

集群性能评估

- 一、压力测试
 - 目标
 - 当前流程
 - 痛点
 - 风险点
 - 报告
 - 采集指标
 - 关注指标
 - 待优化项
 - 参考文档
- 二、集群性能评估
 - 目标
 - 数据对象
 - 数据源
 - 统计指标
 - 流程
- 三、其他
 - 可靠性
 - 验证方式
 - 自动化脚本

一、压力测试

目标

通过自动化压测与模拟压测，确定系统链路流向以及服务单机性能，进而评估促销目标容量与系统稳定性趋势。

当前流程

促销模拟压测

序号	内容	执行人	备注
1	单机QPS预估	性能评估dev	
2	压测容量预估	容量评估dev	
3	通知SRE扩容	压测dev	
4	通知占用LIVETEST	压测dev	在seataalk群 “ livetest对比环境 使用协调 ” 中通知需要占用LIVETEST分支与开关，及其时间段
5	更新LIVETEST分支	业务组PIC	
6	修改LIVETEST开关	业务组PIC	修改appid=sls_api&env=LIVE&cluster={cid}_livetest集群下，ssc.sls.mutable_application namespace中，UseNormalConfigOnDailyVersionCompare 为 true
7	检查开关是否一致	压测dev	检查压测系统中各熔断值是否配置正确+是否生效
8	增量加压	压测dev	加压的过程中，需要各项目组压测PIC同步观察关键指标，如有明显流量/负载上涨，及时沟通上报。 观察项： 1. 关联中间件（MySQL、Codis）和第三方依赖（Spex等）的入口QPS/延迟/错误率/CPU负载指标[Grafana+CAT] 2. 压测服务的错误率/CPU/MEM/Latency等负载指标
9	产出压测报告	压测dev/大促系统	

痛点

场景	现象	原因	改进
日常压测	1. 压测链路放大比与线上偏差大 2. 压测单机QPS预估不准	livetest的流量融合了流量回放引入的流量，导致压测流量接口比例与历史大促有一定差异	放大一定比例的日常压测流量，不再只压单机，减少流量回放的影响
		流量放大的压测方式会导致缓存高命中的现象，压出的性能往往要优于线上真实性能	将压测性能与线上FCST性能进行对比，以线上性能为准
		livetest点线迁移等开关配置与live不一致	短期方案：
		压测代码与线上代码差别较大的代码中合进去很多陈年优化代码	1. 周二上午新增一次自动化压测 [已知每周三是日常版本发版窗口，周二LIVETEST流量回放强占用，配置与代码与上线后保持一致] 2. 每天日常压测之前从apollo和jenkins/dms拉数据，确认开关与代码是否与线上一致。是：标记本次压测为有效压测；反之无效。 长期方案： 隔离流量回放与自动化压测环境
模拟压测	促销规模不一致时，接口比例差异较大，对服务性能影响较大	用户行为不一致	根据不同量级的促销备份多种接口比例 压测平台支持一键更新接口比例

风险点

现象	改进
大规模压测可能会对外部依赖造成较大压力	大规模压测之前预估对第三方依赖（spex、coreserver等）的压力，并邮件/seataalk同步 将依赖方的上限配置为压测任务的熔断值？

报告

采集指标

目标集群	统计指标	统计方式
服务	CPU	MAX, P95, AVG
	MEM	
	LATENCY	
MySQL	Read QPS	MAX
	Connection	
	Thread Running	
	Slow Query	
Codis	QPS	MAX
	CPU	
	MEM	

关注指标

场景	性能指标	计算方式	备注
----	------	------	----

模拟压测	单机QPS	<p>Target = 压测单机QPS * 线上压测差值系数</p> <p>压测单机QPS = QPS by instance / CPU负载 * 目标CPU负载 (62.5% = 五核)</p> <p>差值系数 = (历史大促单机QPS or FCST单机QPS) / 压测单机QPS</p>	<ol style="list-style-type: none"> 1. 模拟压测尽可能的压出各个服务均衡负载时的性能数据 2. 线上单机QPS取线上负载接近5核的真实性能数据，以评定压测性能数据与真实情况的差距 3. 由于FCST/大促时达到正常负载的服务较少，只能抽样计算差值系数，以估算其他无可靠线上性能数据的服务的性能
------	-------	--	---

待优化项

1. 单机QPS计算方式

参考文档

[大促压测流程](#)

二、集群性能评估

目标

结合压测负载数据与线上负载数据，评估各服务、中间件的集群性能与可承载容量。

数据对象

数据源

1. 大促模拟压测性能数据
2. 日常压测性能数据 <历史3天>
3. 历史大促峰值性能数据
4. 历史日常性能数据<FCST>

统计指标

目标集群	统计指标	统计方式	计算方式
Service	单机QPS	MIN	<ol style="list-style-type: none"> 1. 参考数据源2-4，以大促模拟压测数据为准预估单机QPS指标。 2. 标记出数据源中小于平均值30%的最小值，判定该指标是否异常。如果是，去除。 3. 取剩余数据中的最小值作为各服务的单机QPS。
MySQL	读QPS	取最坏情况	无限制，可以通过slave扩展，2w/slave
	写QPS		5000
Codis	QPS		大促前给SRE提单咨询
Dependency	QPS		<ol style="list-style-type: none"> 1. 大促前向下游同步SLS对其的预估压力 2. 咨询可承载最大能力

流程

编号	步骤	说明
1	日常压测	<ul style="list-style-type: none"> • 每周二 10:00-10:30增加一轮发版前得可信自动化压测 • 标记可信/不可信压测数据
2	模拟压测	根据预估流量和预估实例数进行1v1大促模拟压测
3	计算差值系数	<p>取1-3个CPU负载在4核左右的LIVE集群性能数据，与压测数据做对比，计算平均差值系数</p> <p>差值系数 = sum(线上集群QPS/压测集群QPS) / len(集群个数)</p>

4	计算单机数据	预估单机QPS = 压测单机QPS * 差值系数
5	核验单机数据	<ul style="list-style-type: none"> 取最近3次有效日常压测数据和1次一个月内有有效的模拟压测数据 将预估单机QPS数据与上述数据源进行对比： <ul style="list-style-type: none"> 若预估数据 > 1.2 * mim(参考数据), 表示本次压测性能明显提升, 标记该数据为存疑数据, 需人工复核各项数据源是否正正常/近期是否有相关的业务变更/优化 若预估数据 < 0.8 * min(参考数据), 表示本次压测性能明显恶化, 标记该数据为存疑数据, 需人工复核各项数据源是否正正常/近期是否有相关的业务变更 其余数据为可信数据
6	人工校正	人工对不可信数据进行确认&校正
7	促销后验证数据合理性	扫描线上服务的大促历史数据, 拉取负荷在4-6核之间的集群负载数据, 对预估QPS数据进行复盘

三、其他

可靠性

根据线上压测性能与大促模拟压测性能预估服务单机性能, 参考日常压测性能、历史大促性能数据, 甄别是否有性能差距在20%以上的服务。

如有, 需要检查livettest开关、流量配比等压测配置并进行复压。

如果服务性能确实异常, 需与业务组PIC确认性能变化原因。

验证方式

促销之后, 记录线上历史负载数据, 估算服务单机QPS, 与压测结果进行对比, 记录差异在30%以上的服务, 复盘差异原因。

自动化脚本

生成性能结果并同步到容量评估文档中: [api/report/static_evaluation/metric/performance/sync](#)

具体入参含义参考测试用例中的注释: [TestSyncPerfMetrics](#)