哈尔滨工业大学

**计算机科学与技术学院/国家示范性软件学院**

**2021年春季学期**

**《软件架构与中间件》课程**

**实验报告**

**Lab 2：计算层软件架构实验**

|  |  |  |
| --- | --- | --- |
| **姓名** | **学号** | **联系方式** |
| 梅智敏 | 1183710118 | 1044388658@qq.com |
| XXX | 16NNNNNNN | Email/手机号码 |

目 录

[1 实验概述 1](#_Toc69932094)

[1.1 实验目的 1](#_Toc69932095)

[1.2 实验要求 1](#_Toc69932096)

[2 实验内容与过程 1](#_Toc69932097)

[2.1 nginx集群负载均衡实验 1](#_Toc69932098)

[2.2 hadoop分布式计算实验 2](#_Toc69932099)

[2.3 activeMQ异步消息推送实验 2](#_Toc69932100)

[3 结对开发过程记录 2](#_Toc69932101)

[4 实验总结 3](#_Toc69932102)

[5 教师评语 4](#_Toc69932103)

[文档全部完成之后，请在上述区域点击右键，选择“更新域”，在打开的对话框中选择“更新整个目录”]

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **学号：** | **1183710118** | **姓名：** | **梅智敏** | |
| **学号：** |  | **姓名** |  | |
| 实验概述 | | | |  |
| 实验目的 1）学习构建集群的方法  2）学会使用典型的负载均衡框架  3）学会使用Hadoop 进行map-reduce编程  4）能够灵活应用计算层中间件到实际系统 实验要求 1）2人结对成组  2）实验2.1和2.2必做，实验2.3选做  3）结合《软件过程与工具》课程中进销存系统(或其他实际软件系统)进行计算层架构重构，支持海量用户的在线高并发请求场景  4）应给出关键过程的细节 | | | | |
| 实验内容与过程 | | | |  |
| Nginx集群负载均衡实验  1. 请给出搭建单虚拟机Nginx的关键过程，并修改Nginx主机的首页内容从英文变为中文。   我首先要安装nginx    安装成功之后再启动nginx    接下来安装net-tools包    这样就可以使用ifconfig命令来查看本机的IP地址    我们使用浏览器访问该ip地址    接下来找到nginx默认配置文件的位置    以管理员身份修改它（不加sudo的话无权限修改此文件，所以需要加sudo）    在配置文件的Server内容中添加一行charset utf-8    然后前往配置文件中所指的root目录下(/var/www/html)，将index.nginx-debian.html复制一份并重命名为index.html    将index.html进行编辑，将其中的英文内容翻译成中文    更新nginx的配置    重新访问前面的IP地址    显示的内容确实是我们自己创建的index.html，说明更新配置成功！   1. 请详析配置文件中涉及的upstream、location和server块。   upstream：用于进行负载均衡的配置    Location模块：用于进行访问路由的配置    Server模块：用于进行服务器访问信息的配置  可以简单地理解为用来定义一个URL，需要在其中定义server\_name和监听端口号port，以及监听到该URL后执行什么动作。     1. 请描述所搭建的虚拟集群环境，验证虚拟集群可以互相访问。   所搭建的虚拟集群有一台负载均衡服务器，两台Web服务器，负载均衡服务器可以接受多个客户端的请求  我所搭建的虚拟集群如下：一台负载均衡服务器，2太Web服务器    首先我将3个ubuntu虚拟机的连接方式改为桥连接，并且将使用的网卡设置为主机正在使用的网卡：    设置为桥接之后，各个虚拟机将和主机处于同一网段下，且拥有自己独立的ip地址，这保证了它们可以相互ping通  负载均衡服务器ip地址：192.168.43.176    Web\_01服务器IP地址：192.168.43.4    Web\_02服务器IP地址192.168.43.96    负载均衡服务器可以PING通其他2台Web服务器，如下图：    Web01服务器也可以PING通另外2台服务器，如下图：  Web02服务器也可以PING通另外2台服务器，如下图：    综上，可见我们的三台服务器都是可以相互访问的   1. 