云原生架构下的混沌工程实践

周洋（花名：中亭）

阿里云智能事业群-高可用架构团队

GTLC
GLOBAL
TECH LEADERSHIP
CONFERENCE
全球技术领导力峰会

Geekbang. 极客邦科技 | TGO 鲲鹏會

500+高端科技领导者与你一起探讨
技术、管理与商业那些事儿

2019年6月14–15日 | 上海圣诺亚皇冠假日酒店

扫码了解更多信息

自我介绍

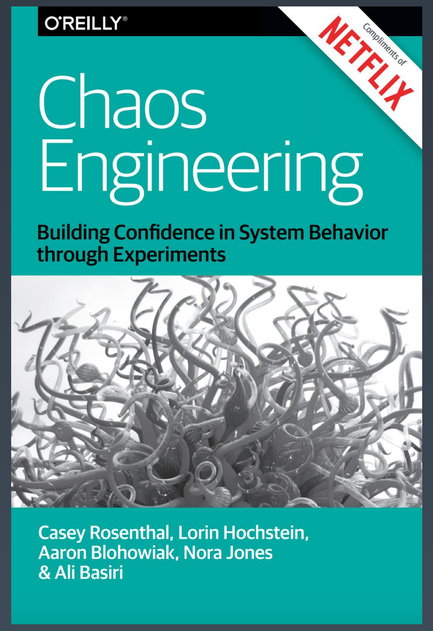
- 阿里高可用架构团队，花名中亭
- 多年高可用保障、产品研发和系统架构经验，2015年双11稳定性负责人
- 目前负责高可用技术云化输出，应用高可用服务（AHAS）、集团突袭演练负责人

- 2017 QCon明星讲师，《阿里电商故障治理和故障演练实践》
- 2019 混沌工程布道师，开源项目ChaosBlade发起人

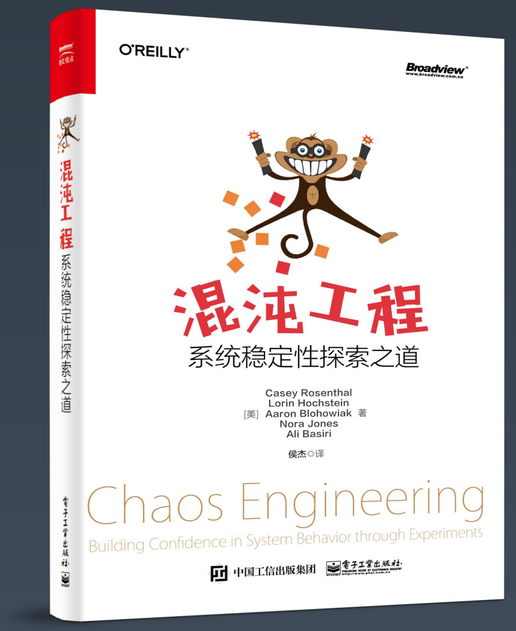
目录

- 混沌工程三连问
  - 混沌工程是什么?
  - 为什么要实施混沌工程?
  - 怎样实施混沌工程?
- 落地案例介绍
- 未来展望

混沌工程的概念



2017.8月出版



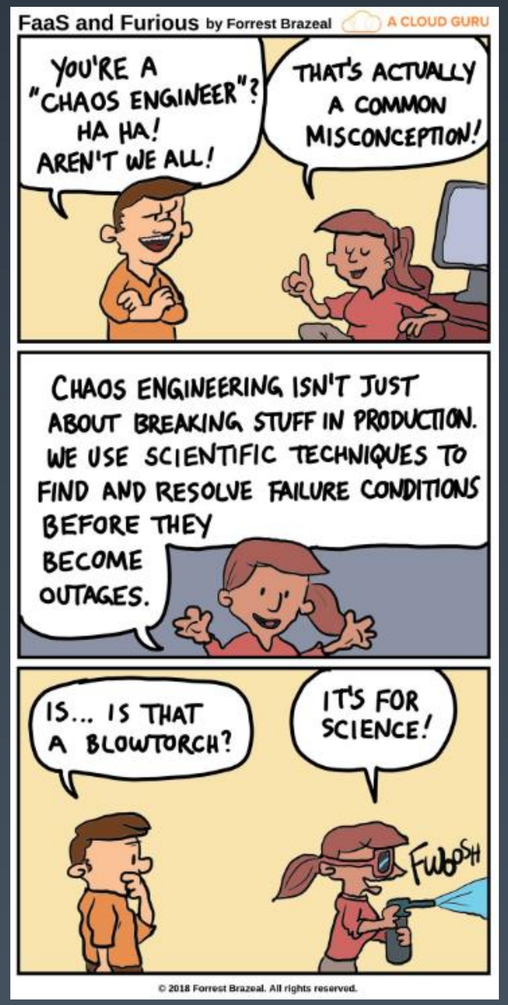
预计2019.6月上市

PRINCIPLES OF CHAOS ENGINEERING

Chaos Engineering is the discipline of experimenting on a distributed system
in order to build confidence in the system's capability
to withstand turbulent conditions in production.

混沌工程是在分布式系统上进行实验的学科，
目的是建立对系统抵御生产环境中失控条件
的能力以及信心。

我对混沌工程的理解



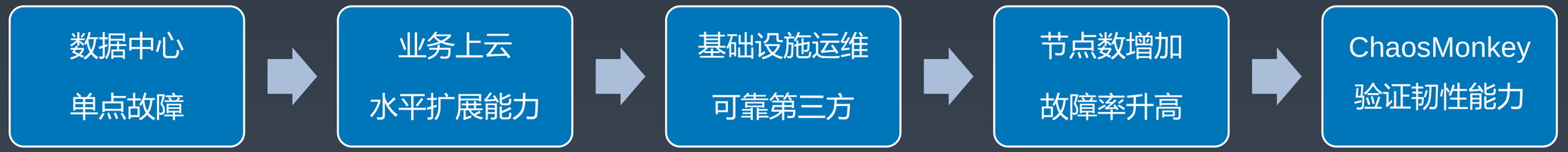
一种拥抱失败的技术文化

一套抽象严谨的实践原则

一种主动防御的稳定性手段

一个高速发展的技术领域

混沌工程的起源



源自《Chaos Engineering》书籍

在过去五年左右的时间里，只有仅有的一次节点掉线影响了我们的服务。当时正是混乱猴子终止了一个由于部署失误而没有冗余的服务节点造成了问题。幸运的是，这个故障发生在白天工作时间，在这个故障的服务刚刚部署不久后，对用户的影响也非常小。

