系统分析与设计

**科技专家资源共享平台**

**数据库设计说明书**

|  |  |  |
| --- | --- | --- |
| **项目组成员信息** | | |
| **小组名称** | **洪洋大小姐不想爆肝** | |
| **小组联系人** | **杨洪洋、林未** | |
| **学号** | **姓名** | **本文档中主要承担的工作内容** |
| 16211028 | 杨洪洋 | 确定分工，审核整个文档 |
| 16211085 | 索宇澄 | 完成数据库管理系统结构和非结构化数据技术方案部分 |
| 16211093 | 林未 | 设计数据库，完成并修改数据库详细信息 |
| 16211079 | 赵双叶 | 设计数据库，完成数据库部署方案部分，完成数据库建表，完善数据库模型图 |
| 16211088 | 凌国辉 | 完成数据库安全措施部分 |
| 16211117 | 王嘉星 | 完成数据存储规模估计部分 |
| 16211077 | 王启鹏 | 设计数据库，完成并修改数据库详细信息 |
| 16211125 | 赵昀睿 | 完善文档内容，整合全部文档 |

2019-04

版本变更历史

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 版本 | 提交日期 | 主要编制人 | 审核人 | 版本说明 |
| 01 | 2019.4.18 | 赵昀睿 | 杨洪洋 | 整合文档 |
| 02 | 2019.4.18 | 杨洪洋 | 杨洪洋 | 修改文档其他细节 |

**目 录**

[1. 概述 1](#_Toc6509272)

[1.1 系统背景 1](#_Toc6509273)

[1.2 文档概述 1](#_Toc6509274)

[2. 数据库物理模型 1](#_Toc6509275)

[2.1 数据库物理模型图 1](#_Toc6509276)

[2.2 数据库表汇总说明 2](#_Toc6509280)

[2.3 数据库表详细信息 3](#_Toc6509281)

[2.3.1 用户表 3](#_Toc6509282)

[2.3.2 认证申请记录表 3](#_Toc6509283)

[2.3.3 科技成果表 4](#_Toc6509284)

[2.3.4 专利表 4](#_Toc6509285)

[2.3.5 论文表 5](#_Toc6509286)

[2.3.6 项目表 5](#_Toc6509287)

[2.3.7 充值记录表 6](#_Toc6509288)

[2.3.8 购买记录表 6](#_Toc6509289)

[2.3.9 转让记录表 6](#_Toc6509290)

[2.3.10 评论表 7](#_Toc6509291)

[2.3.11 浏览记录表 7](#_Toc6509292)

[2.3.12 下载记录表 7](#_Toc6509293)

[2.3.13 搜索记录表 7](#_Toc6509294)

[3. 数据库管理系统结构和非结构化数据技术方案 8](#_Toc6509295)

[3.1 数据库管理系统方案 8](#_Toc6509296)

[3.2 数据库管理工具——Navicat 10](#_Toc6509297)

[3.3 非结构化数据存储方案 10](#_Toc6509298)

[4. 数据存储规模估计 10](#_Toc6509299)

[4.1 结构化数据存储规模估计 11](#_Toc6509300)

[4.2 非结构化数据存储规模估计 11](#_Toc6509301)

[5. 数据部署方案 11](#_Toc6509302)

[5.1 数据库部署方案 11](#_Toc6509303)

[5.1.1 Master HA 11](#_Toc6509304)

[5.1.2 mysql-prox 12](#_Toc6509305)

[5.1.3 LVS 12](#_Toc6509306)

[5.1.4 具体部署方案 12](#_Toc6509307)

[5.2 非结构化数据部署方案 14](#_Toc6509308)

[5.2.1 OpenStack Swift技术 14](#_Toc6509309)

[5.2.2 具体部署方案 14](#_Toc6509310)

[6. 数据库安全措施 15](#_Toc6509311)

[6.1 数据库安全问题 15](#_Toc6509312)

[6.1.1 非授权用户对数据库的恶意存取和破坏 15](#_Toc6509313)

[6.1.2 安全环境的脆弱性 15](#_Toc6509314)

[6.1.3 数据过载 15](#_Toc6509315)

[6.1.4 物理损坏 16](#_Toc6509316)

[6.2 数据库安全措施 16](#_Toc6509317)

[6.2.1 用户身份鉴别 16](#_Toc6509318)