配置集群的负载均衡系统，给出关键过程，验证负载均衡系统是否工作。  * 负载均衡服务器上的操作   首先在负载均衡服务器上新建配置文件      新建/etc/nginx/conf.d/proxy\_params配置文件    修改nginx的主配置文件/etc/nginx/nginx.conf，以把负载均衡服务器的80端口空闲出来     * 对2个Web服务器的操作（下面展示Web01服务器上的操作，Web02与之类似）   在Web01服务器上新建配置文件/etc/nginx/conf.d/server.conf    首先创建node文件夹    再在此文件夹下新建html文件    Web02服务器中的内容为Web02，如下     * 将三个服务器的nginx服务重启     现在在win10主机上访问负载均衡服务器的IP地址    可知负载均衡服务器将我们客户端的请求转发给了Web01服务器来处理    刷新界面之后发现负载均衡服务器将我们的请求转发给了Web02服务器进行处理   1. 配置并验证使用不同集群负载均衡算法（至少三种），给出过程、观测的情况，给出自己对负载均衡算法的理解和认识。      * 默认的轮询方式：在前面的测试中就是轮询方式，每刷新一次页面，负载均衡服务器会切换一个呗访问的服务器，也就是上面所呈现的Web01页面与Web02页面交替出现。 * Weight：权重方式，在轮询策略的基础上指定轮询的几率。   我们修改负载均衡服务器的配置文件，将Web02服务器设置权重为2    然后重启负载均衡服务器的nginx服务，再次从主机访问负载均衡服务器的IP并不断刷新，可以发现显示的内容Web02：Web01正好就是2:1，满足我们所设置的权重比。     * Ip\_hash分配：指定负载均衡器按照基于客户端IP的分配方式，这个方法确保了相同的客户端的请求一直发送到相同的服务器，以保证session会话。这样每个访客都固定访问一个后端服务器，可以解决session不能跨服务器的问题。   首先在配置文件中添加一行”ip\_hash”,指明使用IP\_hash负载均衡策略；同时注意此策略和权重策略不可以同时使用，故需要删除之前添加的权重值。    重启nginx服务器：    再次从主机上访问，可以发现不论怎么刷新，返回的页面都是不变的，说明负载均衡服务器确实按照请求方的IP地址来分配一个确定的Web服务器    总结：   1. 轮询是最基本的配置方法，它是upstream模块默认的负载均衡默认策略。每个请求会按时间顺序逐一分配到不同的后端服务器。 2. 权重策略是在轮询策略的基础上指定轮询的几率。 3. IP\_hash指定负载均衡器按照基于客户端IP的分配方式，这个方法确保了相同的客户端的请求一直发送到相同的服务器，以保证session会话。 4. 应用负载均衡技术改造遗留的“进销存”系统，赋予支持海量用户的在线高并发请求的能力，请给出设计细节并分析负载均衡前后的区别。     设计细节如上图：采用一个负载均衡服务器来帮助缓解单个Web服务器的访问压力，使用IP\_HASH负载均衡策略，来自某个终端的请求会被发送到特定的Web服务器（依据终端的IP地址），服务器处理完之后再将结果返回给终端用户。  下面给出负载均衡之前的整体框架图：    不使用负载均衡服务器的时候，所有用户的请求都将被发送到同一个Web服务器上，这样就会使得这个服务器的访问压力非常大。尤其在访问高发时间段，可能会导致服务器繁忙从而产生高延迟、无响应等现象。  通过对前面2种情况的对比可发现，使用负载均衡服务器可以将来自客户端的大量请求分摊到多个Web服务器上，从而缓解单个Web服务器的访问压力。  7）迁移到真实集群环境重做负载均衡实验(选做)。请给出实验过程中需要注意的问题，比较并分析虚拟集群和真实集群下的异同。 Hadoop分布式计算实验 1）配置集群环境，设置主节点主机名为master，从节点主机名为workerX(X为数字编号)，给出主要过程；同时给出配置SSH免密登录的关键步骤。  首先下载jdk-11.0.11\_linux-x64\_bin.tar.gz，解压到/home/meizhimin/jdk-11.0.11。    