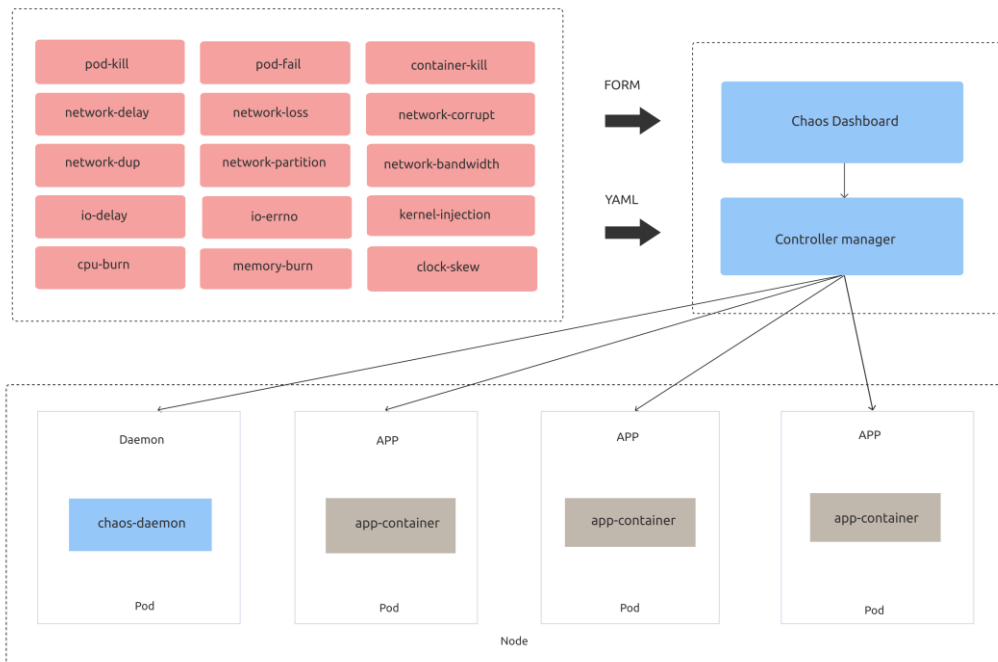
混沌工程平台建设：容器故障注入



Chaos Mesh®

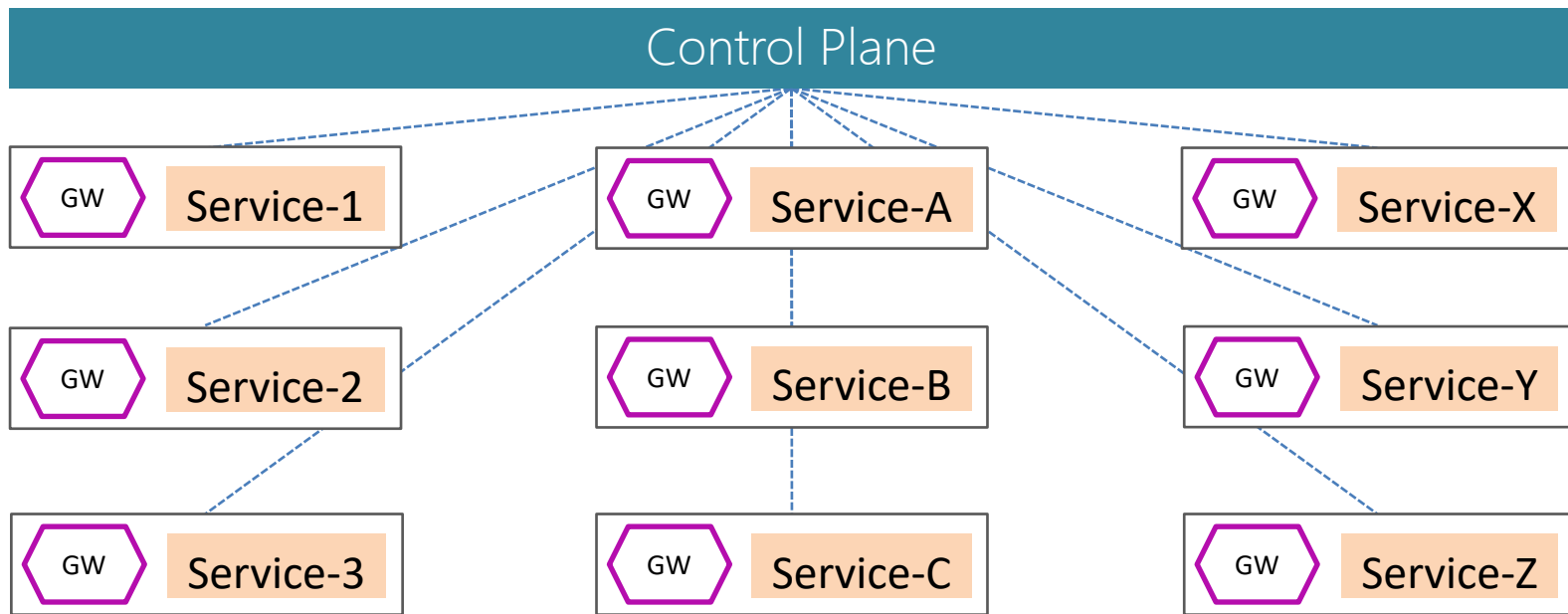


CLOUD NATIVE
COMPUTING FOUNDATION



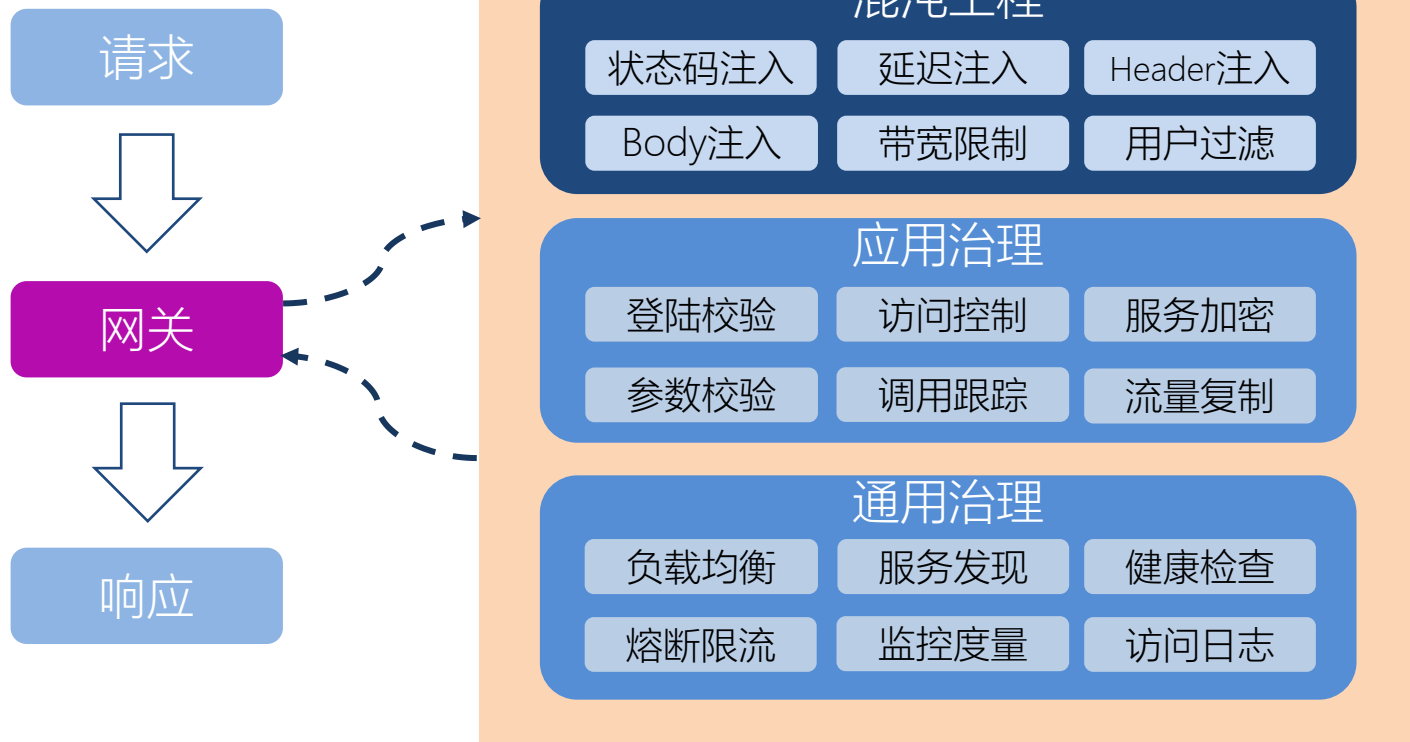
<https://chaos-mesh.org/>

