# **项目话术**

dubbo分布式架构的新闻项目

项目背景:

nt

dubbo 分布式架构+zk 作为注册中心 版本管理控制工具SVN 文件服务器 FastDFS文件存储系统 Es检索技术 Redis做缓存 Freemake页面静态化技术 RabbitMQ异步解耦 MyCat数据库中间件完成读写分离技术

网络信心中心2名科员 阐述目前需要的项目需求及整理项目文档

设计数据库

建立初始表结构存放初始测试数据在数据库内完成各种业务测试,

新闻主体模块服务(Service-api)

后台管理控制台(页面web端)

日志搜集处理服务(Service-api)

数据检索中心服务(Service-Api-Web端)

缓存中心(Service-api)

新闻评论中心(Service-api)

项目开发分为2个周期:

项目开发环境部署:

1台文件服务器

6台redis(6台服务器) 集群部署

Zk 7台 (集群)[5台正常角色 ----2 台obServer]

ES单台部署(一个分片 300万数据 8g内存左右) 500g存储空间

MySQL3台 主从模式 读写分离技术(MyCat实现)

Maven私服一台

3台MQ服务器 (镜像队列)

Nginx服务器(静态页面)

部署时的nginx 服务器(网关)

1.一期

搭建后台管理平台功能:

## **后台模块**

### **后台模块添加新闻数据**

### **业务介绍**

在添加新闻时每个新闻信息会有所对应类型(国内 国际 财经 周报 娱乐.....),先从缓存中心获取类型缓存数据渲染值前台页面中,每一条都对应有图片 视频上传流程(在页面中==上传==时通过异步方式将文件进行上传至后台服务器,将上传后的文件唯一标识返回给前台页面中,在页面中管理员继续填写其他的新闻内容信息后,表单提交时将回显的文件地址一并发送到后台的添加新闻请求中,新闻的图片和视频都是可选的数据内容) ,新闻数据包含 标题 ,作者,来源 ,编辑时间/修改时间[系统指定],新闻是否置顶,是否开放评论 新闻正文[富文本 编辑器进行添加的]

当新闻添加时向后台发送请求,后台接收请求后将调用新闻服务api将进行进行保存到数据库,并根据新闻的是否置顶信息将新闻是否纳入缓存中.

新闻服务Api中包含新闻的 添加 编辑 删除等接口方法

### **问题**

面试官问题1:==介绍下RabbitMQ==

消息队列(RabbitMQ):

Amqp高级消息协议的一个消息队列(消息中间件)

MQ中常用的分为一下几个交互模型:

点对点 直连交换机 (在业务场景中使用最多)

广播 广播交互机

发布/订阅 topic

一个交换机可以绑定多个队列通过绑定key的值进行绑定,在进行消息投递时发送消息是通过交换机进行发送投递的,那么这个消息投递到哪个队列去通过投递时的路由key.

面试官问题2:

==1.rabbitMQ 如何保证消息不丢失?==

消息丢失的场景:

1.消息还未送达到服务器

2.自动确认了消息但是消费方出现了问题导致 消息在队列中也被消费了,而消费方根本没获取到消息

设置消息需要手动确认

,进行确认已经将消息消费 手动ACk

#### **延伸:**

死信队列

一个消息成为死信的情况;

1.消息被拒(没有重新入队)

2.消息过期

3.队列最大长度

死信队列就是用户处理以上三种情况的之一的消息的,进行普通的消息设置时会将一个队列绑定到一个死信交换机中,当普通的队列出现了那三个情况的其中一个时 消息将直接路由到死信交换机中,

大牛逼们的订单:

用户下单后系统将会将库存扣减,等待用户支付(订单创建后需要在2小时内完成支付)

1.将订单写入数据库,每1分钟查询数据库进行查看

2.死信队列 利用消息的过期时间机制将消息投递到死信队列中,

用户下单后创建一个消息扔进一个没有任何消费方接收的对列中(超时时间是2小时),现在的消息一直在队列中,当消息过了2个小时了,该消息称死信进入死信队列,会有一个接收死信队列的消费方消费消息,获取消息根据当前的订单id进行查看该订单是否支付,如果支付了代表订单正常,如果订单新在数据库中处于未支付状态时证明订单支付超时了.

