

字节跳动混沌工程实践

挑战与演进之路 / 童飞帆

我是谁

童飞帆

字节跳动架构工程师

金融业务开发 -> 监控领域 -> 混沌工程

专注于以混沌工程解决高可用问题

降低黑天鹅事件对分布式系统稳定性的影响

目录

1. 为什么&什么是混沌工程
2. 混沌工程实践中的挑战
3. 字节跳动混沌工程演进之路
 - ① 阶段一
 - ② 阶段二
 - ③ 阶段三

为什么&什么是混沌工程

为什么需要混沌工程

- 分布式系统复杂度急剧上升
- 故障发生的时间与区域无法预测
- 故障发生后对系统的影响很难评估
- 黑天鹅事件无法避免

正向推导能解决稳定性问题吗？

引入反向实验的方法？

什么是混沌工程

混沌工程是一门新兴的技术学科，他的初衷是通过实验性的方法，让人们建立对于复杂分布式系统在生产中抵御突发事件能力的信心。

—Netflix, 《Chaos Engineering》

什么是混沌工程

混沌工程实施原则

建立稳定状态的假设

多样化现实世界事件

在生产环境运行实验

持续自动化运行实验

最小化“爆炸半径”

混沌工程实践中的挑战

混沌工程实践中的挑战

- 为什么非得在生产环境实验？我的生产数据很关键。
- 万一我的系统真的挂了怎么办？
- 发生了重大生产事故怎么办？
- 故障注入的影响是否可控？

- 一个集群有上千个实例，模拟故障成本很高。
- 服务涉及多个外部依赖，模拟故障需要多方配合。

- 系统上线一年非常稳定，还需要混沌工程吗？