混沌工程平台建设：应用故障注入



网关MESH：服务治理，链路跟踪。

混沌工程平台建设：应用故障注入



应用层故障注入

▲ 参数位置

参数所在位置

query_parameter

参数名

ID

规则类型

in

规则值

10001,123456

▲ 延迟注入

对服务注入一定的延迟

固定延迟

3s

▲ 注入比例

延迟注入占总请求量的比例

比例分母

HUNDRED

比例分子

100

混沌工程平台建设：实验编排



> 实验原子配置

CPU负载百分比 (%) * 80

运行时长(秒) * 600

高级选项(可选) ☐

> 实验原子配置

延迟时间(毫秒) * 1000

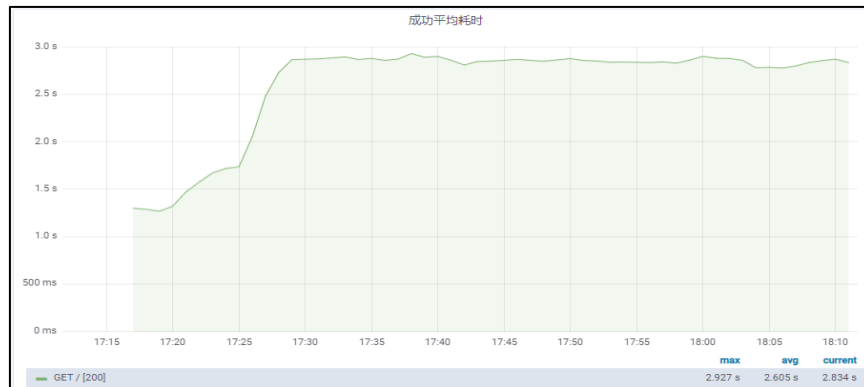
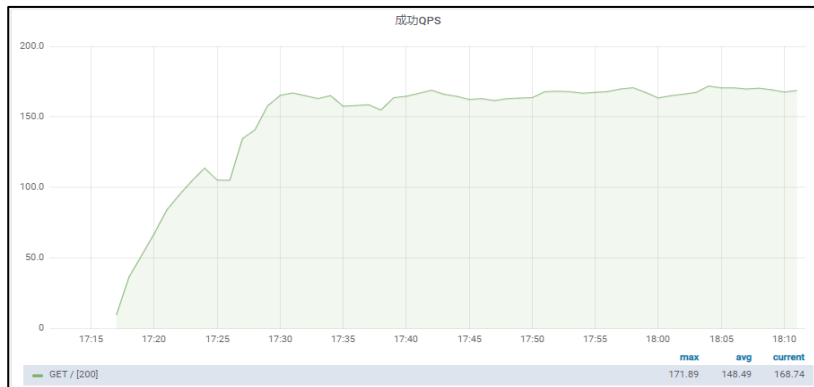
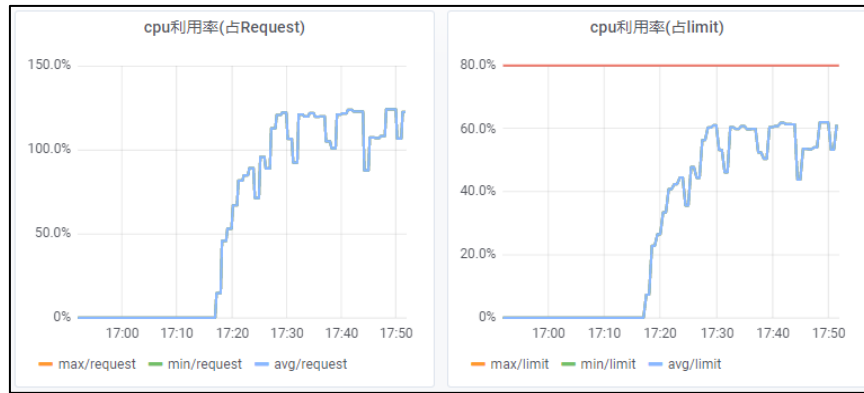
运行时长(秒) * 600