### **后台模块角色管理**

业务介绍:

为了后台管理系统的安全性考虑在后台系统中加入了安全框架,安全框架采用的是shiro,在数据库设计上分为5个表来维护一个安全管理策略[管理员,角色表,权限表,角色权限维护表,角色管理员维护表]

在权限表中内置了后台所有动作的对应权限,在shiro配置文件中根据不同的请求与动作进行配置,当每个动作需要授权时就在shiro.xml中进行配置,通过shiro的认证授权器会根据配置来决定处理该访问请求是否具备权限

问题:

1.shiro的认证执行流程

我们在使用shiro时会实现一个自定义的authenticationRealm,这里实现类中实现2个自己的业务方法,一个是登录认证,另外一个就是认证后的授权方法,在认证中获取用户的登录名密码和自己添加扩展的验证码进行判断,在判断时是通过调用管理员的Api接口方法获取一个管理员对象,在方法中进行处理,如果认证失败会new shiro中的内置异常类,在页面中通过异常类的信息来进行显示给登录者,同时后台管理员在登录时如果出现了登录失败多次的情况后台会将该管理员账号进行锁定,至于锁定时间和失败次数都是通过我们系统的系统参数管理里面进行设置的.

2.shiro中对于权限管理是如何做的?

在管理员登录后我们会将管理员的身份信息存放一个自己编写的身份牌类中,里面最重要的数据就是该管理员的唯一标识,这个标识我们是使用数据库的主键来维护的,因为在数据库表中设计的是主键是唯一的,同时在数据表设计中管理员与角色表有关联(没有主外键关系,数据联系关系),当管理用户需要进行鉴权时会在authenticationRealm中的授权方法中根据登录者的唯一标识进行查询,查询的是根据id获取该用户下的所有权限,通过shiro的授权器与在shiro.xml中定义配置的关系来决定这个用户的动作是否能被正常执行,如果该用户没有想到权限时在页面是不会显示菜单项的.我们的后台管理系统是使用的jsp来做的,这样做是为了能够与shiro标签进行结合.如果没有相应的权限在shiro的配置文件中有设置未授权的地址

当授权完成后会将该权限列表放进缓存中,这样做的目的就是在每个动作时不再执行查询数据库.这样做一旦更新了角色下的权限时是不会立即刷新的需要在缓存失效后或者登陆这退出后再次登入权限才会刷新

3.shiro的几个组件和shiro的执行流程

组件:

securityManager 安全管理器  
AuthenticationRealm 域  
cacheManager 缓存管理  
登入登出的过滤器

### **后台维护首页轮播图**

业务介绍:

在前台的展示页面中有几幅热点图片轮播区,后台能够对该首页轮播图进行设置,设置项包括:首页轮播图片数,图片点击后的跳转地址

问题:

这个功能是如何实现的?

数据库中有一张首页轮播图片表,在表内设计了如下几个字段:

主键 图片名称 唯一标识名称 图片标题 图片跳转地址 是否轮播 1-轮播 2-撤销

通过后台的图片管理来维护的是轮播表中的是否参与轮播的操作,前台数据获取时查询的是图片轮播表中加入轮播的数据,同时该数据是存放到缓存中的,有效时间通常设置为48小时,如有紧急更新时会触发更新缓存数据

## **缓存中心服务:**

==面试官问题:为何要将缓存抽离成模块,在每个业务中使用不是更方便么,这样拆分不是加大了技术曲线么?==

由于多个服务都需要使用redis进行数据缓存,所以在设计上将缓存单独抽离出成为一个缓存中心,每个服务需要使用缓存时直接调用缓存中心的api 接口即可,当需要进行新闻 及新闻类型的缓存和维护时缓存中心将直接访问MySQL数据库,根据不同的调度来完成缓存添加或修改

面试官问题:

==在新闻添加后如何调用的缓存中心?= u[]\出现问题将直接导致我上游调用者的业务正常进行,所以采取异步方式进行调用无需结果.

