

《农产品批发市场大数据分析系统》

用 户 需 求 说 明 书

v1.0

|  |
| --- |
|  |

**版 本 历 史**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本/状态 | 作者 | 参与者 | 日期 | 备注 |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |
|  |  |  |  |  |

**目 录**

第一部分 引言1

一、说明1

二、定义1

第二部分 综述3

一、项目背景3

二、建设目标3

三、建设原则4

四、用户业务需求说明6

（一）、整体业务需求示意图6

（二）、需求详细说明6

2.1.用户查询业务流程描述6

（1）、价格查询6

（2）、价格分析预测7

2.2.后台数据处理业务流程描述7

（1）、数据采集7

（2）、数据存储7

（3）、数据处理8

2.3.用户查询业务流程描述8

（1）、权限管理8

（2）、数据管理8

第三部分 需求分析9

一、用例分析9

（一）、 用户用例9

1.1、用户注册用例描述10

1.2、用户登录用例描述10

1.3、请求监控数据用例描述10

1.4、抓取量查询用例描述11

1.5、实时数据监控用例描述11

1.6、数据查询用例描述12

1.7、价格查询用例描述12

1.8、获取价格走势用例描述13

1.9、价格对比用例描述13

1.10、获取预测价格用例描述14

（二）、 后台管理员用例15

2.1、管理员登陆用例描述15

2.2、数据维护用例描述16

2.3、数据检索用例描述16

二、界面风格17

第四部分 验收标准18

一、功能范围定义18

二、性能指标定义18

第五部分 环境和部署要求19

一、网络部署图19

二、应用部署图20

三、运行环境说明20

（一）、服务器20

（二）、客户机器环境21

**第一部分 引言**

一、说明

编写本说明书的目的是为了准确阐述项目具体业务需求和需求边界，本说明书的作者是【农产品批发市场大数据分析系统】项目组，本说明书的确认者是【项目经理】负责人，本说明书的读者是项目所有直接干系人。

本说明书是指导项目实施的重要指导性文件，也是用户最后进行验收（终验）的依据，说明书中内容一旦确认双方将以此为基础开展工作。如果需要变更说明书内容，必须走变更流程，变更必须得到甲乙双方书面确认，最后变更内容将作为本文的一部分，在项目实施过程中得以体现。

二、定义

（一）、农产品批发市场

是以粮油、畜禽肉、禽蛋、水产、蔬菜、水果、茶叶、香辛料、花卉、棉花、天然橡胶等农产品及其加工品为交易对象，为买卖双方提供长期、固定、公开的批发交易设施设备，并具备商品集散、信息公示、结算、价格形成等服务功能的交易场所，按交易商品的种类范围，农产品批发市场分为综合型批发市场（交易农产品品类三大类以上）和专业型批发市场（交易农产品两种以下）两种。

（二）、再销售者

再销售者是指通过转卖商品而获得利润的营利性组织。包括二次及以下的批发商以及零售商。再销售者是批发商业的基本用户，其主要目的是为了卖而买，通过赚取买卖价差即毛利而获得利润。对再销售者来说，其最关心的问题就是所采购的商品能否有理想的毛利率与好的销路。这就要求批发商业尽量控制成本和销售价格，以便让用户有利可图；同时还要认真研究消费者的需求动向，以便组织消费者所需要的商品。

（三）、农产品交易大数据

农产品交易大数据是指大数据大数据理念、技术和方法在农产品商贸领域的实践。涉及农产品的交易、分析、决策等方面，是跨行业、跨专业的数据分析与挖掘。

（四）、差价

差价是指不同等级，不同交割月份，不同商品，不同交割地点的期货价格差异。也指同一商品因各种条件不同而产生的价格差别，如批发和零售的差价、地区差价、季节差价。 进货与卖出货的差额是一些商人盈利的主要部分。