修改/etc/profile文件，配置JDK信息    然后执行source /etc/profile使得配置文件生效    然后为meizhimin用户添加权限：    修改 /etc/hostname 文件： master节点的主机设置为 master ，其他两个虚拟机分别设  置为slave1、 slave2  并且修改/etc/hosts文件如下：    安装ssh    创间.ssh目录并赋予当前用户权限    生成RSA密钥对以及集群内共享密钥：      配置集群内的3个节点可以通过ssh无密码访问：      现在测试无密码ssh访问      成功！  2）请配置Hadoop集群计算环境，要求数据块副本数为3。给出主要过程，并验证正确性。  首先下载hadoop3.2.2并解压到用户目录下    然后给hadoop配置环境变量：    使用source命令使得配置文件生效：    修改Hadoop安装目录/etc/hadoop/目录下的Hadoop-env.sh文件，添加：    修改core-site.xml中主要内容：    hdfs-site.xml中主要内容：value字段设为3，表示数据块副本数为3    mapred-site.xml中主要内容(MapReduce的配置文件)    配置三台主机的Hadoop文件(就是创建masters和workers文件并编辑)      向 slave1 和 slave2 节点复制 hadoop-3.2.2 整个目录至相同的位置，使用如下命令（将sam换成自己的用户名，hadoop版本也换成自己的即可）    在master节点上执行hdfs namenode -format    再执行start-all.sh验证正确性    使用jps查看    故成功开启！  3）请使用Hadoop集群环境计算英文版[Stray Birds](https://en.wikisource.org/wiki/Stray_Birds)正文中每个单词出现的次数。  进入hadoop的安装目录    下载Stray Birds文件，并上传到hdfs文件系统    调用wordcount函数来对输入的文件进行处理，结果存储在output中    结果如下：    4）请使用Hadoop集群环境计算遗留的“进销存”系统中海量用户的log日志中的访问统计。  在我们的分布式系统中会有一个专门的logInfo文件用以保存用户的登录信息，其内容如下所示：      将其上传到HDFS服务器并且调用wordcount函数来对输入的文件进行处理，统计出各个用户的访问次数，结果如下：    在我们采用Hadoop的map-reduce编程之后，系统将会充分利用我们设置的三个数据块的计算资源，从而更为快速地完成分布式统计计算任务，大大提高我们系统的计算高性能。 ActiveMQ异步消息推送实验 1）安装并配置ActiveMQ 5.15.9，验证正确性。  2）请先创建名为asm和se的主题，然后面向这些主题，给出发送端和接收端的交互流程，最后编写流程对的代码。  3）请完成如下要求：如果发送的消息是object，该如何处理；如何保证消息处理成功之后，会发送消息确认。  4）请比较实验1自主开发的消息中间件与ActiveMQ的异同，重点给出如何改进自己程序的不足。 | | | | |
| 结对开发过程记录 | | | |  |
| **（1）角色切换与任务分工**  表1-1结对开发角色与任务分工   | 日期 | 时间(HH:MM - HH:MM) | 驾驶员角色 | 领航员角色 | 本段时间的任务 | | --- | --- | --- | --- | --- | |  |  |  |  |  | |  |  |  |  |  | |  |  |  |  |  |   【注意】该表格可自行增加更多的行。  **（2）工作日志**  由领航员负责记录，记录结对开发期间的遇到的问题、两人如何通过交流合作解决每个问题的。  表1-2 结对开发工作日志   | 日期/时间 | 问题描述 | 最终解决方法 | 交流过程 | | --- | --- | --- | --- | |  |  |  |  | |  |  |  |  | |  |  |  |  |   【注意】该表格可自行增加更多的行  **（3）结对开发工作现场照片、或视频及文件沟通截图**  请其他同学帮助拍摄结对开发现场照片至少2张。  13071045_wLhN 1  结对开发现场照片1 结对开发现场照片2 | | | | |
| 实验总结 | | | |  |
| 【结合前期课程项目系统或其他实际软件系统，解决其中的消息订阅/分发问题等】 | | | | |
| 教师评语 | | | |  |
|  | | | | |