[6.2.2 存取控制 16](#_Toc6509319)

[6.2.3 数据加密 16](#_Toc6509320)

[6.2.4 数据备份 16](#_Toc6509321)

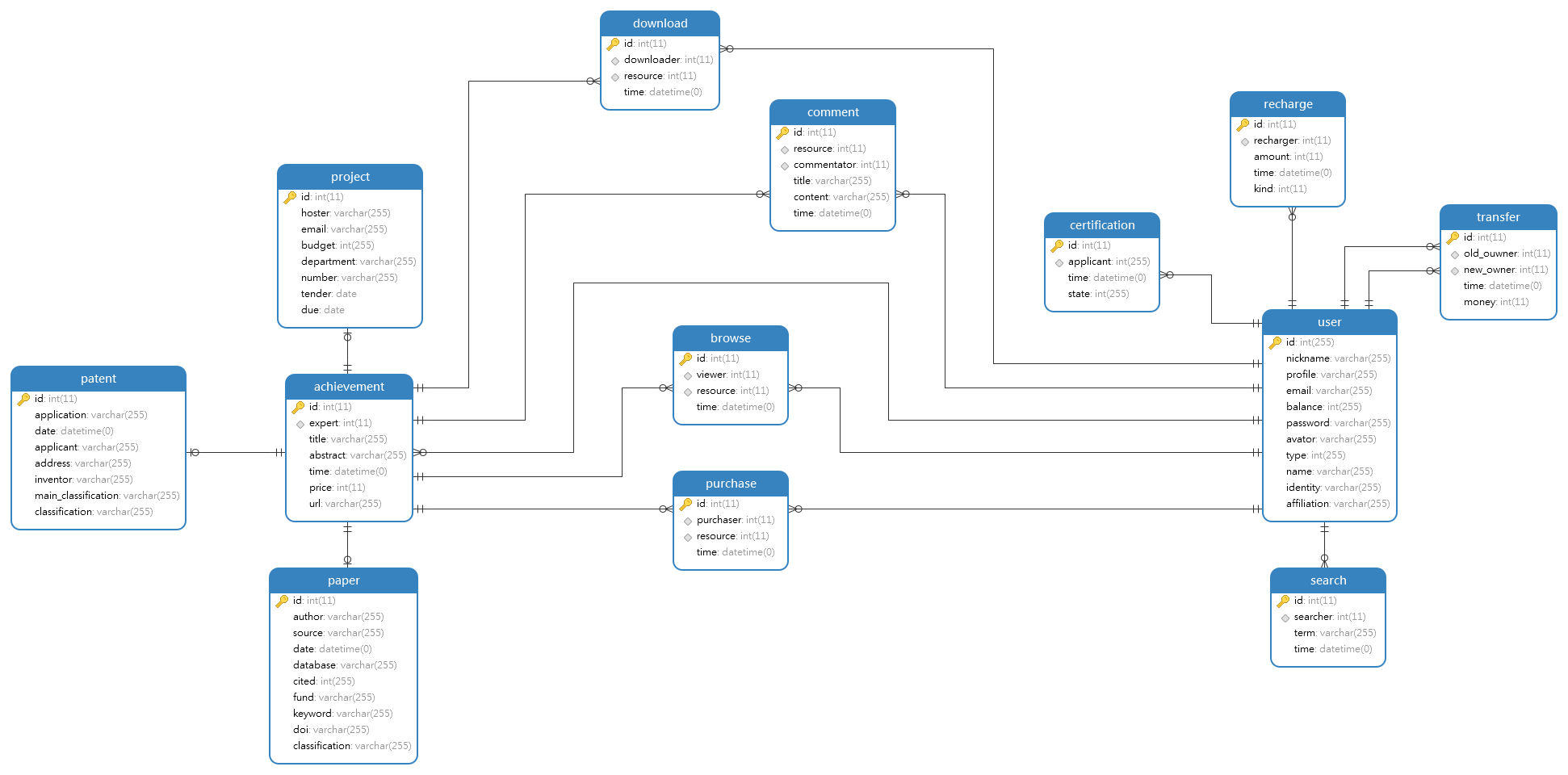
1. 概述
   1. 系统背景

科技专家资源共享平台主要用于有效地推动科研成果合作和科技资源共享，通过在统一管理的基础上建立科技专家门户系统以实现管理和维护科技专家资源。平台基本的运营和管理包括：获得专家科技资源成果、科技专家申请管理和维护自己的门户系统、普通用户检索和查看科技专家各类资源、对各类专家资源进行有效的统计和分析。通过进行需求调研、建立需求模型和分析模型，可以建立软件的架构设计。在完成需求分析和软件架构后，进行数据库设计，以有效地存储数据，满足用户使用需求，为后来的平台开发奠定基础。