**第二部分 综述**

一、项目背景

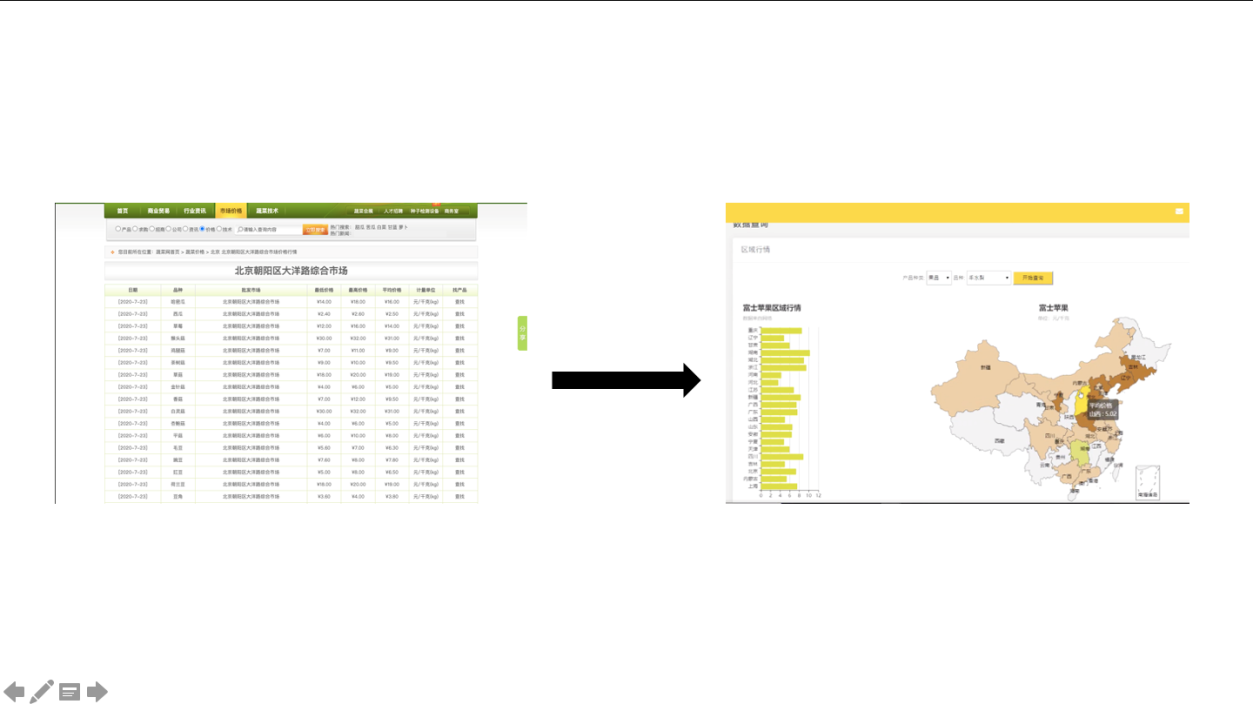
近几年随着电子商务的兴起，农产品的网上销售也模式也逐渐发展起来，各种涉农信息的发布也改变了传统模式，直接利用网络资源进行宣传与销售。而原有的大型果蔬农产品批发市场，也逐步改变原有的管理销售模式，采用新的电子商务销售系统，对农产品交易和农资信息发布进行有效的整合，为人们在线进行农产品交易提供虚拟网络环境和保障农产品电子商务顺利运营的管理环境。

然而我国土地面积大，各地气候环境不一，这就造成了人们所需求的农产品不可能完全在某一个地区实现供给，需要各地区农产品批发市场的负责人在全国各地进行农产品的采购，在这种情况下，一个实时整合分析全国各地农产品价格的数据分析系统的需求也就越来越高。

在此，为实现一个能为各地果蔬交易商提供及时有效的农产品信息，为了能够以更快更便捷的方式向大家展示农产品的各类信息，在保证信息真实可靠的前提下，决定利用各地各大果蔬农产品批发市场的交易系统所提供的交易信息，开发一款农产品交易市场的大数据分析系统。