2.通过MQ方式来进行异步调用,新闻数据添加到MYSQL后将唯一标识[动作标记]通过MQ进行投递到队列中,缓存中心同样监听绑定缓存的消息队列,从队列中获取唯一表进行缓存从数据库中读取该数据根据不同的情况来进行处理

问题:

==1.redis用在哪个位置了==

新闻类别 热点数据 (新闻置顶数据) 后台管理员的权限信息

==2.redis集群的介绍==

redis用的是三主三从 ,三个主节点三个从节点(主从模式变化来的)主节点发送宕机后对应的从节点会接替该主节点从salve节点变为master节点,当宕机的master节点再次上线时她的角色还是salve节点不能在恢复到原来的master节点了,redis集群中如果出现2个主节点宕机时集群模式将失败,热点集群模式后将不在有redis16个库的概念,取而代之的是哈希操,redis集群会有16384个槽位,分布在不同的集群节点上,每个节点会维护自己的槽位,当有key 进行set时 redis会通过将key进行crc16算法取模进行计算来维护该key落在哪个槽点上.在集群i 时如果设置的key没有在该槽点时redis将重定向到目标维护槽位的节点位置上.

==3.redis持久化的策略==

redis持久化分为2个策略 rdb快照 aop 文件追加方式进行维护 当前目录下 config get dir

1.rdb方式是属于快照方式,在redis中默认有3个或的关系来进行生成快照 ,可以通过手动的方式来进行快照save /bgsave

2.aof文件方式 每个指令将追入到aop文件中,追加策略分为3中模式 ,每秒 实时 根据系统决定 同时为了防止aof文件过大redis还有一个aof重写

面试官问题:

==介绍下FastDFS和如何实现文件上传的==

上传的文件存储系统FastDFS:

FastDFS分为2个服务:1.文件追踪服务,2.文件存储服务,客户连接文件追踪服务,由文件追踪服务联系文件存在,文件上传后文件追踪服务不负责保存文件名称,文件唯一标示名称也不需要在客户段手动维护,FD自动生成,客户端只是将上传后的文件地址进行保存,

### **日志记录**

在对接需求时网信中心主任要求在后台进行操作管理时希望能够对每个行为作出记录,所以在后台关系的系统中我们又加入了日志记录的功能,记录的日志包括 登录时间 登录ip 失败次数登录的信息,同时当登录后的管理员在进行后台的信息维护时每个请求我们都进行了记录操作,将记录的日志存放MySQL和Es中索引.

问题:

==你的这个日志记录怎么写的?==

我们实现方式是利用SpringAop的方式来进行切入记录的,在用户发请求的时候我们的切入点为每个控制层,同时也排除了一些不需要记录的方法其中包括 跳转那个页面这种请求我们没有记录,注意记录的是 在业务中的数据操作为主,在切面类中我们记录的该管理员的 执行时间 执行参数 登录的内部ip地址等信息进行保存设置,同时为了更好能够溯源还添加来日志查询,在日志查询时我们是通过ES方式来进行数据检索的,MySQL只是进行物理数据保存,没过一段时间这些日志将清理一次,但是近2周内的日志都会保留.

==谈谈你对Aop的理解==

几个通知类型

什么是aop

配置Aop

## **数据检索模块**

为了能够更好对新闻数据信息进行查询我们加入了ES,最开始设计的是不使用任何的搜索技术,就是单纯的MySQL数据库进行查询,但是这样查询会导致查询数据不准确,所以在和甲方沟通后加入了Es检索技术. 我们检索实现是在后台对新闻数据的添加后会将新闻的id扔进MQ队列中在检索服务中监听该消息,当获取到消息后检索服务之间查询数据库获取该新闻数据进行索引,在索引中我们只索引 标题 作者来源 新闻类型信息 同时为了能够在检索数据后能够在页面展示出新闻的封面图片同时又另外存储了图片的地址,但是这个图片地址不是能够检索的属于业务上的冗余索引列,在搜索服务中可以根据标题进行全文检索,同时类型 作者来源没有进行全文检索在设置上我们是使用的keyword.用户在前台页面总根据要检索的新闻标题关键字进行新闻搜索,查询到的结果设置了高亮显示功能.

问题:

1.你们的ES是如何分页的?

首先es中提供了from 和size 俩个参数设置,可以根据这个来完成分页功能,在进行业务分页显示检索时获取到的数据结果要分析处理.先根据条件获取source中的total取出共检索到多少条数据,再根据这个数据来进行分页

延伸问题(分页):你们的es中有多少热点检索数据,包含经常被检索的数据和没有定期清理的数据大概在1w-3w之间,

那么数据量超过 1w了 有该如何分页?