* 1. 文档概述

本文档首先介绍了数据库的物理模型，明确了数据库的各个表名、字段名等数据信息；然后分析并提出了数据库管理系统管理方案与非结构化数据存储方案；接下来对数据库的规模进行了估计，提出并详细阐述了数据库的部署方案；最后分析了数据库潜在的安全问题与对应的防护措施。

1. 数据库物理模型
   1. 数据库物理模型图



**图2.1 数据库物理模型图**

图2.1中论文表（papers）、专利表（patents）和项目表（projects）继承自成果表（achievement）。（注：此处图片经压缩后无法辨识，可查看源文件“数据库物理模型图.png”）



## 数据库表汇总说明

表2.2 数据库表汇总说明表

|  |  |  |
| --- | --- | --- |
| 英文表名 | 中文表名 | 功能说明 |
| user | 用户 | 保存普通用户、科技专家和平台管理员的详细信息，包括用户名、密码、邮箱、积分等信息，是用户登录、认证申请、管理科技成果等重要动作相关的信息表 |
| certification | 认证申请记录 | 保存申请专家认证的信息记录，普通用户申请科技专家认证，需要经过平台管理员审核认证 |
| achievement | 科技成果 | 保存科技成果的详细信息，包括标题、定价、摘要等信息，是系统的核心实体，对科技成果进行交易、转让、评价等重要动作相关的信息表 |
| patent | 专利 | 科技成果的一种，保存专利的额外信息，包括申请日、主分类号、专利号等重要信息 |
| paper | 论文 | 科技成果的一种，保存论文的额外信息，包括发表日期、被引量等重要信息 |
| project | 项目 | 科技成果的一种，保存项目的额外信息，包括申请经费、项目编号等重要信息 |
| recharge | 充值记录 | 保存用户充值积分的详细信息，包括充值时间、充值积分数、充值用户等信息，注册用户进行积分充值的重要记录表 |
| purchase | 购买记录 | 保存用户购买科技成果的详细信息，包括付费时间、付费积分数、付费对象等信息，注册用户进行成果购买的重要记录表 |
| transfer | 转让记录 | 保存科技专家转让科技成果的详细信息，包括转让时间、转让专家与被转让专家等信息，科技专家进行成果转让的重要记录表 |
| comment | 评论 | 保存用户对科技成果的评价信息，包括标题、内容等信息，用户对科技成果的反馈 |
| browse | 浏览记录 | 保存用户对科技成果的浏览信息，包括浏览时间、浏览成果等信息 |
| download | 下载记录 | 保存用户对科技成果的下载信息，包括下载者、下载时间等信息 |
| search | 搜索记录 | 保存用户对科技成果的搜索信息，包括检索词、搜索时间等信息 |

## 数据库表详细信息

### 用户表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 用户ID | int |  | 否 | 主键 |
| nickname | 用户名 | varchar(255) |  | 否 |  |
| profile | 个人简介 | varchar(255) |  | 否 |  |
| email | 邮箱 | varchar(255) |  | 否 |  |
| balance | 积分 | int | 0 | 否 |  |
| password | 密码 | varchar(255) |  | 否 |  |
| avator | 头像URL | varchar(255) | rsp/avator.jpg | 否 |  |
| type | 身份 | int |  | 否 |  |
| name | 姓名 | varchar(255) |  | 是 |  |
| identity | 身份证号 | varchar(255) |  | 是 |  |
| affiliation | 所属机构 | varchar(255) |  | 是 |  |

### 认证申请记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 记录号 | int |  | 否 | 主键 |
| applicant | 申请人ID | int |  | 否 | 外键 |
| time | 申请时间 | datetime |  | 否 |  |
| state | 状态 | int |  | 否 |  |