- 只是验证了故障发生之后的系统表现，还有什么收益？

风险成本高？

时间成本高？

收益不明显？

字节跳动混沌工程演进之路

阶段一

字节跳动混沌工程演进之路 - 阶段一

目标

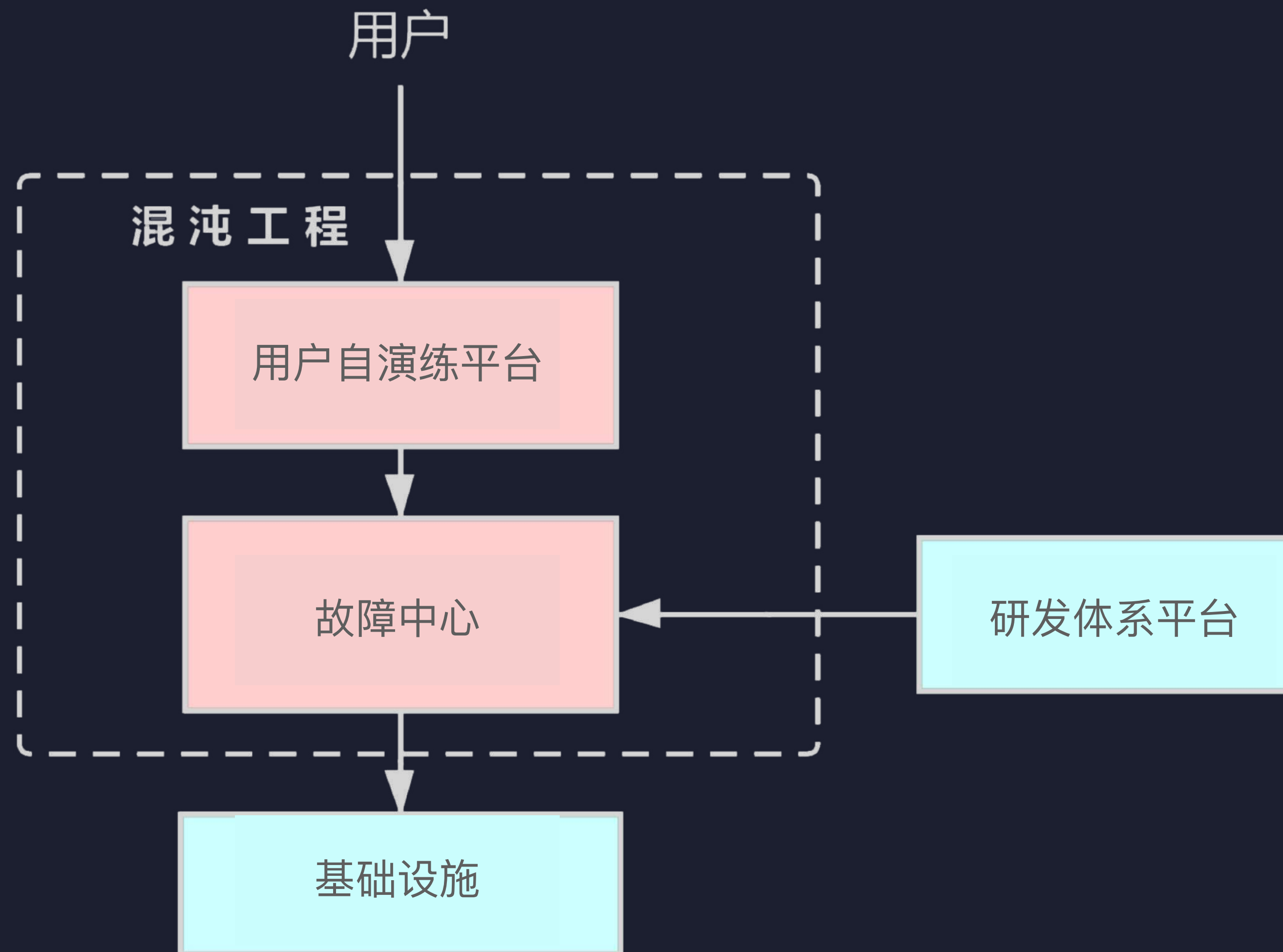
设计可扩展的故障中心，实现精准可控的爆炸半径控制。降低风险成本。

屏蔽业务系统部署细节，抽象故障模型，提供声明接口，使得故障注入简单明了。降低时间成本

解耦故障实现与实践活动管理。为探索更大的收益提供灵活性。

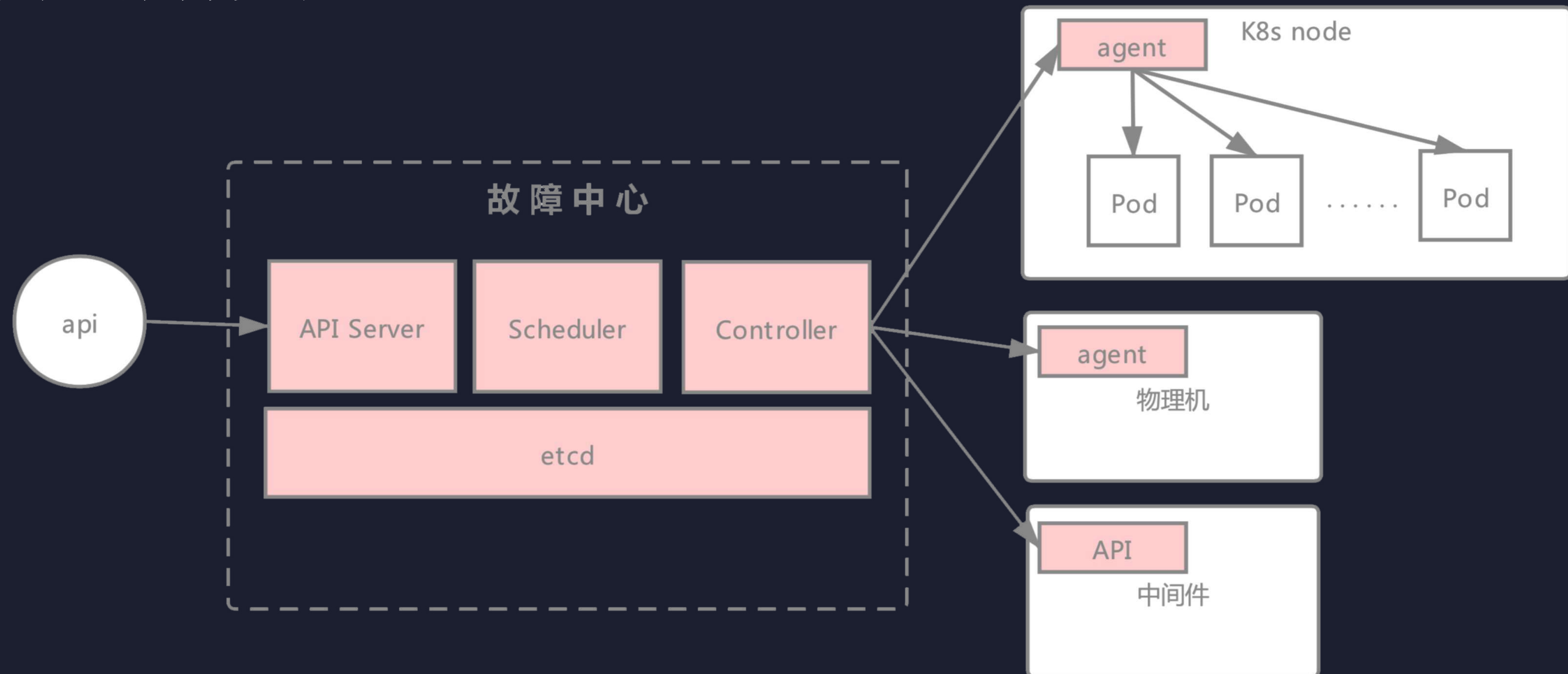
字节跳动混沌工程演进之路 – 阶段一

整体设计