混乱猴子的美妙之处就在于此，它能尽可能地将服务节点失效的痛苦提到最前，同时让所有工程师在构建一个具有足够弹性应对失败的系统上，达成一个一致的目标。

混沌工程原则

建立一个围绕稳定状态行为的假说

- 关注可测量输出，而不是系统内部属性。
- 短时间内的度量结果，代表了系统的稳定状态。
- 验证系统是否工作，而不是如何工作。

多样化真实世界的事件

- 混沌变量反映了现实世界中的事件。
- 通过潜在影响或预估频率排定事件的优先级。
- 任何能够破坏稳态的事件都是混沌实验中的一个潜在变量。

在生产环境中运行实验

- 系统的行为会根据环境和流量模式有所不同。
- 为了保证系统执行方式的真实性与当前部署系统的相关性，混沌工程强烈推荐直接采用生产环境流量进行实验。

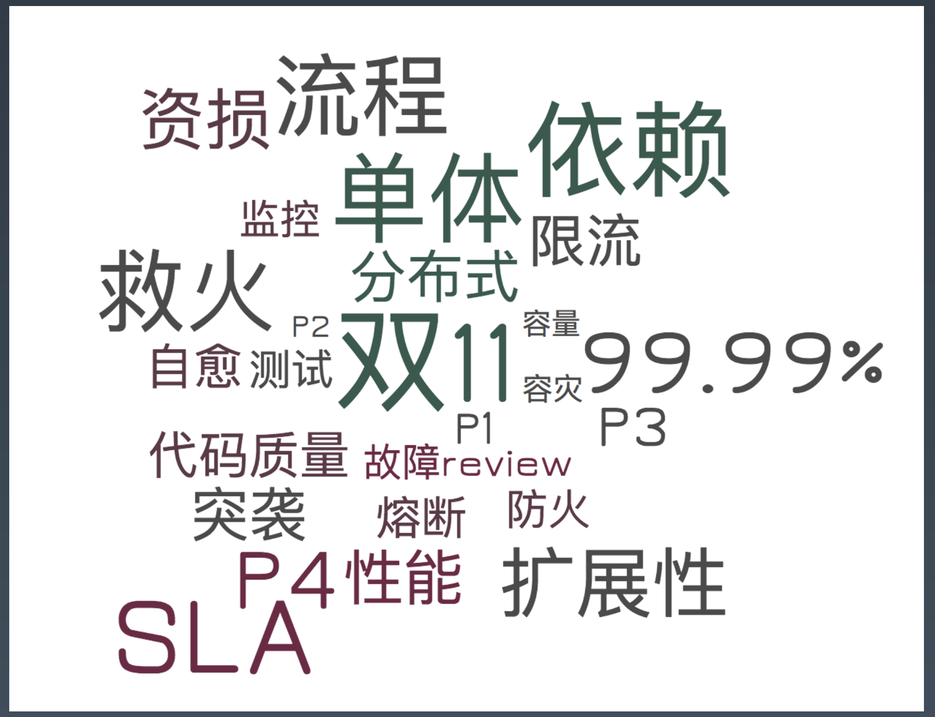
持续自动化运行实验

- 手动运行实验是劳动密集型的，最终是不可持续的，所以我们要把实验自动化并持续运行。
- 混沌工程要在系统中构建自动化的编排和分析。

最小化爆炸半径

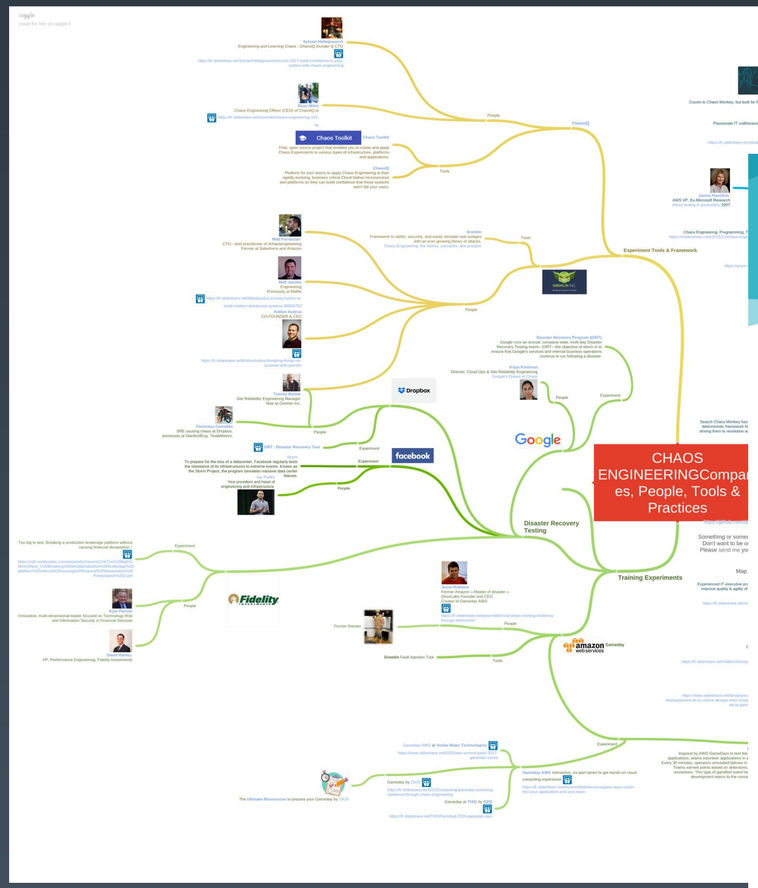
- 在生产中进行试验可能会造成不必要的客户投诉。但混沌工程师的责任和义务是确保这些后续影响最小化且被考虑到。

从"故障驱动"到"故障"驱动


资损
流程
依赖
单体
监控
限流
救火
分布式
P2
双11
容量
自愈
测试
容灾
99.99%
P1
P3
代码质量
故障review
突袭
熔断
防火
P4
性能
扩展性
SLA


监控
发现
持续
集成
限流
预案
架构
治理
混沌工程
资损
防控
故障管理
活动
保障
故障
应急
稳定性
度量

一个高速发展的技术领域



技术

采用

1. Lightweight Architecture Decision Records

试验

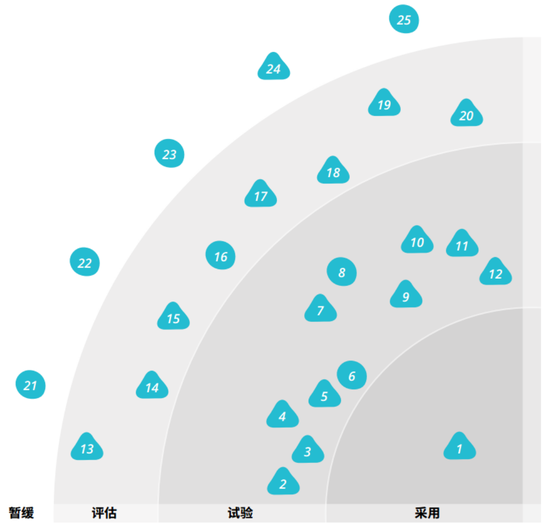
2. Applying product management to internal platforms *NEW*
3. Architectural fitness function *NEW*
4. Autonomous bubble pattern *NEW*
5. Chaos Engineering *NEW*
6. Decoupling secret management from source code
7. DesignOps *NEW*
8. Legacy in a box
9. Micro frontends
10. Pipelines for infrastructure as code *NEW*
11. Serverless architecture
12. TDD'ing containers *NEW*

评估

13. Algorithmic IT operations *NEW*
14. Ethereum for decentralized applications *NEW*
15. Event streaming as the source of truth *NEW*
16. Platform engineering product teams
17. Polycloud *NEW*
18. Service mesh *NEW*
19. Sidecars for endpoint security *NEW*
20. The three Rs of security *NEW*

暂缓

21. A single CI instance for all teams
22. CI theatre
23. Enterprise-wide integration test environments
24. Recreating ESB antipatterns with Kafka *NEW*
25. Spec-based codegen