### 科技成果表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 科技成果ID | int |  | 否 | 主键 |
| expert | 所属专家 | int |  | 否 | 外键 |
| title | 标题 | varchar(255) |  | 否 |  |
| abstract | 摘要 | varchar(255) |  | 否 |  |
| time | 发布时间 | datetime |  | 否 |  |
| price | 定价 | int | 0 | 否 |  |
| url | 成果URL | varchar(255) |  | 是 |  |

### 专利表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 专利ID | int |  | 否 | 主键/外键 |
| application | 申请号 | varchar(255) |  | 否 |  |
| date | 申请日 | date |  | 否 |  |
| applicant | 申请人 | varchar(255) |  | 否 |  |
| address | 地址 | varchar(255) |  | 否 |  |
| inventor | 发明（设计）人 | varchar(255) |  | 否 |  |
| main\_classification | 主分类号 | varchar(255) |  | 否 |  |
| classification | 分类号 | varchar(255) |  | 否 |  |

### 论文表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 论文ID | int |  | 否 | 主键/外键 |
| author | 作者 | varchar(255) |  | 否 |  |
| source | 来源 | varchar(255) |  | 否 |  |
| date | 发表日期 | date |  | 否 |  |
| database | 数据库 | varchar(255) |  | 否 |  |
| cited | 被引量 | int |  | 否 |  |
| fund | 基金 | varchar(255) |  | 是 |  |
| keyword | 关键词 | varchar(255) |  | 是 |  |
| doi | DOI | varchar(255) |  | 是 |  |
| classification | 分类号 | varchar(255) |  | 是 |  |

### 项目表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 项目ID | int |  | 否 | 主键/外键 |
| hoster | 项目主持人 | varchar(255) |  | 否 |  |
| email | 邮箱 | varchar(255) |  | 否 |  |
| budget | 申请经费 | int |  | 否 |  |
| department | 所在单位 | varchar(255) |  | 否 |  |
| number | 项目编号 | varchar(255) |  | 否 |  |
| tender | 招标时间 | date |  | 是 |  |
| due | 交付时间 | date |  | 是 |  |

### 充值记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 充值记录ID | int |  | 否 | 主键 |
| recharger | 充值者 | int |  | 否 | 外键 |
| amount | 充值积分数 | int |  | 否 |  |
| time | 充值时间 | datetime |  | 否 |  |
| kind | 第三方支付系统类别 | int |  | 否 |  |

### 购买记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 购买记录ID | int |  | 否 | 主键 |
| purchaser | 购买者 | int |  | 否 | 外键 |
| resource | 购买成果 | int |  | 否 | 外键 |
| time | 购买时间 | datetime |  | 否 |  |

### 转让记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 转让记录ID | int |  | 否 | 主键 |
| old\_owner | 转让专家 | int |  | 否 | 外键 |
| new\_owner | 被转让专家 | int |  | 否 | 外键 |
| time | 转让时间 | datetime |  | 否 |  |
| money | 转让积分 | int |  | 否 |  |

### 评论表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 评论ID | int |  | 否 | 主键 |
| resource | 评论对象 | int |  | 否 | 外键 |
| commentator | 评论者 | int |  | 否 | 外键 |
| title | 标题 | varchar(255) |  | 否 |  |
| content | 内容 | varchar(255) |  | 否 |  |
| time | 评论时间 | datetime |  | 否 |  |

### 浏览记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 浏览记录ID | int |  | 否 | 主键 |
| viewer | 浏览者 | int |  | 否 | 外键 |
| resource | 浏览成果 | int |  | 否 | 外键 |
| time | 浏览时间 | datetime |  | 否 |  |

### 下载记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 下载记录ID | int |  | 否 | 主键 |
| downloader | 下载者 | int |  | 否 | 外键 |
| resource | 下载成果 | int |  | 否 | 外键 |
| time | 下载时间 | datetime |  | 否 |  |

### 搜索记录表

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 英文字段名 | 中文字段名 | 数据类型 | 默认值 | 允许为空 | 主键/外键 |
| id | 搜索记录ID | int |  | 否 | 主键 |
| searcher | 搜索者 | int |  | 否 | 外键 |
| term | 检索词 | varchar(255) |  | 否 |  |
| time | 搜索时间 | datetime |  | 否 |  |

