目录
数据库
常见框架
分布式
高并发
高可用：如何设计...
高可用：负载均衡...
高性能：池化技术...
高性能：零拷贝为...
高性能：有哪些常...
高可用：降级和熔...
高可用：灰度发布...
服务器
Devops
技术面试题自测篇
面经篇
练级攻略篇

高可用：降级和熔断有什么区别？

大纲
什么是降级？
降级方式有哪些？
服务降级有哪些分类？
大规模分布式系统如何降级？
什么是熔断？
降级和熔断有什么区别？
有哪些现成解决方案？
推荐阅读
参考

什么是降级？

降级是从系统功能优先级的角度考虑如何应对系统故障。

服务降级指的是当服务器压力剧增的情况下，根据当前业务情况及流量对一些服务和页面有策略的降级，以此释放服务器资源以保证核心任务的正常运行。

降级服务的特征如下：

1. 原因：整体负荷超出整体负载承受能力。
2. 目的：保证重要或基本服务正常运行，非重要服务延迟使用或暂停使用
3. 大小：降低服务粒度，要考虑整体模块粒度的大小，将粒度控制在合适的范围内
4. 可控性：在服务粒度大小的基础上增加服务的可控性，后台服务开关的功能是一项必要配置（单机可配置文件，其他可视用数据库和缓存），可分为手动控制和自动控制。
5. 次序：一般从外围延伸服务开始降级，需要有一定的配置项，重要性低的优先降级，比如可以分组设置等级 1-10，当服务需要降级到某一个级别时，进行相关配置

降级方式有哪些？

- 延迟服务：比如发表了评论，重要服务，比如在文章中显示正常，但是延迟给用户增加积分，只是放到一个缓存中，等服务平稳之后再执行。
- 在粒度范围内关闭服务（片段降级或服务功能降级）：比如关闭相关文章的推荐，直接关闭推荐区
- 页面异步请求降级：比如商品详情页上有推荐信息/配送至等异步加载的请求，如果这些信息响应慢或者后端服务有问题，可以进行降级；
- 页面跳转（页面降级）：比如可以有相关文章推荐，但是更多的页面则直接跳转到某一个地址
- 写降级：比如秒杀抢购，我们可以只进行 Cache 的更新，然后异步同步扣减库存到 DB，保证最终一致性即可，此时可以将 DB 降级为 Cache。
- 读降级：比如多级缓存模式，如果后端服务有问题，可以降级为只读缓存，这种方式适用于对读一致性要求不高的场景。

服务降级有哪些分类？

降级按照是否自动化可分为：

- 自动开关降级（超时、失败次数、故障、限流）
- 人工开关降级（秒杀、电商大促等）

自动降级分类又分为：

1. 超时降级：主要配置好超时时间和超时重试次数和机制，并使用异步机制探测回复情况
2. 失败次数降级：主要是一些不稳定的 api，当失败调用次数达到一定阀值自动降级，同样要使用异步机制探测回复情况
3. 故障降级：比如要调用的远程服务挂掉了（网络故障、DNS 故障、http 服务返回错误的状态码、rpc 服务抛出异常），则可以直接降级。降级后的处理方案有：默认值（比如库存服务挂了，返回默认现货）、兜底数据（比如广告挂了，返回提前准备好的一些静态页面）、缓存（之前暂存的一些缓存数据）
4. 限流降级：当我们去秒杀或者抢购一些限购商品时，此时可能会因为访问量太大而导致系统崩溃，此时开发者会使用限流来进行限制访问量，当达到限流阈值，后续请求会被降级；降级后的处理方案可以是：排队页面（将用户导流到排队页面等一会重试）、无货（直接告知用户没货了）、错误页（如活动太火爆了，稍后重试）

大规模分布式系统如何降级？

在大规模分布式系统中，经常会有成百上千的服务。在大促前往往会根据业务的重要程度和业务间的关系批量降级。这就需要技术和产品提前对业务和系统进行梳理，根据梳理结果确定哪些服务可以降级，哪些服务不可以降级，降级策略是什么，降级顺序怎么样。大型互联网公司基本都会有自己的降级平台，大部分降级都在平台上操作，比如手动降级开关，批量降级顺序管理，熔断阈值动态设置，限流阈值动态设置等等。