小结

混沌工程

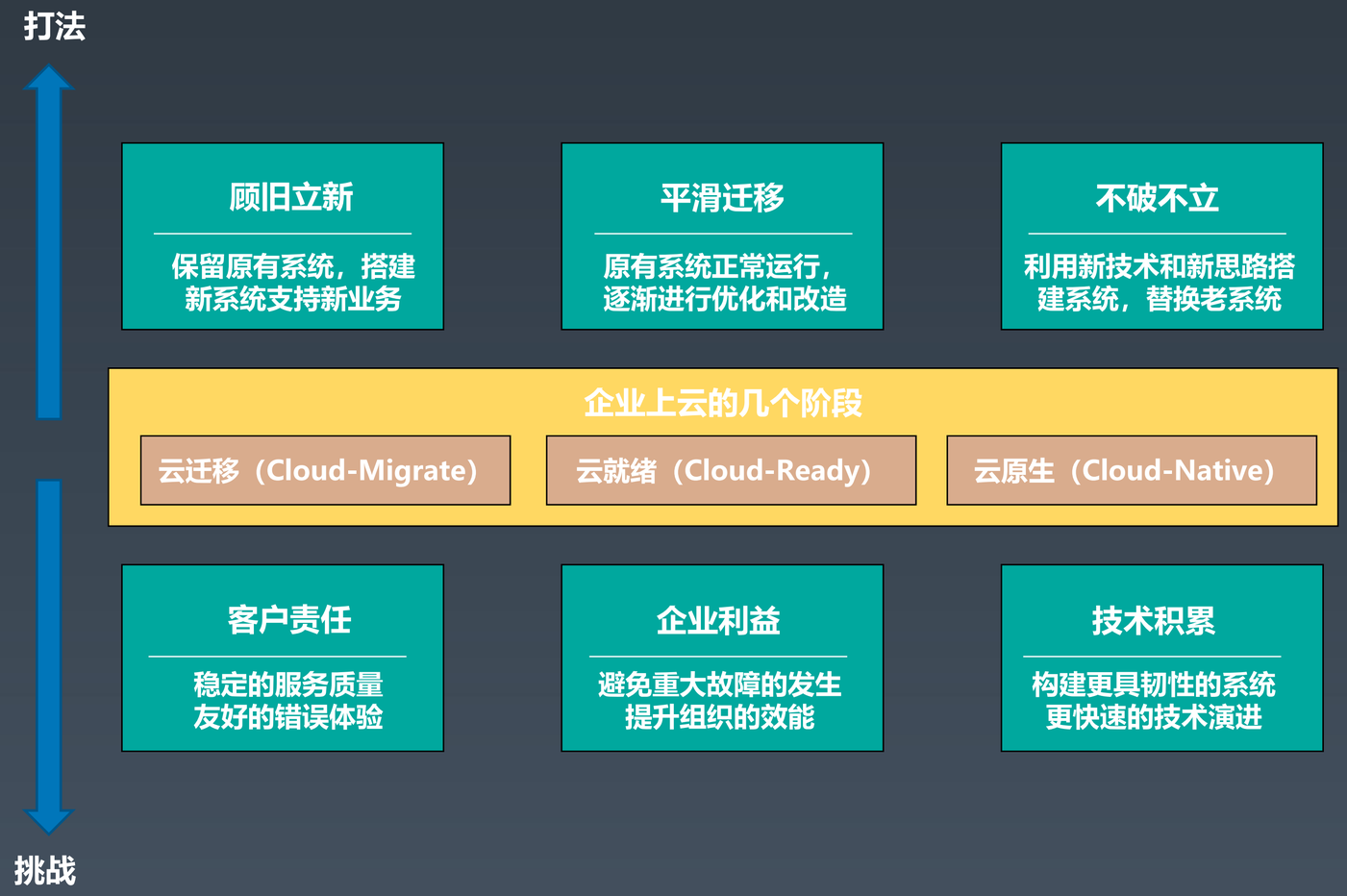
作为一个蓬勃发展的技术领域，

体现了一种反脆弱的技术思想，

提供了一套严谨的实践原则，

帮助企业更主动的提升稳定性

为什么要实施混沌工程



减小业务损失，让重大风险在可控范围提前暴露

提升系统弹性，持续验证系统对极端场景的容错能力

增强团队信心，验证稳定性措施有效性，量化团队价值

企业如何开始实施混沌工程?

- 混沌工程的引入（0 -> 1）

  结合技术架构，选择实验工具

  最小爆炸半径，控制实验风险

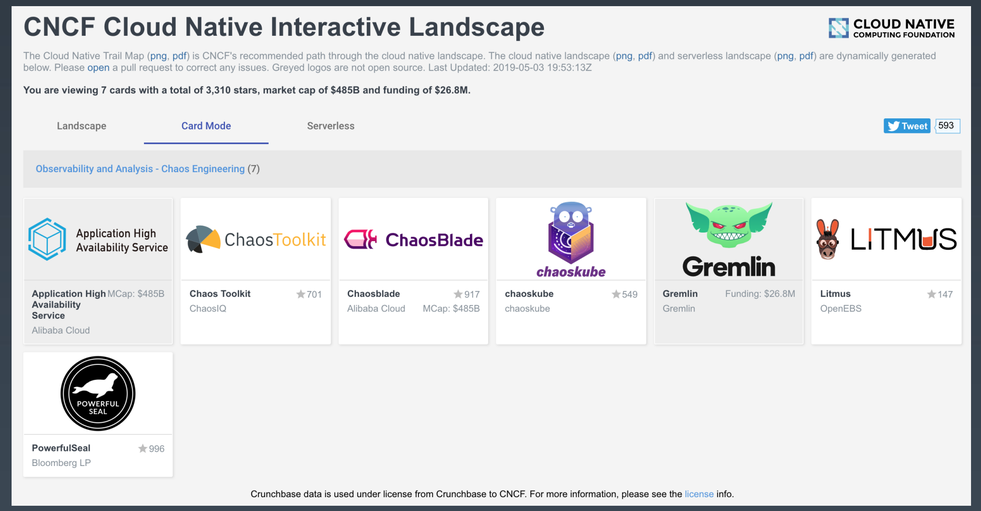
- 混沌工程的推广（1 -> N ）

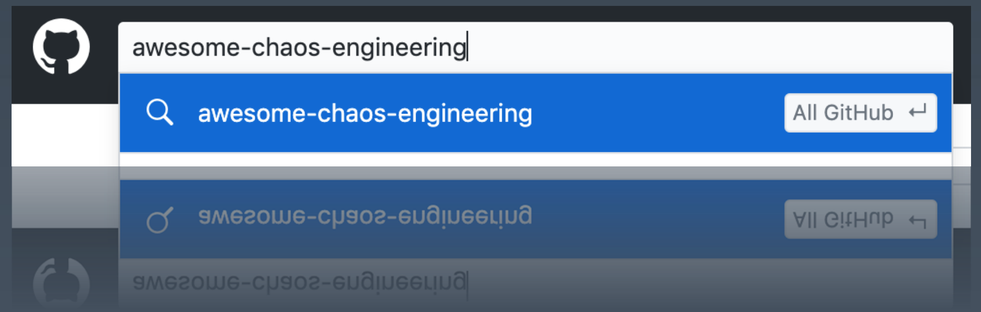
  建立面向失败设计的技术文化

  围绕战略制定目标，围绕目标设计组织

  复用成熟产品，提升效能

一款好的实验工具需要满足哪些条件?