二、建设目标

1、利用数据可视化的技术，将刻板枯燥的各个省份各个市场各个品类不同日期的价格以更加形象直观的图表以及表格的形式展现出来，方便用户以及种植商通过页面观看最近5天的农产品的价格走势，方便用户以及种植商做出适当的决定，使其利益最大化。



2、采用大数据技术，利用数据分析的结果，对部分地区的部分农产品进行价格预测，从大数据中获取农产品价格的变化趋势，发现市场规律，掌握市场行情。

3、 从宏观方面说，本产品的建设目标是为了不断推进农业经济的优化，实现可持续的产业发展和区域产业结构优化，进一步推动智慧农业的建设进程以及全面及时掌握农业的发展动态。

三、建设原则

（一）实用有用

系统将以权威的信息网站作为来源基础，及时的从各大农产品的信息网站获取各类信息，并不时地进行数据更新，以保证提供的农产品数据信息实用、可用。

（二）灵活先进

系统将采用更加优化的算法进行数据的分析，以保证系统能够更快地对用户的需求做出响应。系统应充分考虑到信息的不断变化的需求，对信息的处理以及人员权限等参数进行调整设置，在系统的设计方法、运行模式上充分利用现有的先进技术，最大限度地发挥现有网络与设备的优势，并为软件未来因新技术的出现的升级留下空间。

（三）界面友好

本项目在遵循以用户为中心原则、一致性原则、简洁与明确原则、体现特色原则、兼顾不同的浏览器原则以及明确的导航设计原则的基础之上，还将遵循界面友好的设计原则。本项目将采用清新的UI设计风格，主要界面用于展示农产品的价格以及走势，使用户集中精力获取想要的信息，加之适当的动画，提升用户的界面体验。同时，设计的界面也将尽可能减少用户的操作，让用户在尽可能少的操作中就获取到自己需要的信息。