字节跳动混沌工程演进之路 – 阶段一

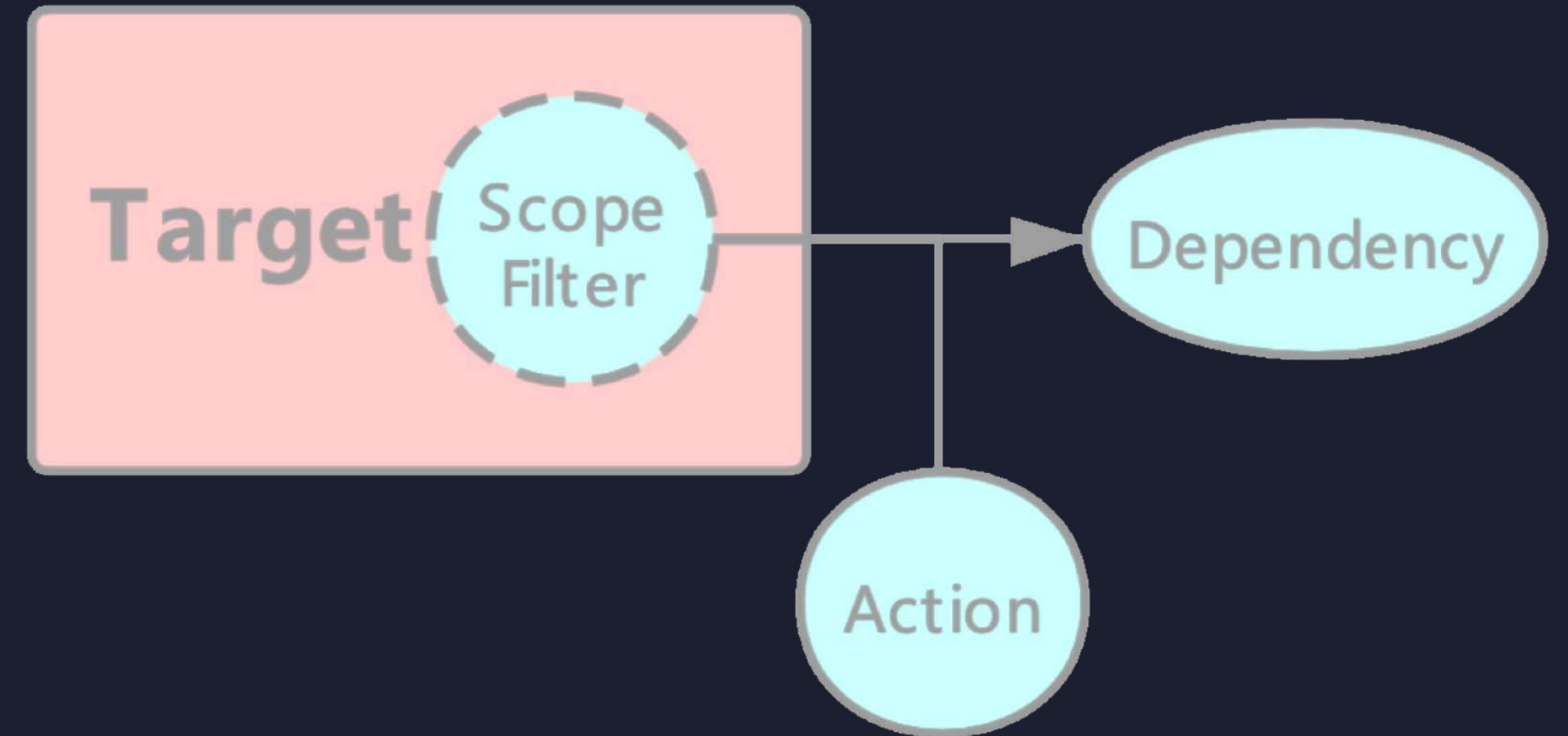
故障中心架构设计



字节跳动混沌工程演进之路 – 阶段一

故障模型设计

- Target – 目标服务，即被故障影响的服务，该服务为主要观察目标。
- Scope Filter – 对应混沌工程概念中的爆炸半径。通常为某一机房、集群，甚至精确到实例级别乃至流量级别。
- Dependency – 依赖，即所有可能影响服务本身的异常来源，包括中间件，某下游服务，也包括所依赖的cpu，磁盘，网络等。
- Action – 故障事件，即发生了何种故障，比如下游服务返回拒绝，发生丢包；又比如磁盘写异常，cpu被抢占等。