关于表索引的说明： 所有数据库表的主键外键都有自动索引，候选键（可唯一标识记录）也有索引。

# 数据库管理系统结构和非结构化数据技术方案

## 数据库管理系统方案

数据库管理系统使用MySQL。MySQL 是最流行的关系型数据库管理系统之一，在 WEB 应用方面，MySQL是最好的 RDBMS (Relational Database Management System，关系数据库管理系统) 应用软件。MySQL是一种关系数据库管理系统，关系数据库将数据保存在不同的表中，而不是将所有数据放在一个大仓库内，这样就增加了速度并提高了灵活性。我们之前已经在专业课上学习过SQL语法，因此学习成本最低。

在将 MySQL 与其他数据库系统进行比较时，所要考虑的最重要的因素是性能、支持、特性（与 SQL 的一致性、扩展等等）、认证条件和约束条件、价格等。相比之下，MySQL 具有许多吸引人之处：

1. 简单易用

MySQL 是一个高性能且相对简单的数据库系统，与一些更大系统的设置和管理相比，其复杂程度较低。

2. 价格

MySQL 对多数个人用户来说是免费的。

3. 小巧

MySQL 4.1.1发行版仅仅只有21M，安装完成也仅仅51M。

4. 支持查询语言

MySQL 可以利用 SQL（结构化查询语言），SQL 是一种所有现代数据库系统都选用的语言。也可以利用支持 ODBC（开放式数据库连接）的应用程序，ODBC 是 Microsoft 开发的一种数据库通信协议。

5. 性能

许多客户机可同时连接到服务器。MySQL数据库没有用户数的限制，多个客户机可同时使用同一个数据库。可利用几个输入查询并查看结果的界面来交互 式地访问 MySQL。这些界面为：命令行客户机程序、Web 浏览器或 X Window System 客户机程序。此外，还有由各种语言（如C, C++, Eiffel, Java, Perl, PHP, Python, Ruby, and Tcl）编写的界面。因此，可以选择使用已编好的客户机程序或编写自己的客户机应用程序。

6. 连接性和安全性

MySQL 是完全网络化的，其数据库可在因特网上的任何地方访问，因此，可以和任何地方的任何人共享数据库。而且 MySQL 还能进行访问控制，可以控制哪些人不能看到您的数据。

7. 可移植性

MySQL 可运行在各种版本的 UNIX 以及其他非 UNIX 的系统（如 Windows 和 OS/2）上。MySQL 可运行在从家用 PC 到高级的服务器上。

8. 开放式的分发

MySQL 容易获得；只要使用 Web 浏览器即可。如果不能理解某样东西是如何起作用的，或者对某个算法感到好奇，可以将其源代码取来，对源代码进行分析。如果不喜欢某些东西，则可以更改它。

9. 速度

MySQL 运行速度很快。开发者声称 MySQL 可能是目前能得到的最快的数据库。

下面是MySQL数据库管理系统的架构：

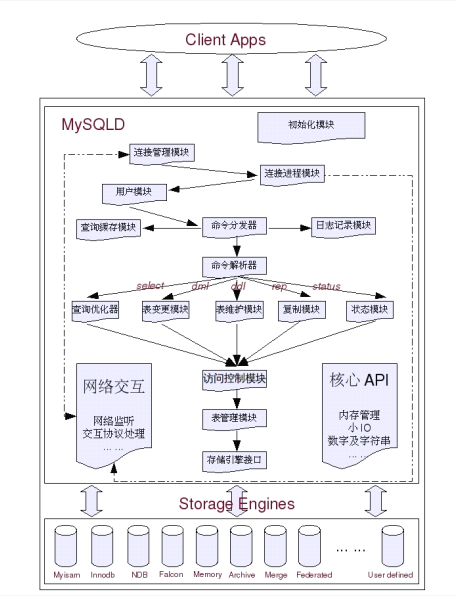


图3.1 MySQL管理系统结构图

## 数据库管理工具——Navicat

Navicat是一套快速、可靠并价格相当便宜的数据库管理工具，专为简化数据库的管理及降低系统管理成本而设。它的设计符合数据库管理员、开发人员及中小企业的需要。Navicat 是以直觉化的图形用户界面而建的，让你可以以安全并且简单的方式创建、组织、访问并共用信息。