（四）兼容扩展

具有兼容性，项目需要能够在Chrome、Bing、Firefox、百度等不同浏览器中都能运行。

方便扩展，在未来项目能够方便集成更多新的特性以及新的功能，具有良好的可扩展能力。

（五）安全可靠

系统运行具有较高的稳定性，确保数据的收集的安全、可靠与保密。

（1）系统保障在本系统内注册的用户的信息安全，谢绝外单位的非管理员访问，具体的个人信息谢绝各类人员访问；

（2）数据库内存放的各类数据信息不允许从外部网络进行访问，需要管理人员在内部网络进行登录才可进入管理界面；

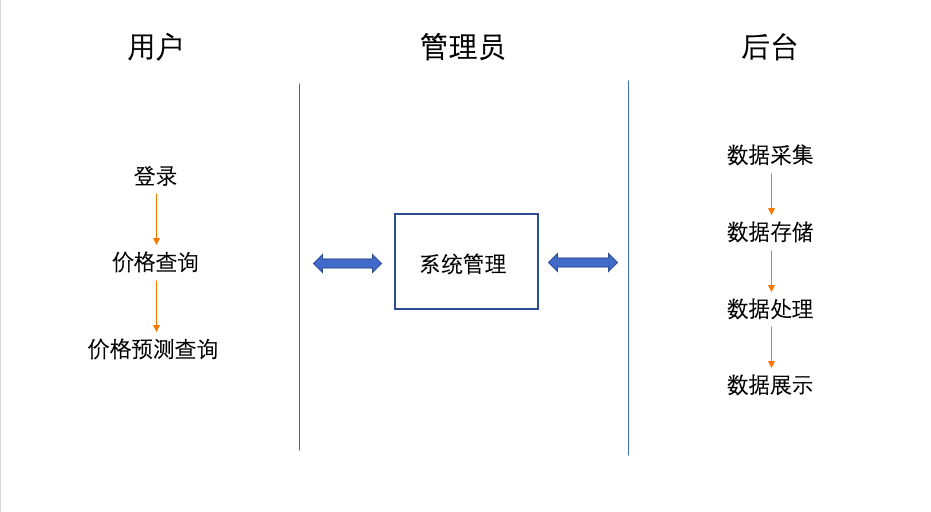
（3）系统的核心程序文件不允许被访问，需要专业人员以特殊的方式登录才可访问；

（4）数据备份，保证数据的安全；

（5）多维度保证系统运转的可靠性。

四、用户业务需求说明

（一）、整体业务需求示意图



（二）、需求详细说明

2.1、用户查询业务流程描述

（1）、价格查询

用户可以通过网站的用户登录进入网站，不登陆以游客的方式进入网站也可以，通过点击网站上相应的按钮进入到农产品的价格查询界面，用户可以查看到全国的各个农产品的全国平均 价格，用户可以点击某一产品，可以进入以省份分类的各省各大农产品批发市场关于该产品的价格。用户同样可以在网页的上方输入特定省份特定农产品的名称进行查询。

在查询界面用户可以自行选择数据的显示方式，系统默认以列表的方式呈现，，用户还可以通过点击上方各类显示方式以柱状图等方式呈现。

（2）、价格分析预测

用户使用本功能需要进行注册登录，在登陆成功后即可通过点击界面的预测按钮查看到该农产品的历史价格，以及经过本系统数据分析后得出的未来的数据变化趋势。同样在本界面用户可以自行选择显示的方式，系统默认以折线图的方式呈现。

2.2、后台数据处理业务流程描述

（1）、数据采集

本项目的暂定数据源为<http://www.vegnet.com.cn>，通过Java爬虫项目从中国蔬菜网的网页中抓取到各个省份各个市场各个品种的蔬菜历史价格与实时价格，历史数据只抓取一次，实时数据通过定时器每天定时抓取。

（2）、数据存储

采集到的数据需要存储起来，数据存储主要有两个部分，一部分是抓取下来的数据存储到分布式系统HDFS中上面，另一部分是存储在Hive里面，通过Hive查询将结果导入到MySQL中。

（3）、数据处理

数据处理包括两个部分。第一部分是用MapReduce对分布式系统HDFS上的数据进行清洗，在清洗阶段，对不同的数据定义不同的清洗规则来清洗数据。第二部分是利用大数据技术，对某些地区某些品种蔬菜的数据进行特定的处理，从而实现改种蔬菜的价格预测功能。