网卡设备 (如eth1) * eth1

高级选项(可选) ☐

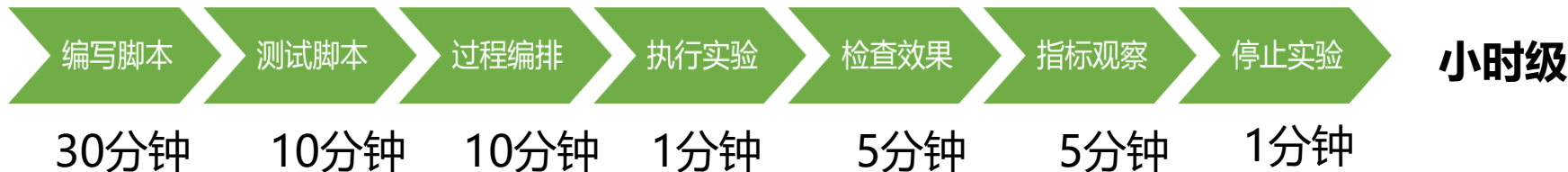
混沌工程平台建设：实验观测

- 基础监控系统
- 业务监控系统
- Prometheus



混沌工程平台建设：收益

传统故障演练：



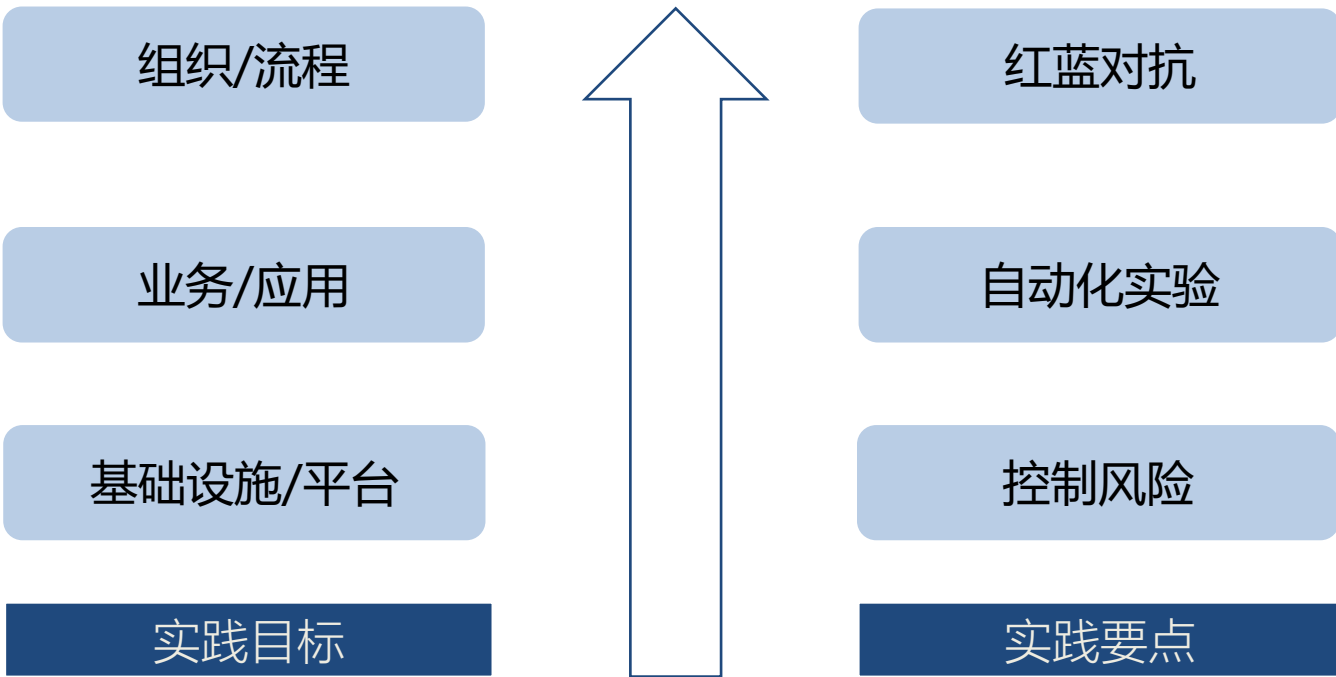
混沌平台演练：



功能齐全的自助化和自动化工具



混沌工程实践



环境递进

演习环境，预生产环境，生产环境

范围递进

指定用户，指定微服务，指定大区

实验防护

一键终止，触发阈值自动终止

时机选择

避免流量高峰期，人齐且通知到位



切尔诺贝利核反应堆废墟

混沌工程实践：风险控制