Navicat具有很多优点，例如：

1. 用户界面简单易用

Navicat 设计简单，易于使用，不管是数据库开发人员或新手用户，都会认为Navicat 非常容易使用。

1. 丰富的功能

Navicat 提供了丰富的功能，使数据库管理更轻松更快捷。除了常规的数据库管理功能，如编辑或设计表、数据写入、SQL 转储，和创建或编辑用户，其他有用的功能包括：导入/导出向导、报表创建工具、数据模型、查询创建工具、自动完成代码、数据/结构同步，批处理作业计划，HTTP/SSH 通道连接、备份以及更多。

1. 支持SSH通道/HTTP连接

Navicat 支持通过 SSH 通道和 HTTP 通道连接到数据库。SSH 能防止安全漏洞，可以访问远端服务器而安全性不受影响。

1. 支持最全面的文件类型的导入/导出

Navicat 支持最全面的数据格式迁移。Navicat 的导入向导允许数据从多达 11 种数据格式导入到数据库，同时 Navicat 的导出向导支持数据导出到 11 种数据格式，包括 MS Access、MS Excel、XML、HTML、TXT、CSV 及更多。

1. 价格便宜

Navicat 提供一个公平的定价模式。为不同的预算和项目的要求，提供了批量许可证和区域许可证的折扣优惠。对于非商业组织则会有一个特别价格。

## 非结构化数据存储方案

在我们系统之中，非结构化数据主要为资源数据，包括论文、专利以及项目文件等，还有一部分是图片数据。我们所使用的存储方案是：将资源文件存储在服务器的文件系统之中，在数据库的表中存储资源的绝对路径。这样，当我们需要取用某个资源时，根据它在磁盘上的路径便可很快找到它。

# 数据存储规模估计

数据存储为本项目重要部分，无论是用户相关信息还是科技成果相关信息的存储都依赖于数据库。下面将根据需求调研结果以及服务器限制结合本项目组期望的效果对数据库存储规模进行估计。预估本项目的普通用户数量为50万人，专家用户为2万人，平均每人在平台上有10项科技成果。

将本平台的数据划归为结构化数据和非结构化数据两部分。结构化数据即储存在数据库中的数据，包括用户信息、评论等各类事件的记录等。非结构化数据包含科技成果的源文件、来源网站地址等。以下将对这两方面进行分析。

## 结构化数据存储规模估计

结构化数据即储存在数据库中的数据，数据库中的数据内容主要为文本类信息（用户头像除外），所以在结构化数据存储上花费的空间较非结构化数据小。

普通用户信息包括用户个人信息、浏览记录、搜索记录、充值记录、认证申请、下载记录等，每位普通用户的信息总量约为600KB，其中用户头像大小约为100KB。科技专家继承于普通用户，包含以上普通用户所包含的全部信息，除此之外还包含转让信息、项目信息、专利信息等，估算每位科技专家信息总量为5MB。

因此专家的信息总量为100GB，普通用户的信息总量为300GB，其余数据库信息总量估计在10GB以内。结构化数据存储所需空间总量约为410GB。

## 非结构化数据存储规模估计

非结构化信息包含专家资源和网站资源，下面将根据这两部分进行分析。

专家资源部分：项目预估专家数量为2万人，每位专家的科技资源数量为10项，科技资源分为论文、专利和项目。论文和专利为PDF形式，大小在2MB左右。项目的大小应根据项目具体种类而定，因此估计每位专家科技资源总量占用50MB存储空间，所有专家数据存储所需空间为1000GB。

网站资源部分：网站资源分为网站图片资源和网站日志资源。网站日志资源部分为文本类信息，空间占用量小于网站图片资源。该项目中网站图片资源主要包括专家个人主页背景，网站其他图片等，预计共占用空间3MB，因此网站资源数据总量储存空间大小约为30GB。

综上所述，本系统非结构化数据存储所需要的空间为1030GB。

据以上分析，结构化和非结构化数据存储共需要空间1030GB。另外，为保证可扩展性以及一些未预料到的情况，我们在此基础上增加了10%的空间作为缓冲。因此，本系统数据存储大概需要1133GB的数据存储空间。