- 丰富度
  资源、主机、容器、应用 ……
- 易用性
  开发框架、实验工具、产品平台
- 开放程度
  闭源、OpenAPI 、支持扩展、开源
- 集成方式
  代码依赖、架构依赖、无依赖
- 多语言
  Java、Go、C++、语言无关 ……
- 活跃状态
  已停滞、维护、活跃

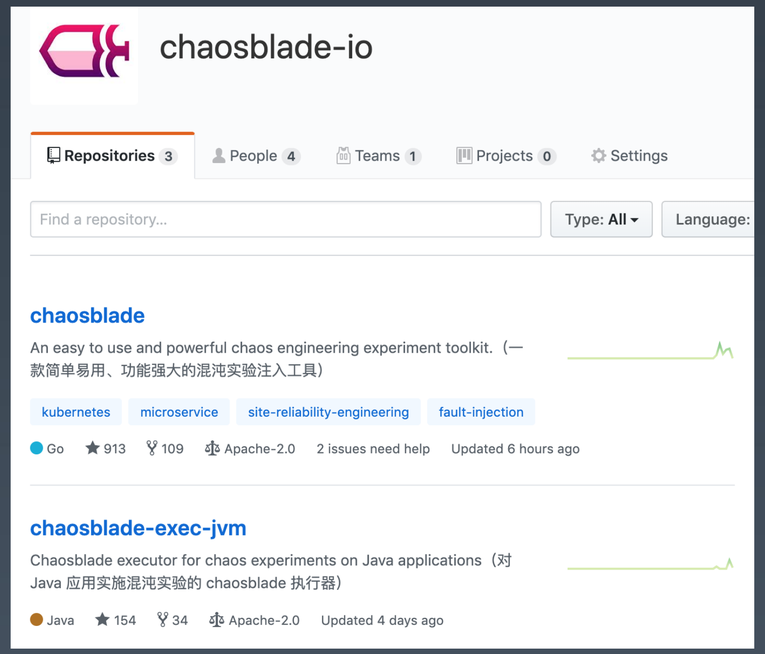
阿里混沌工程的技术演进路线

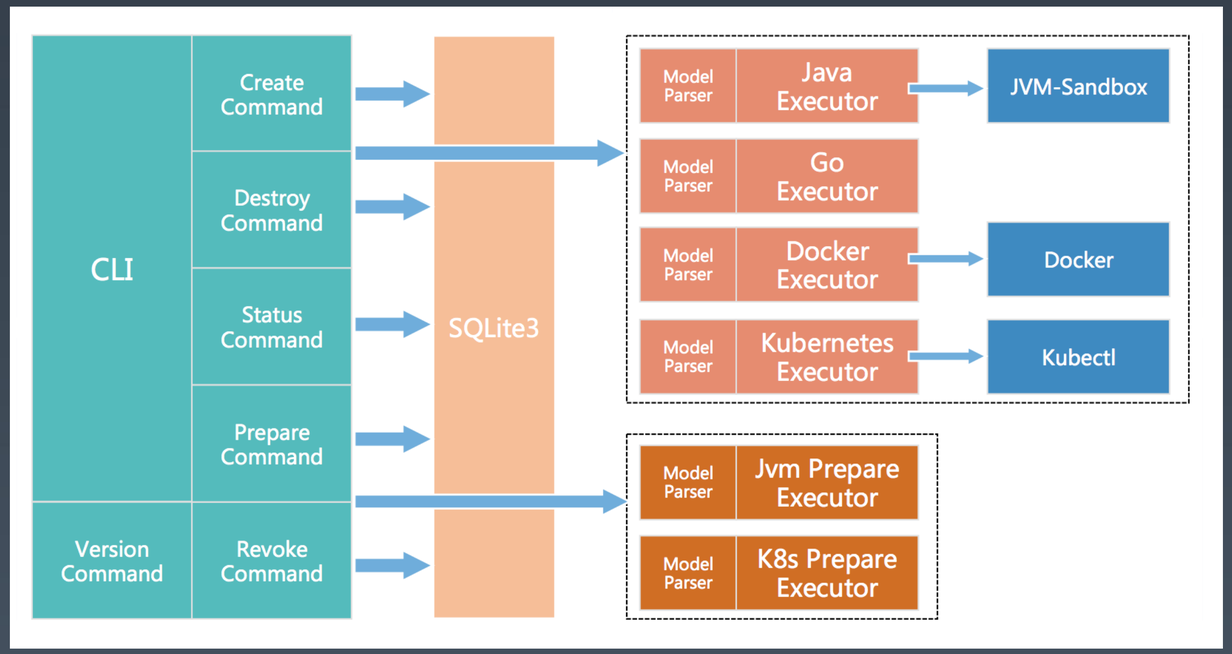
延伸阅读：《阿里电商故障治理和故障演练实践》

开源工具——混沌之刃（ChaosBlade）

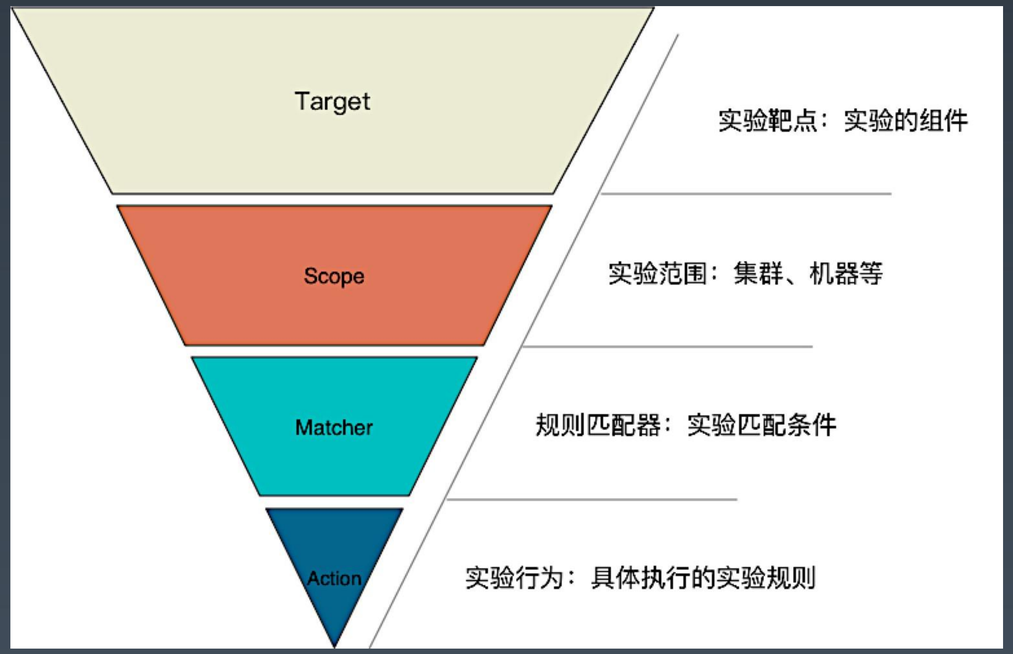
ChaosBlade是一款遵循混沌实验模型，提供丰富故障场景实现，帮助分布式系统提升容错性和可恢复性的混沌工程工具，它的特点是操作简洁、无侵入、扩展性强。