因为我们是单机部署,不存在集群分片等查询分页问题所以在使用上依旧是正常分页处理

2.你们的ES是集群还是单机部署?

3.如果单机Es发生了宕机怎么办?(单机没有办法来处理宕机问题,因为单机一个分片没有副本分片宕机后只能通过MySQL进行查询处理,等待ES恢复)

4.MySQL中的新闻数据一旦更新后Es中是否会同步该数据?

我们在后台更新新闻时有2个更新操作,一个是更新新闻内容,一个是变更维护新闻标题等内容.如果新闻内容中出现了错别字获取其他误会 歧义性的内容时通过修改正文才操作,但是在Es索引中并不会索引新闻的内容正文,更新内容后不会触发更新索引,只有在编辑新闻标题 类型这种导致检索条件发生变化的数据时才会更新维护Es中的索引数据.

5.在使用Es时是用的什么方式技术进行使用的?

Es官方提供的api,并没有使用Spring中的Data Es来检索,因为Es中只要更新一个Es版本就会同时推出一个相应的java Api,我们用的是Es的高阶Api来完成处理的.

6.在进行数据检索时你都结果高亮是如何实现的?

在es中通api进行设置高亮时并不能直接从检索的结果中获取高亮显示信息,要想设置高亮显示结果需要在构建查询时加入真毒哪个field进行设置,当执行查询返回结果后需要进行手动的判断处理结果中是否包含高亮字段,如果包含高亮字段设置就需要手动去获取解析然后在返回处理实体对象中set变换才能拿到带有自己设置的高亮显示内容.

7.分词是如何实现的?

我们在Es的插件中安装中文分词器,同时还加入了扩展词库,扩展词库是同nginx方式来配置完成的一个远端扩展词,每当有新的词汇时我们会在扩展词汇的文件将该词汇进行加入.分词采用的 最细分词算法

## **前端页面渲染**

前台的新闻数据展示我们用的是页面静态化技术,在进行新闻数据检索后用户根据检索到的相关性新闻标题可以点击进新闻详情页面,这个页面是html展示的,在添加新闻数据后会直接生成一个新闻内容的详细信息页面,这个页面部署在nginx中,用户只是点击标题即可直接跳转到对应的页面中显示,在页面中下方会使用异步方式加载该新闻的评论内容信息,根据甲方沟通一些新闻类型不能够添加评论[国内 政治 类的这些新闻是不能够添加评论内容]

freemaker页面布局我们是由前端人员来编写的 样式和js脚本 后台只是根据不同的模板进行设置生成html静态页面,

当新闻内容信息和标题新闻等修改时需要重新生成一个新的新闻详情页面

问题:

1.你们和前端人员是如何交互的?

前端开发和我们都是一起在甲方进行协同开发的所有在交互沟通上都在直接协调,但是项目后期进行验收时后补齐的相关文档.

人员配置:

5个人 :

1个前端开发工程师

2后台开发人员(小组)

1产品需求对接人员

项目经理整体的项目把控

你的角色:

1.如果你对这项目业务上没有任何太大问题那就就说你是小组领导,职责就是搭建项目架构选型 进行负载小组的员分配等工作

1.1 考虑到主观因素在甲方要求构建一套分布式体系架构的项目(政绩)

2.2 如果是采取SpringCloud整理架构的话组件比Dubbo更加完善,但是小组中人员水平有限有限开发人员对Cloud体系没过整体开发流程,为了不耽误项目进度所以在设计架构时采取的Dubbo ,第二 项目内部使用在并发上 安全性并没有那么高所以就选用了dubbo技术架构

3.在这个项目中遇到的技术难题有哪些? 你是如何解决的?

项目初期设计上缓存中心的数据是通过RPC直接调用业务接口来获取的,这样一来导致调用出现一定的混乱,业务链路稍长,所以在后面将缓存中心的消费者角色取消掉每次缓存或更新数据是都在直接通过自己查询数据库,哪方在使用缓存时直接调用服务api接口 由api接口调用缓存层

[管理端]--->[缓存]---->[业务]

2.你对这个项目的模块不是很清晰只写过其中的几个简单的功能那就写该功能

业务数据的清理:

在设计上每天凌晨3点对业务数据进行清理,清理的条件是根据当前时间以天为计算查询大于现在时间的2周的数据进行清除,清除后包括 索引数据 和缓存数据