2.3、系统管理业务流程描述

（1）、权限管理

用户在网站注册后，只有登录的权限，查看网站内的各种农产品信息，但无法获得对系统进行管理的权限，如需添加管理员，则需要给该用户进行权限分配。

权限分配须在管理员的界面进行设定，也就是说，需要其他的管理员为其进行用户类别的更改，并且得到其他管理员组中其他管理员的确定后方可生效 。

（2）、数据管理

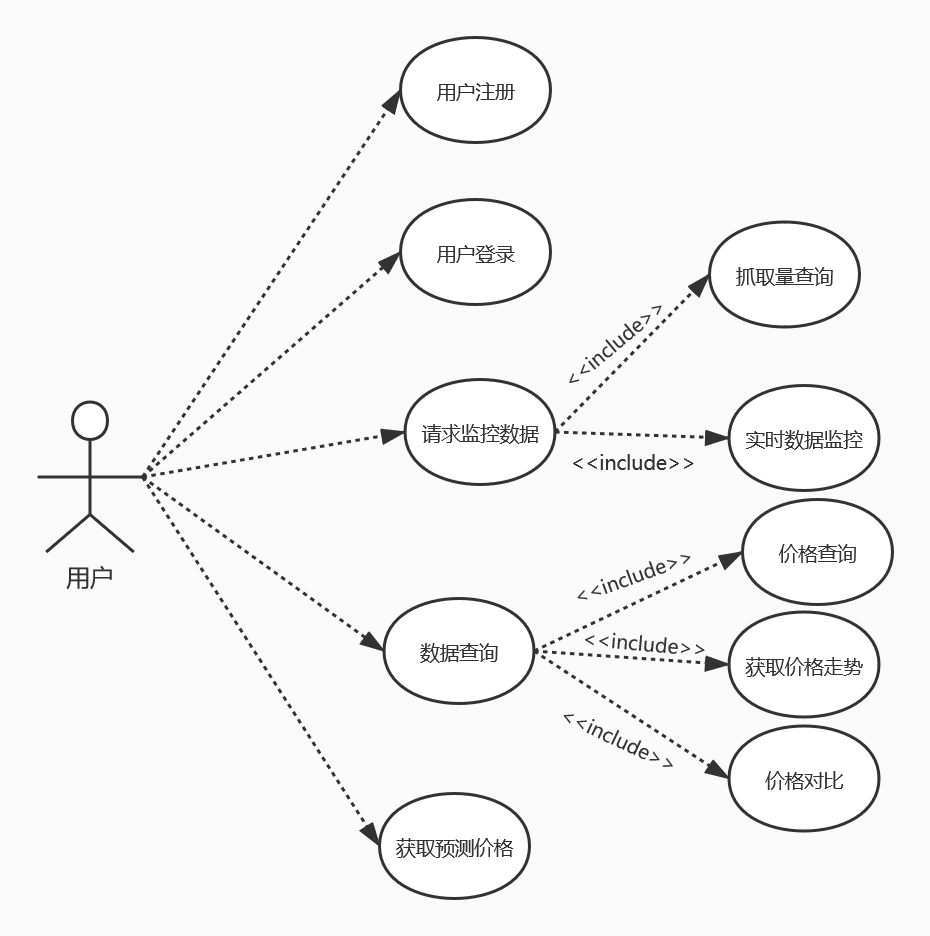
系统管理员在爬取数据后，要将数据保存至分布式系统中，并利用Hive查询将结果保存至MySQL数据库中，保存时涉及到增加数据、删除超过时间的数据等操作。

同时，在系统运行过程中，系统管理员需要根据系统的运行情况对数据进行管理。当系统中出现了错误数据，系统管理员要从后台手动删除错误数据；当获得了其他数据（不是从网页中爬取的，而是以其他方式获得，如csv文件等），系统管理员需要手动将数据导入到后台数据库中。同时，系统管理员还需要对服务器进行维护。

**第三部分 需求分析**

一、用例分析

（一）、 用户用例



1.1、用户注册用例描述

|  |  |
| --- | --- |
| ID | YH0000 |
| 用例名称 | 用户注册 |
| 父用例ID |  |
| 主要执行者 | 用户 |
| 前置条件 | 无 |
| 事件流 | a.在登陆界面选择注册跳转至注册界面。  b.账号输入框内输入用户账号，在密码输入框内输入用户密码。  c.按照网页给定验证码图输入验证码。  d.点击注册按钮，注册成功。  e.存储该用户信息。 |
| 可选事件流 | 无 |
| 异常事件流 | d1.若用户已经存在账号，则显示注册失败，跳转至登陆界面。 |
| 后置条件 | 无 |

1.2、用户登录用例描述

|  |  |
| --- | --- |
| ID | YH0001 |
| 用例名称 | 用户登录 |
| 父用例ID |  |
| 主要执行者 | 用户 |
| 前置条件 | 用户已注册 |
| 事件流 | a.打开登录页面，选择用户登录.  b.在用户名输入框内输入用户名。  c.在密码输入框内输入密码。  d.点击登录按钮登录。  e.登陆成功跳转至用户主页。 |
| 可选事件流 | 无 |
| 异常事件流 | d1.若用户名密码不匹配，显示登陆失败。 |
| 后置条件 | 无 |

1.3、请求监控数据用例描述

|  |  |
| --- | --- |
| ID | YH0002 |
| 用例名称 | 请求监控数据 |
| 父用例ID |  |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录成功 |
| 事件流：  包含事件 | 1.抓取量查询  2.实时数据监控 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