GitHub 地址：https://github.com/chaosblade-io





统一实验模型，描述混沌事件



实验示例：

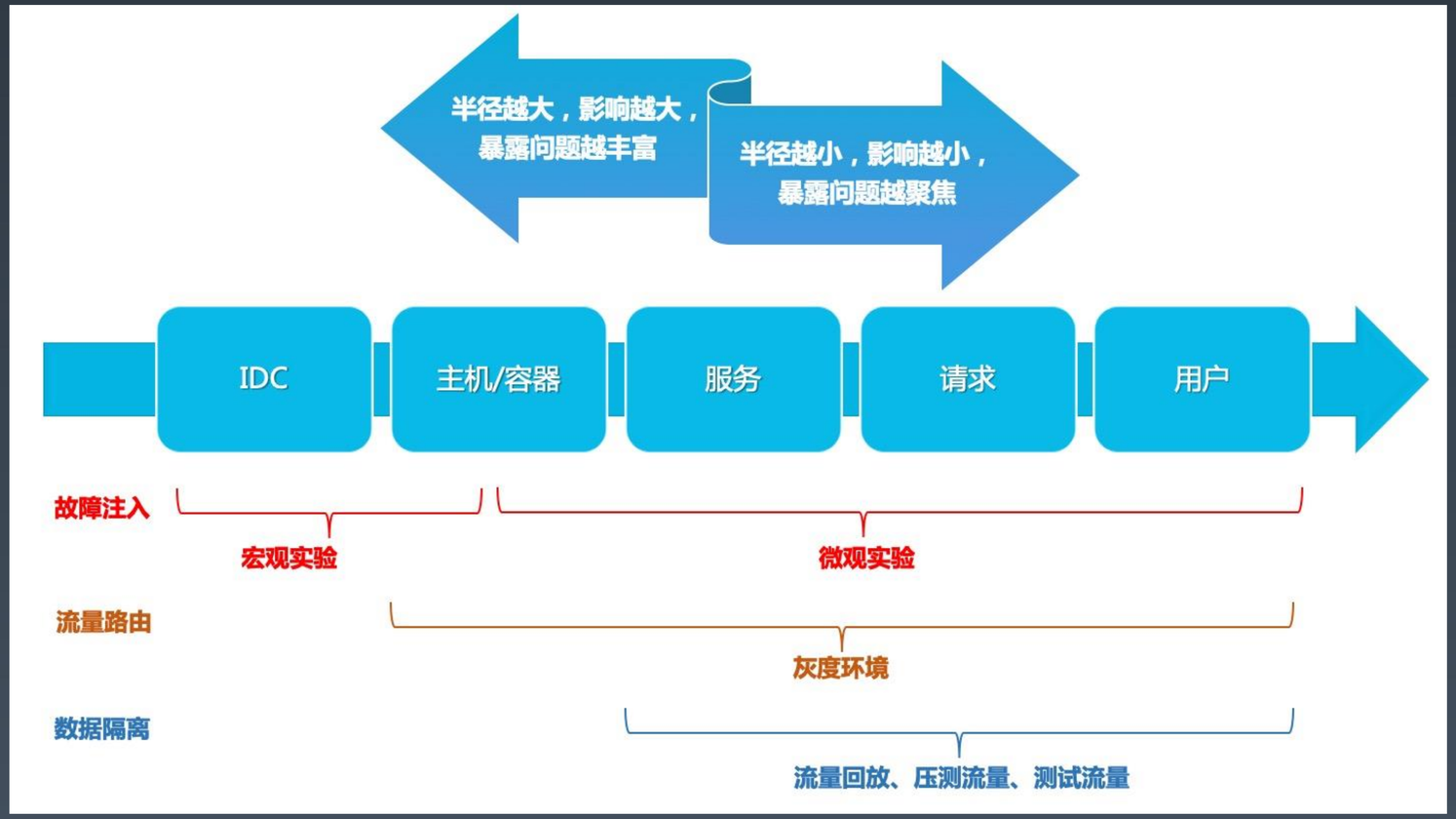
```
blade create dubbo delay --time 3000
--consumer
--service com.example.HelloService
--version 1.0.0

blade create cpu  fullload --cpu-count 4
```

结束示例：

```
blade destroy 6435335635bbaca5（实验ID）
```

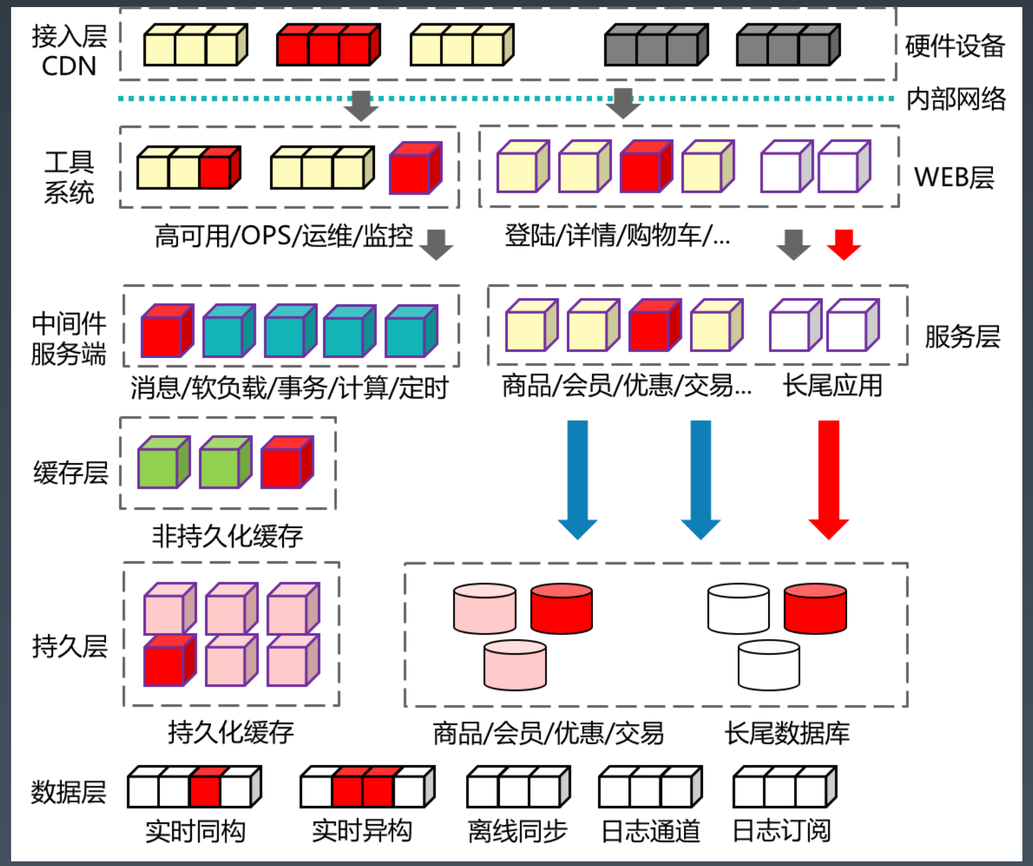
```
{
      "code": 200,
      "success": true,
      "result": "command: cpu fullload --cpu-
count 4 --debug false --help false"
}
```