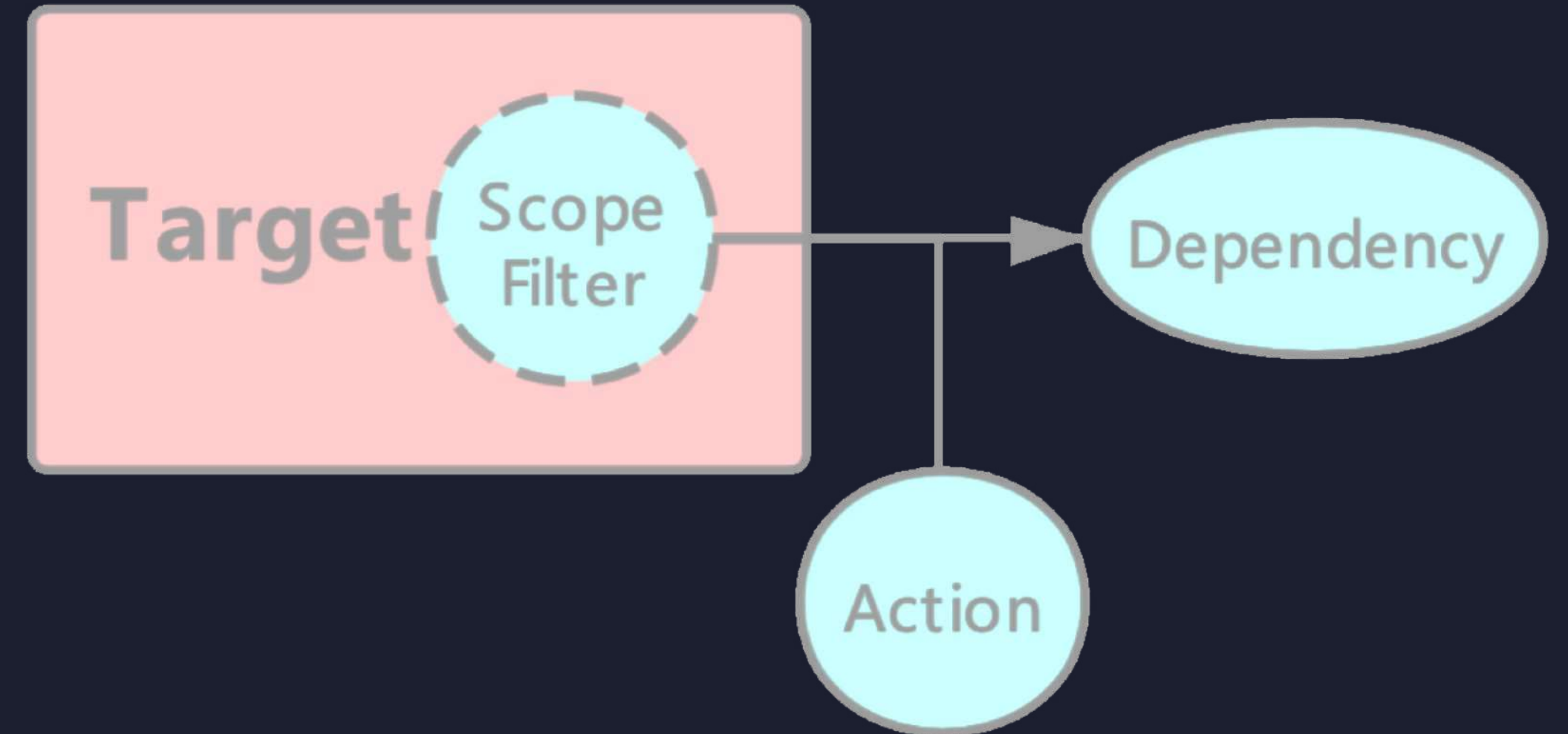


字节跳动混沌工程演进之路 – 阶段一

故障模型设计

故障声明伪代码

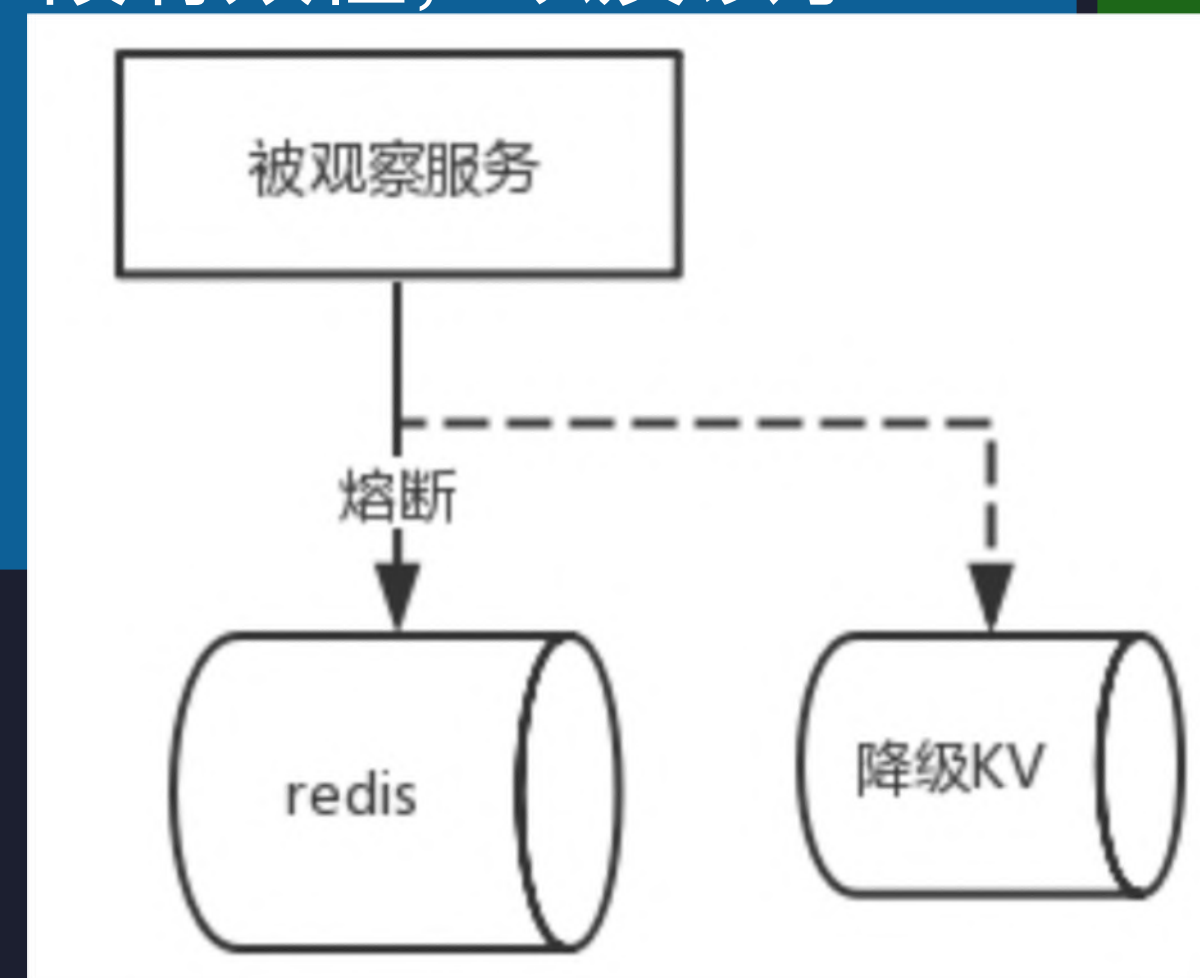
```
1  # 微服务application A 的 cluster1集群内10%的实例cpu突然满载
2  spec.
3      tareget("application A").
4          cluster_scope_filter("cluster1").
5              percent_scope_filter("10%").
6                  dependency("cpu").
7                      action("cpu_burn").
8                          end_at("2020-04-19 13:36:23")
9
10 # 服务application B 的 cluster2集群所依赖的下游application C突然延时增加100ms
11 spec.
12     tareget("application B").
13         cluster_scope_filter("cluster2").
14             dependency("application C").
15                 action("delay, 100ms").
16                     end_at("2020-04-19 13:36:23")
```