另一方面，为防止数据丢失，还需要在远程服务器上对系统数据进行备份。因此，我们需要两倍的数据存储空间，也就是2266GB的存储空间。

# 数据部署方案

## 数据库部署方案

### Master HA

MHA（Master HA）是一款开源的MySQL的高可用程序，它为MySQL主从复制架构提供了自动故障切换功能。MHA在监控到master节点故障时，会提升其中拥有最新数据的slave节点成为新的master节点，在此期间，MHA会通过与其它从节点获取额外信息来避免一致性方面的问题。MHA还提供了master节点的在线切换功能，即按需切换master/slave节点。

相较于其它HA软件，MHA的目的在于维持MySQL Replication中Master库的高可用性，其最大特点是可以修复多个Slave之间的差异日志，最终使所有Slave保持数据一致，然后从中选择一个充当新的Master，并将其它Slave指向它。

### mysql-prox

Proxy是官方提供的MySQL中间件产品，处于客户端应用程序和MySQL服务器之间，通过截断、改变并转发客户端和后端数据库之间的通信来实现其功能，例如负载平衡，读写分离等。mysql-proxy是一个中间层代理，负责将前台应用的连接请求转发给后台的数据库，并且可以实现复杂的连接控制和过滤，从而实现读写分离和负载平衡。

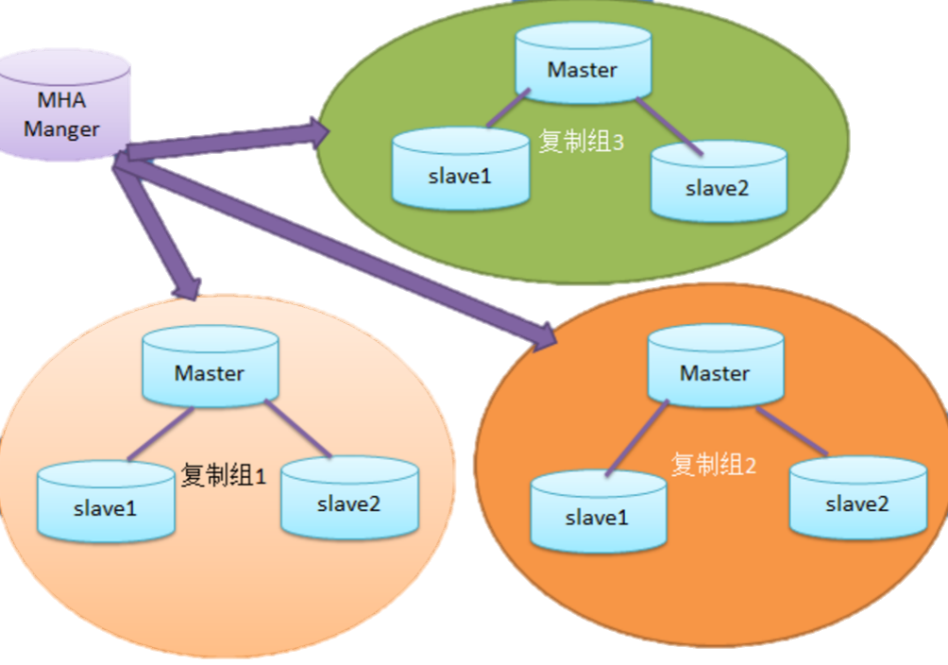
### LVS

LVS即Linux Virtual Server，是我们国家的章文嵩博士的一个开源项目。在Linux 2.6内核中，它已经成为内核的一部分。LVS主要用于多服务器的负载均衡。它工作在网络层，可以实现高性能，高可用的服务器集群技术。它廉价，可把许多低性能的服务器组合在一起形成一个超级服务器。它易用，配置非常简单，且有多种负载均衡的方法。它稳定可靠，即使在集群的服务器中某台服务器无法正常工作，也不影响整体效果。另外可扩展性也非常好。