控制爆炸半径，减小实施风险

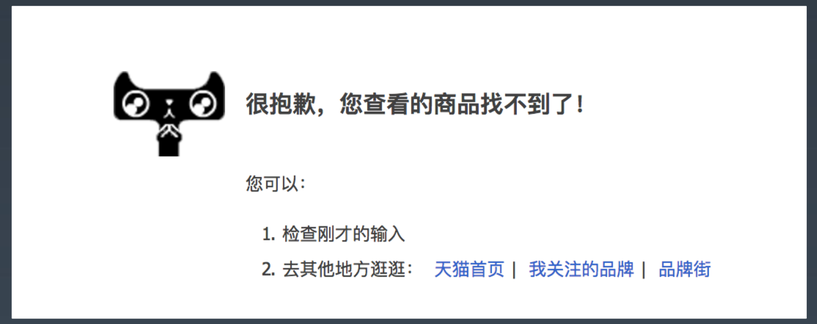


建立面向失败设计的技术文化

面向失败设计，因为每一块硬盘，每一个业务系统，每一种技术组件都有出错的可能！



分布式系统需要制定分级策略，防止非核心业务拖垮核心业务！！



工具系统需要优先实现容灾！！！

故障处理流程和人员能力也非常重要！！！！

围绕企业战略制定项目目标

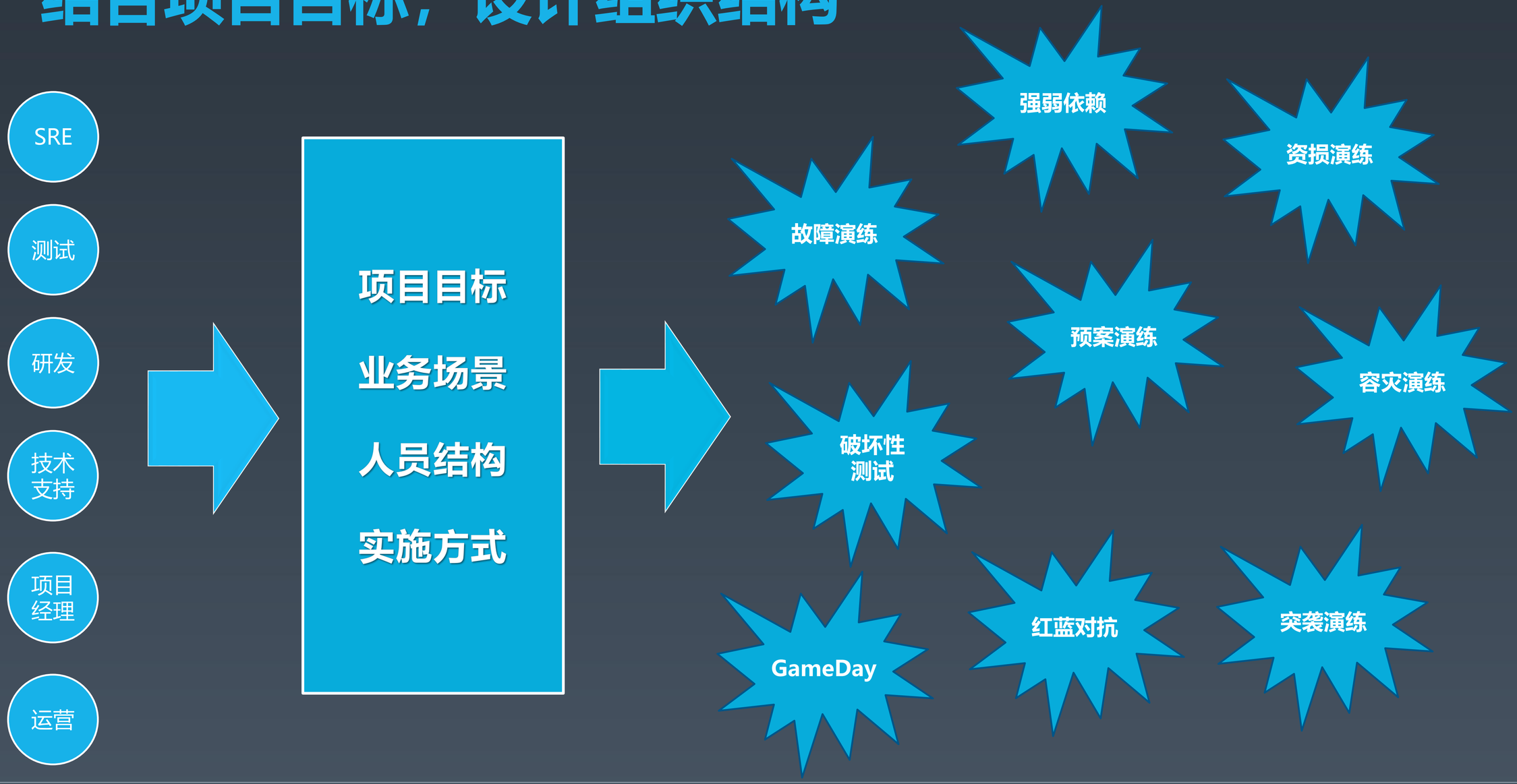
- 服务可用性
- 异常体验满意度
- API成功率
- …

- 资损故障预防
- 历史故障覆盖率
- …

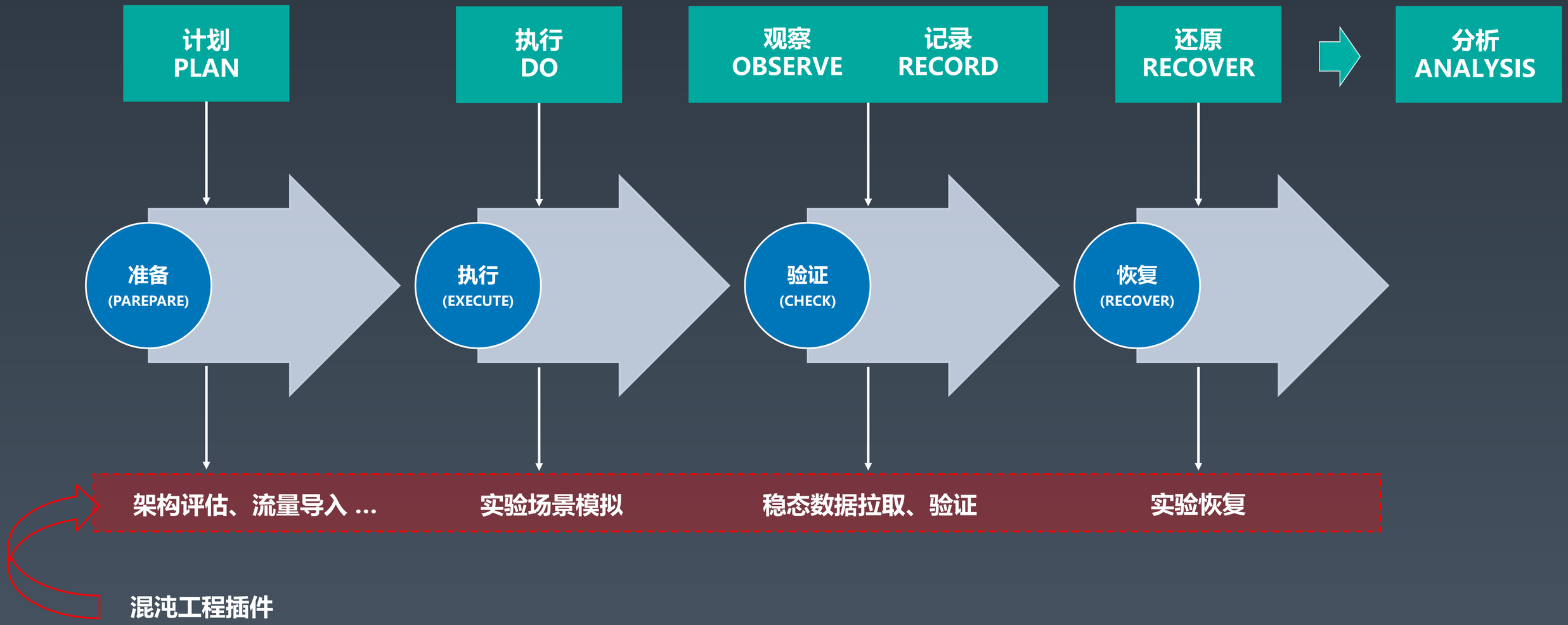
- 监控发现率
- 故障处理时长
- 预案有效性

- 架构容灾
- 依赖治理
- 故障自愈
- …

结合项目目标，设计组织结构



通过平台能力，标准化实验流程



建设实验平台，提升规模化能力

小结

引入和推广混沌工程，您需要

结合企业特点，选择适合当下的工具或产品；

最小爆炸半径，控制实验风险；

建立面向失败的技术文化，接受不确定性；

围绕企业战略，有针对性的设计组织和实施；

建设实验平台，提升规模化能力；

落地场景举例

新零售

云服务

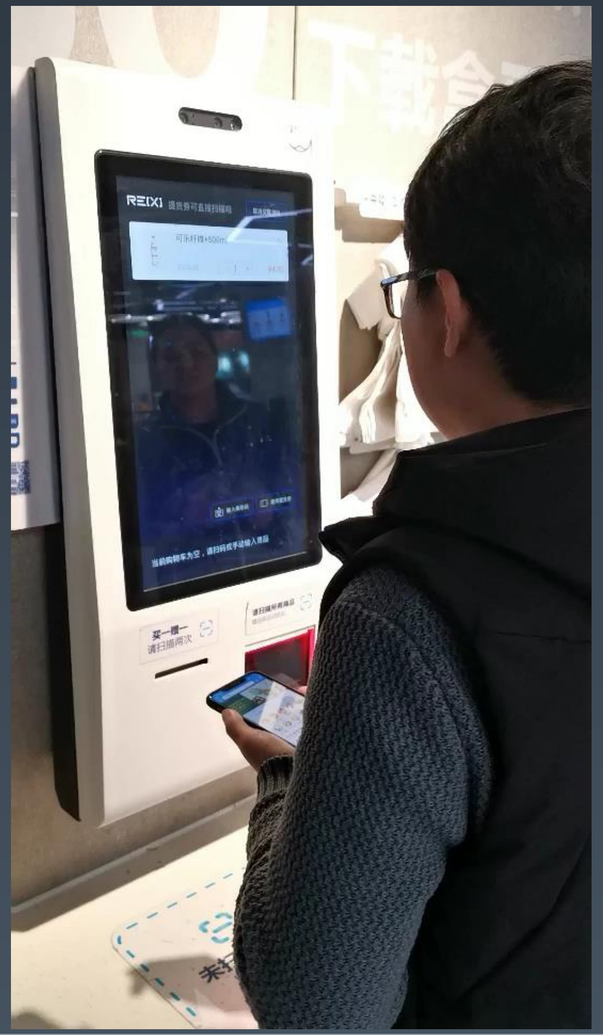
云业务

新零售业务稳定性的挑战

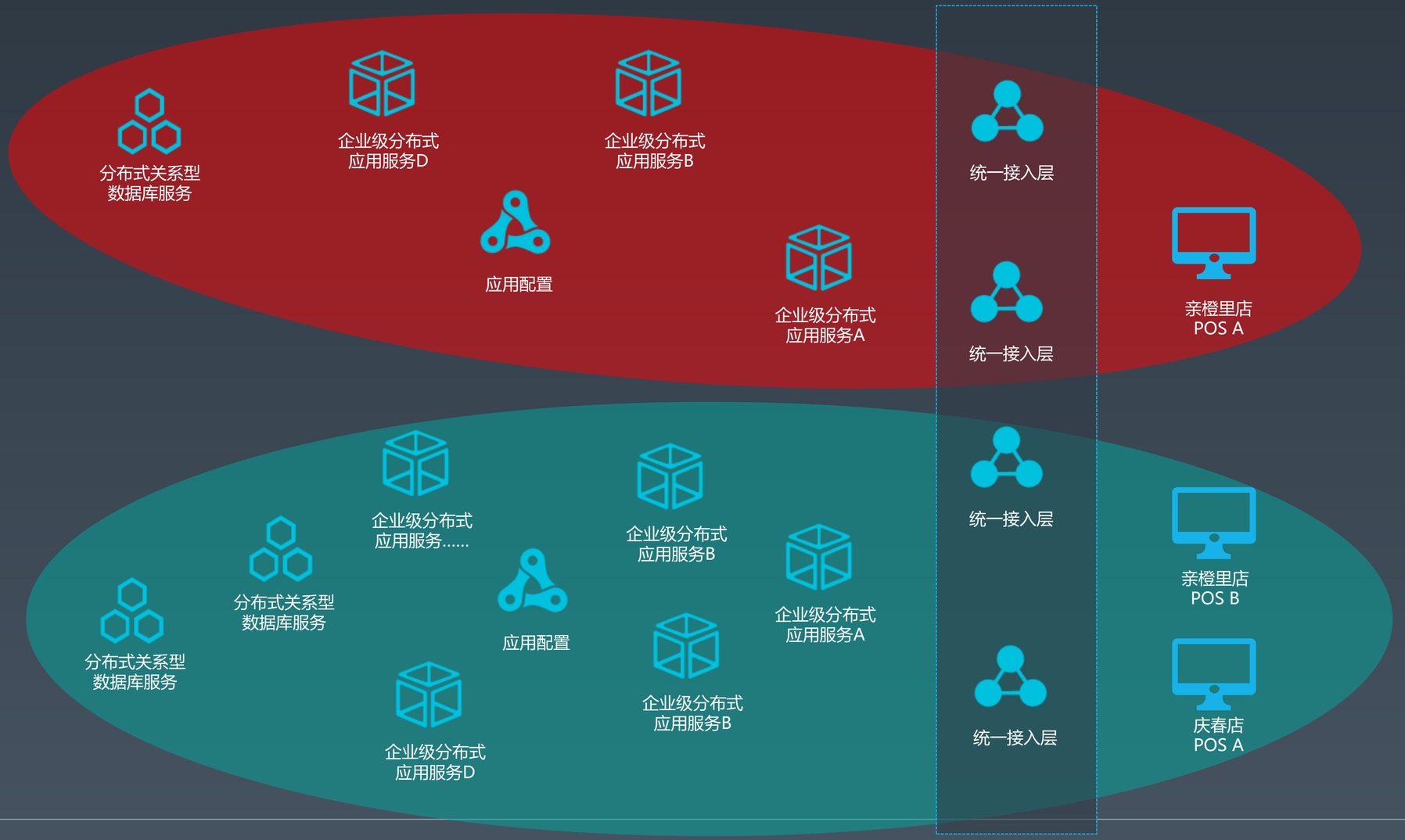
挑战
- 线下场景，用户对故障容忍度低
- 无法彻底规避网络问题，尤其是门店网络

要求
- 云端服务要具备较高的可用性
- 终端的异常提示要面向用户友好
- 现场人员要熟悉处理手段

难点
- 如何证明稳定性措施有效性？
- 如何减少实验对业务带来的影响？