字节跳动混沌工程演进之路 – 阶段一

阶段性成果

- 各业务方可自行发起混沌工程实验
- 有能力快速验证故障恢复手段有效性，以及该手段副作用
- 有能力尝试探索故障拐点



产生新需求

- 缺少以业务链路为视角的混沌工程
- 是否可以更贴近生产真实情况的故障模拟
- 是否可以做面向人的混沌工程

字节跳动混沌工程演进之路 - 阶段二

字节跳动混沌工程演进之路 – 阶段二

目标

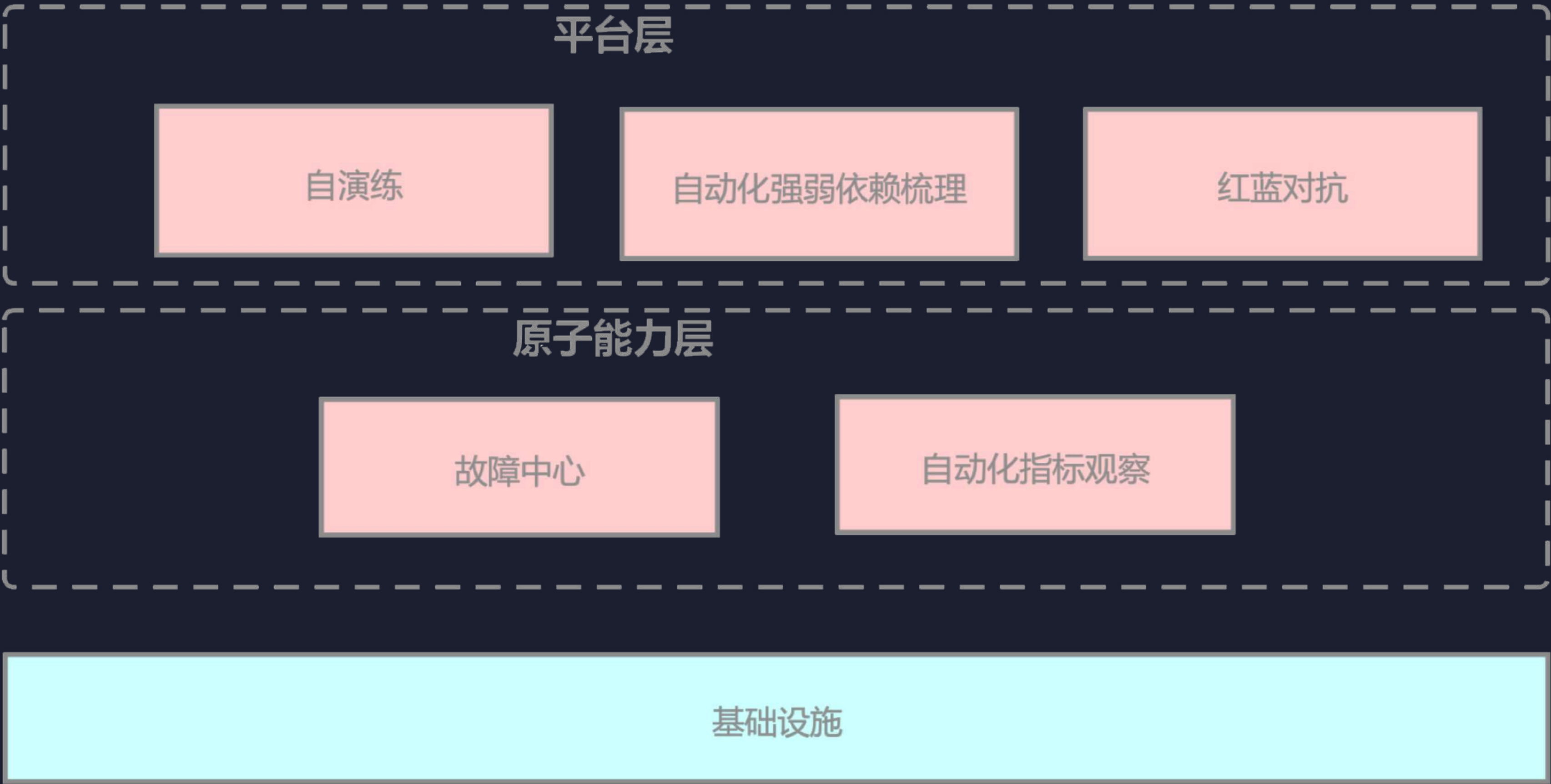
提供强弱依赖自动化分析与红蓝对抗平台。探索更大的收益。

提供自动化指标分析能力，根据稳定性检测自动停止。降低风险成本。

进一步丰富故障的类型。降低时间成本。

字节跳动混沌工程演进之路 – 阶段二

整体设计



字节跳动混沌工程演进之路 – 阶段二

红蓝对抗实践特性

- 角色划分为蓝军（攻击方），红军（防守方）
- 对抗期间蓝军会无通知随机化故障注入，相对更接近平日真实场景。
- 红军需要在规定时间内感知定位故障，并将故障恢复。
- 不只是对系统能力的混沌工程，同时也是对人的混沌工程

字节跳动混沌工程演进之路 – 阶段二

红蓝对抗整体流程



字节跳动混沌工程演进之路 – 阶段二

红蓝对抗实践 单场红蓝对抗数据汇总



字节跳动混沌工程演进之路 – 阶段二

红蓝对抗实践 单场红蓝对抗结果汇总

问题汇总				
问题类型	问题详情			状态
监控问题	告警红灯占仍然未修复，当前仍存在问题			待解决
演练问题	只支持...，需要新增支持			待解决
告警问题	可...的熔断机制而失效，跟进			待解决
预案未生效	只...支持			解决 by ...
兜底问题	...，...			待解决