1.4、抓取量查询用例描述

|  |  |
| --- | --- |
| ID | YH0003 |
| 用例名称 | 抓取量查询 |
| 父用例ID | WMS0002 |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流 | a.在数据监控模块选择抓取量查询。  b.用户选择某农产品批发市场。  c.显示对于该批发市场抓取到的数据总量信息，如品种总数、蔬菜或水果的占比等。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

1.5、实时数据监控用例描述

|  |  |
| --- | --- |
| ID | YH0004 |
| 用例名称 | 实时数据监控 |
| 父用例ID | WMS0002 |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流 | a.在数据监控模块选择实时数据监控。  b.显示当前批发市场的实时监测结果。  c.显示当前农产品的实时监测结果。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

1.6、数据查询用例描述

|  |  |
| --- | --- |
| ID | YH0005 |
| 用例名称 | 数据查询 |
| 父用例ID |  |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流:  包括事件 | 1.价格查询  2.获取价格走势 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

1.7、价格查询用例描述

|  |  |
| --- | --- |
| ID | YH0006 |
| 用例名称 | 价格查询 |
| 父用例ID | YH0005 |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流 | a.在数据查询模块选择价格查询。  b.选择批发市场后，种类中会加载爬取到的农产品品种。  c.选择一个品种后，点击查询按钮。  d.用图表形式显示一段时间内改农产品的价格。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

1.8、获取价格走势用例描述

|  |  |
| --- | --- |
| ID | YH0007 |
| 用例名称 | 获取价格走势 |
| 父用例ID | YH0005 |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流: | a.在数据查询模块选择价格走势。  b.选择一个农产品品种后，点击查询按钮。  c.用图表形式显示该农产品的一段时间内的价格走势。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

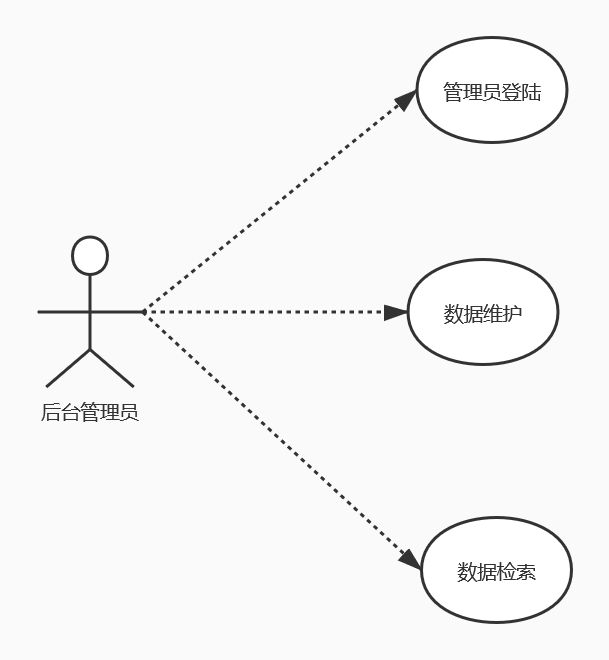
1.9、价格对比用例描述

|  |  |
| --- | --- |
| ID | YH0008 |
| 用例名称 | 价格对比 |
| 父用例ID | YH0005 |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流: | a.在数据查询模块选择价格对比。  b.选择一个农产品品种后，点击查询按钮。  c.用图表形式显示各批发市场该农产品的价格对比图。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

1.10、获取预测价格用例描述

|  |  |
| --- | --- |
| ID | YH0009 |
| 用例名称 | 获取预测价格 |
| 父用例ID |  |
| 主要执行者 | 用户 |
| 前置条件 | 用户登录 |
| 事件流: | a.进入价格预测模块。  b.用户选择农产品品种。  c.图表中显示一段时间中该农产品的平均价格变化趋势与基于计算预测的价格。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