演习环境

可控性：高

不会影响线上业务

演习有效性：中

与线上完全隔离，有效性一般

人力成本：低

可自动化，无需人力参与

预发布环境

可控性：较高

对线上业务基本上不会有影响

演习有效性：较高

接近线上的环境，效果较为有效

人力成本：低

可自动化，无需人力参与

生产环境

可控性：低

风险系数高，可能造成业务损失

演习有效性：高

演习效果真实有效

人力成本：高

演习需要多方人力参与

结论：演习环境自测 + 预发布环境（每日自动化实验） + 生产环境（半年一次）相结合

混沌工程实践：实验防护

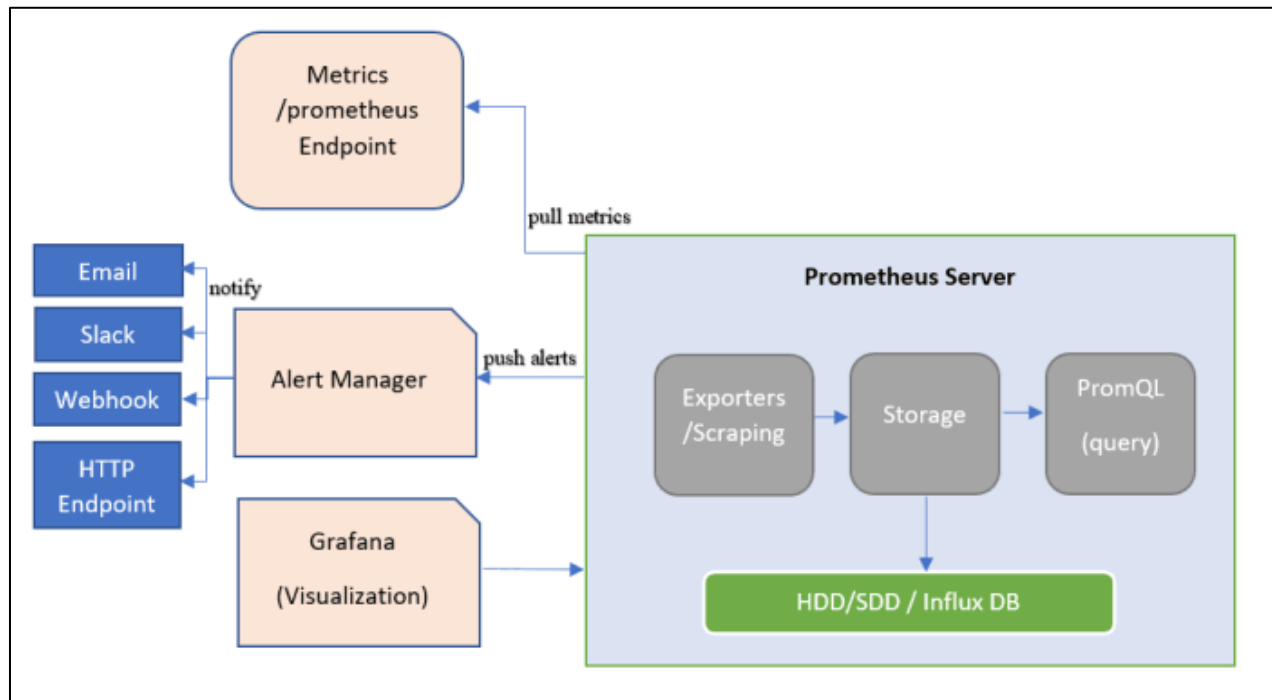
登记exporter, 上报稳态指标到混沌实验平台