- 如何常态化的实施实验？



最小化爆炸半径，实现常态化的实验



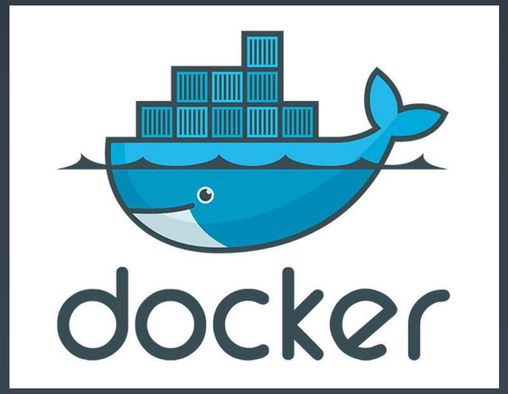
云服务稳定性——专有云混沌工程实践

运行态	运行时	服务强制退出	服务优雅退出	用户态CPU异常	OOM	
	服务交互	Request 参数空 / 特殊Query / 协议错误	用户流量徒增 / 雪崩	Connect 超时/失败	Connect 连接数满	Response 异常
	运行环境	K8S异常仿真	ECS IOhang 模拟	Docker异常仿真		
系统层	系统异常	CPU / MEM / 磁盘Inode 各类资源耗尽	文件系统ReadOnly	磁盘IO Hang < Gray failure >	系统Load高 < Gray failure >	
		NTP / YUM / DNS 异常	系统配置异常	系统权限异常	DPDK 网络异常	网络延迟、包乱序、丢包 < Gray failure >
	内核异常	内存页分配错误	内核futex死锁模拟 < Gray failure >	内核级函数错误模拟	Kernel Panic	系统参数异常
网络层	特定流不通	特定流五元组流量不通	特定IP流量不通 < Gray failure >	特定源和目的访问单通		
	黑洞/随机丢包	设备级别流量黑洞 < Gray failure >	SLOT级别流量黑洞	随机流量丢弃	网络设备单端口拥塞	
硬件层	服务器	服务器宕机 / 掉盘	PCIE Degrade < Gray failure >	NVME SSD 异常 < Gray failure >	RMDA 异常 < Gray failure >	SPDK 异常 < Gray failure >
	交换机	整机故障	异常重启	交换机 上行 / 下行 端口（RANGE）异常 < Gray failure >		