什么是熔断？

熔断是应对微服务雪崩效应的一种链路保护机制，类似股市、保险丝

微服务之间的数据交互是通过远程调用来完成的。服务 A 调用服务 B，服务 B 调用服务 C，某一时间链路上对服务 C 的调用响应时间过长或者服务 C 不可用，随着时间的增长，对服务 C 的调用也越来越多，然后服务 C 崩溃了，但是链路调用还在，对服务 B 的调用也在持续增多，然后服务 B 崩溃，随之 A 也崩溃，导致雪崩效应

服务熔断是应对雪崩效应的一种微服务链路保护机制。例如在高压电路中，如果某个地方的电压过高，熔断器就会熔断，对电路进行保护。同样，在微服务架构中，熔断机制也是起着类似的作用。当调用链路的某个微服务不可用或者响应时间太长时，会进行服务熔断，不再有该节点微服务的调用，快速返回错误的响应信息。当检测到该节点微服务调用响应正常后，恢复调用链路。

服务熔断的作用类似于我们家用的保险丝，当某服务出现不可用或响应超时的情况时，为了防止整个系统出现雪崩，暂时停止对该服务的调用。

降级和熔断有什么区别？

熔断和降级是两个比较容易混淆的概念，两者的含义并不相同。

降级的目的在于应对系统自身的故障，而熔断的目的在于应对当前系统依赖的外部系统或者第三方系统的故障。

有哪些现成解决方案？

Spring Cloud 官方目前推荐的熔断器组件如下：

- Hystrix
- Resilience4J
- Sentinel
- Spring Retry

我们单独拎出 Sentinel 和 Hystrix 来说一下（*没记错的话，Hystrix 目前已经没有维护了。*）。

Hystrix 是 Netflix 开源的熔断降级组件。Sentinel 是阿里中间件团队开源的一款不光具有熔断降级功能，同时还支持系统负载保护的组件。

简单来说，**两者都是主要做熔断降级的**，那么两者到底有啥异同呢？该如何选择呢？

Sentinel 的 wiki 中已经详细描述了其与 Hystrix 的区别，地址：https://github.com/alibaba/Sentinel/wiki/Sentinel-与-Hystrix-的对比。

下面这个详细的表格就来自 Sentinel 的 wiki。

	Sentinel	Hystrix
隔离策略	信号量隔离	线程池隔
熔断降级策略	基于响应时间或失败比率	基于失
实时指标实现	滑动窗口	滑动窗口
规则配置	支持多种数据源	支持多种
扩展性	多个扩展点	插件的
基于注解的支持	支持	支持
限流	基于 QPS，支持基于调用关系的限流	有限的
流量整形	支持慢启动、匀速器模式	不支持
系统负载保护	支持	不支持
控制台	开箱即用，可配置规则、查看秒级监控、机器发现等	不完善
常见框架的适配	Servlet、Spring Cloud、Dubbo、gRPC 等	Servlet

如果你想了解 Sentinel、Hystrix、resilience4j 三者的对比的话，可以查看 Sentinel 的相关 wiki：https://github.com/alibaba/Sentinel/wiki/Guideline:-从-Hystrix-迁移到-Sentinel#功能对比。

推荐阅读

- 服务降级方案
- 服务熔断处理

参考

- 高并发之服务降级与熔断
- 揭开服务降级的面纱！！！

若有收获，就点个赞吧

SnailClimb 08-03 21:47 4752 0 IP 属地:湖北 举报

上一篇
高性能：有哪些常见的 SQL 优化手段？
下一篇
高可用：灰度发布和回滚有什么用？

加入语雀，参与知识分享与交流
注册 或 登录 语雀进行评论
立即加入
注册 或 登录 语雀进行评论
语雀 | 关于语雀 使用帮助 数据安全 服务协议 | English