点击后跳转到故障现场

对抗全记录				
				全展开
m	下游依赖 5%不可用 持续180s 2020-04-15 10:42:40 ~ 2020-04-15 10:45:46	指标表现： 3项	发现问题： 3项问题	防守结果： ● 成功
发现问题		异常指标		
1. 监		1. ...		
2. ...		2. ...		
3. ...		3. ...		
m	下游依赖 ...%不可用 持续180s 2020-04-15 10:42:40 ~ 2020-04-15 10:45:46	指标表现： 0项	发现问题： 3项问题	防守结果： ● 成功
i...	...不可用 持续180s 2020-04-15 10:42:40 ~ 2020-04-15 10:45:46	指标表现： 0项	发现问题： 3项问题	防守结果： ● 成功
n	...%不可用 持续180s 2020-04-15 10:42:40 ~ 2020-04-15 10:45:46	指标表现： 0项	发现问题： 3项问题	防守结果： ● 成功

< 1 2 3 4 5 6 7 8 9 > 10条/页 跳至 5 页

字节跳动混沌工程演进之路 – 阶段二

阶段性成果

- 可以从业务链路视角进行混沌工程
- 探索了系统感知故障的能力(监控、告警)
- 探索了开发者定位故障与解决故障的综合能力

产生新需求

- 探索最佳实践，推广更大范围业务使用
- 如何衡量系统稳定性建设程度

字节跳动混沌工程演进之路 – 阶段三

字节跳动混沌工程演进之路 – 阶段三

目标

固化混沌工程最佳实践，常态化混沌工程实验体系。

进一步丰富故障的类型。降低时间成本。

与预案平台联动，固化预案执行手册。

字节跳动混沌工程演进之路 – 阶段三

固化混沌工程最佳实践



字节跳动混沌工程演进之路 – 阶段三

字节跳动混沌工程成熟度定义

成熟度	演练烈度	故障自身的等级	间隔频率	重叠程度	时间范围	单故障持续时长	故障触发通知	故障恢复能力
初步	温柔的猴子 (自助演练)	不触发高于P2故障	业务owner自行决定	无重叠	无限制	自行决定	知晓触发时间，知晓故障细节	预案执行/恢复手段依赖经验&口口相传
入门	悠闲的猴子 (红蓝对抗)	不触发高于P2故障	不低于15分钟	无重叠	1小时	不大于2分钟	告知触发时间，不告知故障细节	初步整理预案执行手册，能应对普通故障
常规	微怒的猴子 (红蓝对抗)	不触发高于P1故障	不低于5分钟	不超过2	5小时	不大于5分钟	不告知触发时间，不告知故障细节	预案执行手册清晰可理解易执行，外部团队也能应对普通故障
熟练	暴躁的猴子 (红蓝对抗)	不触发高于P0故障	不低于1分钟	不超过3	1天	不大于10分钟	不告知触发时间，不告知故障细节	预案执行手册清晰可理解易执行，外部团队也能应对黑天鹅事件
大师	喜怒无常的猴子 (突袭演练)	任意故障	不低于1分钟	不超过3	任意时间	任意时长	不告知触发时间，不告知故障细节	大部分故障能自动降级/恢复，不惧各类黑天鹅事件

字节跳动混沌工程演进之路 – 总结

总结

低成本的故障注入能力，可帮助业务快速验证故障下的系统表现与恢复手段

混沌工程不止是对系统故障恢复能力的验证，也是面向开发者感知故障、排除故障能力的锻炼

不同系统处在不同的阶段，需要以不同的混沌工程实验方式帮助稳定性建设
(字节跳动混沌工程成熟度定义)

交流&招贤纳士



THANKS

QCon⁺ 案例研习社