云业务稳定性保障的挑战

微服务

容器



开源组件

云服务

有什么

有哪些第三方组件，用了哪些云服务，有哪些自开发的应用服务，他们之间的关系是什么样的，他们和底层的容器，云服务器的关系是什么样子的？

做什么

我的第三方组件，云服务和应用服务需要具备哪些高可用能力

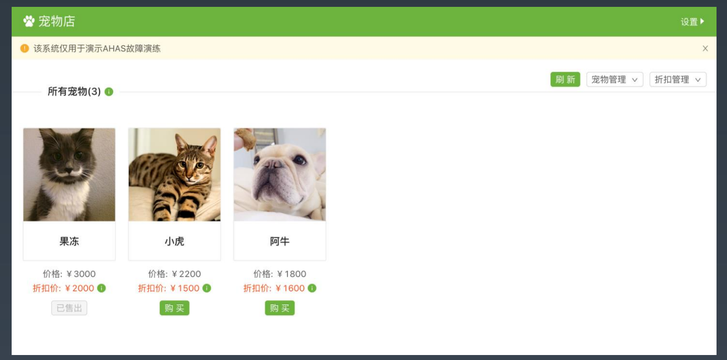
怎么做

如何提高这些组件的高可用能力

面向云业务的高可用服务

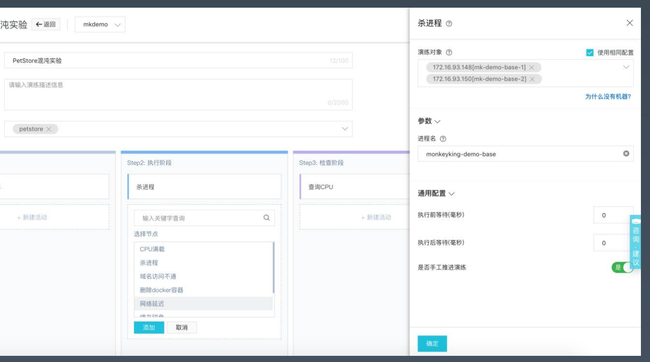
阿里云菜单搜索『 AHAS 』，免费公测中

云业务混沌实验方案

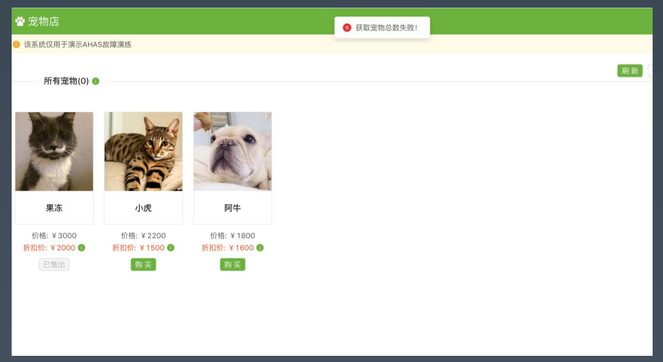


云原生业务（稳态分析）

架构感知&组件识别



白屏化式实验



云原生业务（稳态验证）

未来规划

方案

帮助云原生业务提升高可用能力的云服务（AHAS）

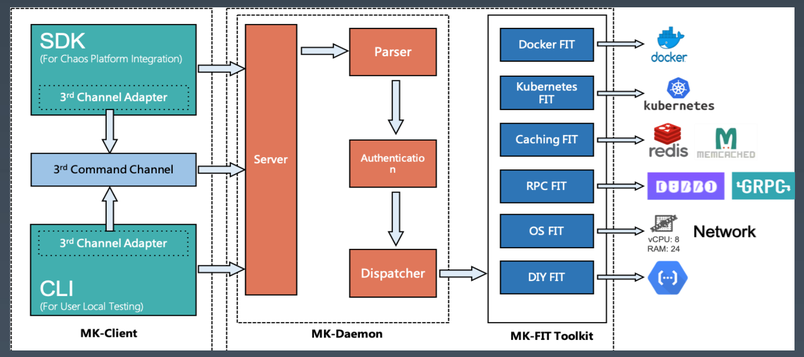
特性

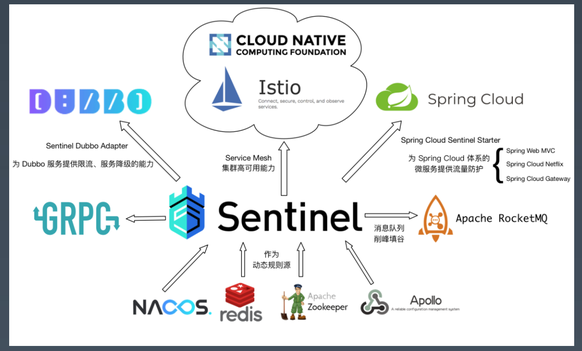
- 架构感知
- 故障演练
- 限流降级
- 高可用分析

连接

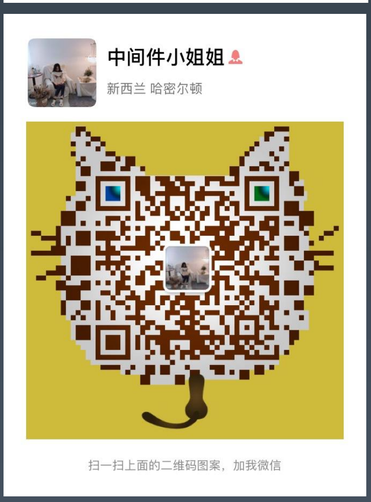
- 容器服务
- 日志服务
- 云监控
- 其他...

开源





社区建设



口令：chaos

Geekbang 极客邦科技 | TGO 鲲鹏會

TGO鲲鹏会

汇聚全球科技领导者的高端社群

全球12大城市

850+高端科技领导者

使命 Mission

为社会输送更多优秀的科技领导者

愿景 Vision

构建全球领先的有技术背景优秀人才的学习成长平台

扫描二维码，了解更多内容

Geekbang. 极客邦科技 InfoQ

极客邦科技
会议推荐2019

5月

QCon 北京
全球软件开发大会
大会：5月6–8日
培训：5月9–10日

QCon 广州
全球软件开发大会
培训：5月25–26日
大会：5月27–28日

6月

GTLC GLOBAL TECH LEADERSHIP CONFERENCE 上海
技术领导力峰会
时间：6月14–15日

GMTC 北京
全球大前端技术大会
大会：6月20–21日
培训：6月22–23日

7月

ArchSummit 深圳
全球架构师峰会
大会：7月12–13日
培训：7月14–15日

10月

QCon 上海
全球软件开发大会
大会：10月17–19日
培训：10月20–21日

11月

GMTC 深圳
全球大前端技术大会
大会：11月8–9日
培训：11月10–11日

AiCon 北京
全球人工智能与机器学习大会
大会：11月21–22日
培训：11月23–24日

12月

ArchSummit 北京
全球架构师峰会
大会：12月6–7日
培训：12月8–9日

THANKS!

QCon 10th