2、 后台管理员用例



2.1、管理员登陆用例描述

|  |  |
| --- | --- |
| ID | GLY0000 |
| 用例名称 | 管理员登陆 |
| 父用例ID |  |
| 主要执行者 | 后台管理员 |
| 前置条件 |  |
| 事件流 | a.登录页面，选择管理员登陆  b.在用户名输入框内输入管理员用户名。  c.在密码输入框内输入管理员密码。  d.点击登录按钮登录。  e.登陆成功跳转至管理员主页。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

2.2、数据维护用例描述

|  |  |
| --- | --- |
| ID | GLY0001 |
| 用例名称 | 数据存储 |
| 父用例ID |  |
| 主要执行者 | 后台管理员 |
| 前置条件 |  |
| 事件流 | a.在爬取数据后，将数据保存至分布式系统中。  b.利用hive查询将结果保存至MySQL数据库中。  c.保存时涉及到增加数据、删除超过时间的数据等操作。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

2.3、数据检索用例描述

|  |  |
| --- | --- |
| ID | GLY0002 |
| 用例名称 | 数据检索 |
| 父用例ID |  |
| 主要执行者 | 后台管理员 |
| 前置条件 |  |
| 事件流 | a.获取查询数据目标，生成查询指令。  b.在HDFS或MySQL数据库中进行相应数据的检索。  c.返回检索到的数据。 |
| 可选事件流 | 无 |
| 异常事件流 |  |
| 后置条件 | 无 |

二、界面风格

**WEB登录界面：**



**主页面：**



**第四部分 验收标准**

**一、功能范围定义**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **#** | **产品** | **模块** | **组件** | **规格/型号** | **角色** | **接入** |
| 1 | 农产品批发市场大数据分析系统 | 系统管理单元 | 系统管理模块 | 用户管理、资源库管理、内容管理 | 管理员 | web |
| 3 | 数据管理单元 | 数据管理模块 | 数据挖掘、数据清洗、数据分析、数据维护 | 管理员 | web |
| 11 | 查询功能管理单元 | 查询功能模块 | 抓取量查询、数据监控、价格查询、获取价格走势、价格对比查询 | 普通用户 | web |

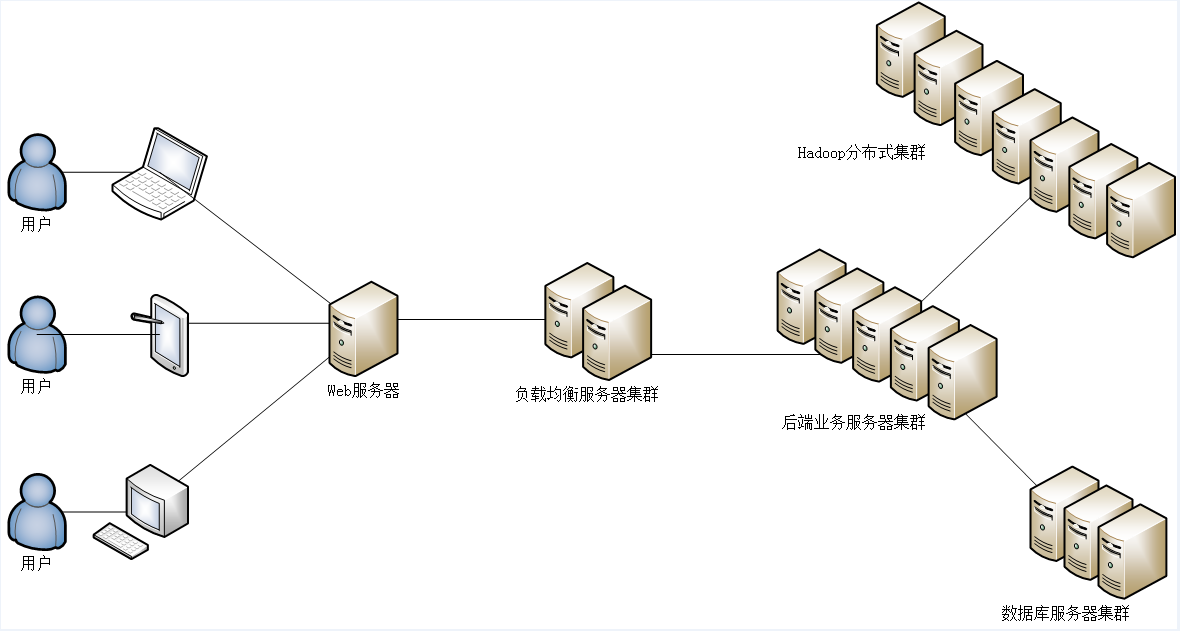
**二、性能指标定义**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **#** | **产品** | **模块** | **组件** | **规格/型号** | **性能级别** |
| 1 | 农产品批发市场大数据分析系统 | 系统管理单元 | 系统管理模块 | 用户管理、资源库管理、内容管理 | A |
| 6 | 数据管理单元 | 数据管理模块 | 数据挖掘、数据清洗、数据分析、数据维护 | A |
| 10 | 查询功能管理单元 | 查询功能模块 | 抓取量查询、数据监控、价格查询、获取价格走势、价格对比查询 | A |