配置熔断规则（阈值、自定义PROMSQL）

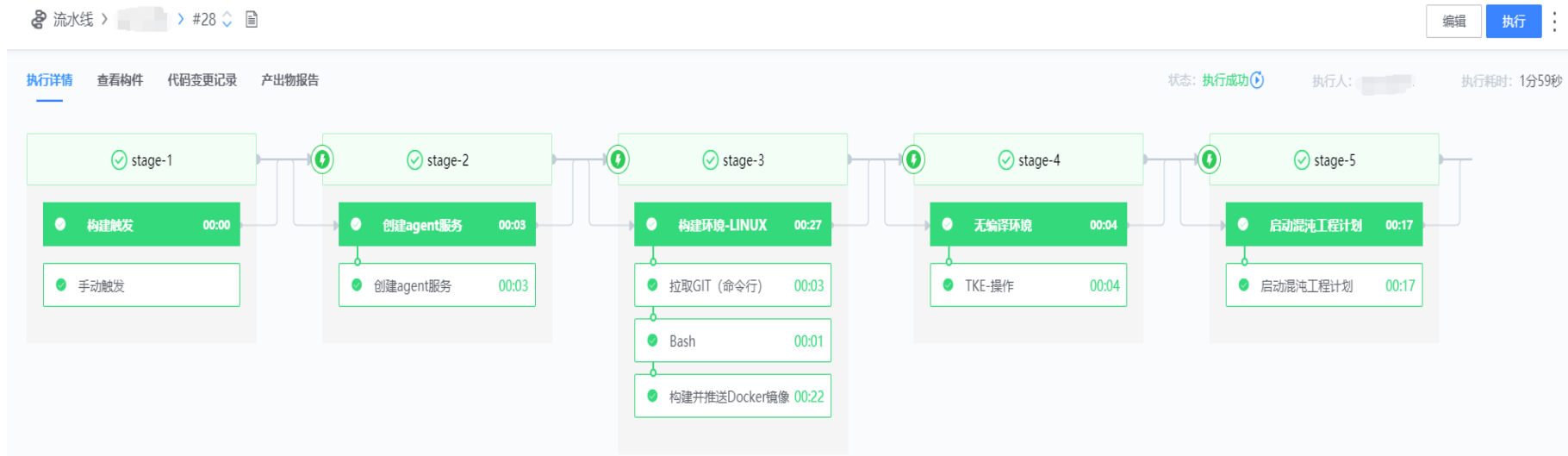


超过阈值，触发熔断，立即终止实验



图片来自网络

混沌工程实践：自动化混沌实验



发布流水线集成混沌实验套餐，自动引用、执行。

混沌工程实践：实验内容

实验分类	实验目标
单点故障	检测故障隔离、主备切换、健康探针有效性
告警验证	检测告警系统的有效性，检验组织协作响应机制
强弱依赖	检测不合理的依赖关系，验证高内聚低耦合架构
网络抖动	检测快速失败、失败重试策略
机房故障	检测异地容灾、故障隔离策略的有效性
第三方故障	检测降级、熔断策略以及本地缓存的有效性
过载保护	检测防刷、拒绝服务攻击、流控规则的有效性



全球敏捷运维峰会

THANK YOU !