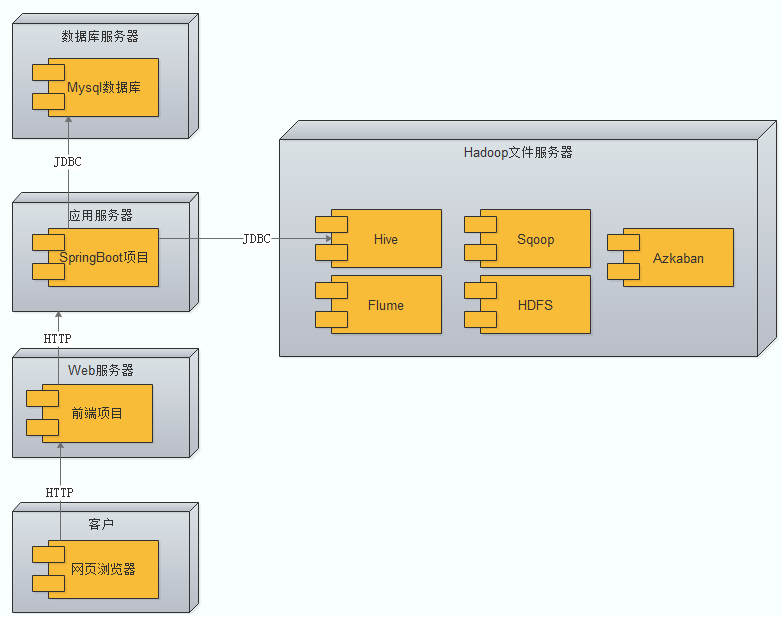
**说明：**级别（A:表示非常重要必须达到的技术性能要求,B:表示重要推荐达到的技术性能要求,C：表示非重要可以弱化的技术性能要求.）

**第五部分 环境和部署要求**

**一、网络部署图**



**二、应用部署图**



**三、运行环境说明**

**（一）、服务器**

Web服务器配置：

* CPU 2核
* 内存 4G
* OS CentOS8
* 磁盘空间 40G

负载均衡服务器配置：

* CPU 2核
* 内存 4G
* OS CentOS8
* 磁盘空间 40G

后端服务器配置：

* CPU 4核
* 内存 16G
* OS CentOS8
* 磁盘空间 200G

数据库服务器配置：

* CPU 4核
* 内存 16G
* OS CentOS8
* 磁盘空间 500T

Hadoop分布式服务器配置：

* CPU 4核
* 内存 16G
* OS CentOS8
* 磁盘空间 500T

服务器上软件版本如下：

* JDK 1.8
* MySQL 8.0.19
* Hadoop 3.2.1
* Hive  3.1.2
* Sqoop 2
* Flume 1.9.0
* Nginx 1.16.1
* Zeppelin 0.8.1
* Azkaban

**（二）、客户机器环境**

Windows/Linux/MacOS/Android/IOS

由于是B/S架构，所以支持浏览器的机器即可