### 具体部署方案

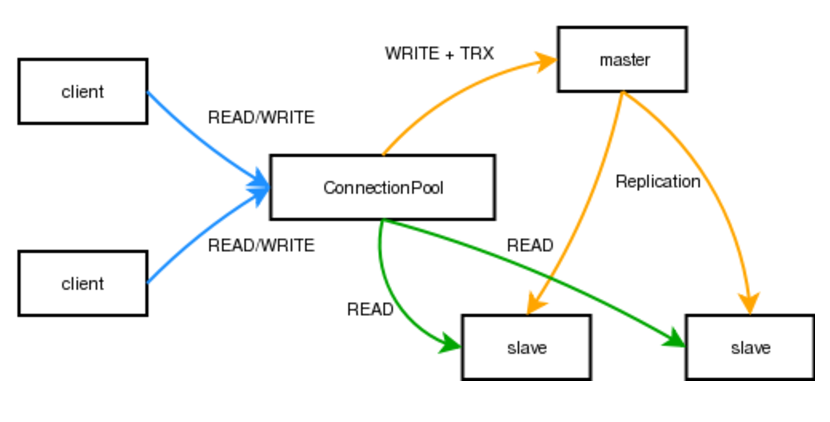
根据数据规模估计结果，我们使用一主多从的数据库集群部署方案。我们计划部署有3个节点的集群，建立mysql的主从结构，部署1台服务器作为master节点，部署2台服务器作为slave节点。

使用MHA完成一主多从架构，实现数据库的高可用。当Master故障时，MHA会从多个Slave中随机选择一个充当新的Master，也可在配置文件中指定某一个Slave优先成为Master。



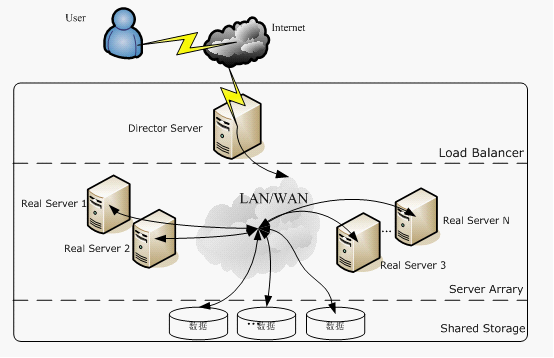
**图5.1-1 MHA原理图**

使用mysql-proxy完成部署数据库的读写分离。mysql-proxy将介入MySQL客户端和服务器之间的通信协议。首先mysql-proxy以服务器的身份接受客户端请求，根据配置对这些请求进行分析处理，然后以客户端的身份转发给相应的后端数据库服务器，带服务器消息返回，接受服务器的信息，返回给客户端。



**图5.1-2 mysql-proxy原理图**

使用LVS实现部署数据库接受请求的负载均衡问题。作为LVS调度器，当接收客户端的读请求并根据LVS调度规则，把请求调度到后方节点。



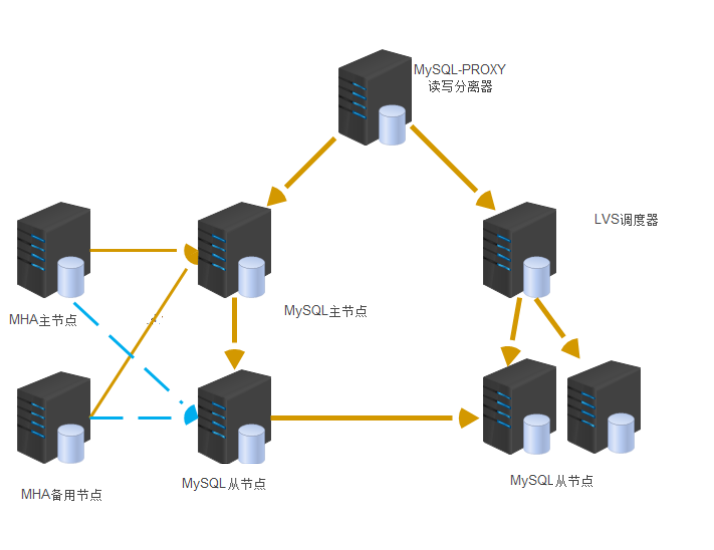
**图5.1-3 LVS原理图**

配置MHA主节点，监控mysql的各个节点，实现MySQL的高可用，在下方设置一台服务器作为MHA的备用节点，在MHA主服务器端运行MHA监测脚本，当主MHA服务器宕机，自动使用备用MHA节点。

配置MySQL集群的主节点，用来执行客户端的写操作，采用主从同步的模式，实时同步主节点中的内容。

配置mysql-proxy实现读写分离，提升整个MySQL集群结构的性能。

最终部署图如下图所示。

****

**图5.1-4 数据库部署架构图**

## 非结构化数据部署方案

### OpenStack Swift技术

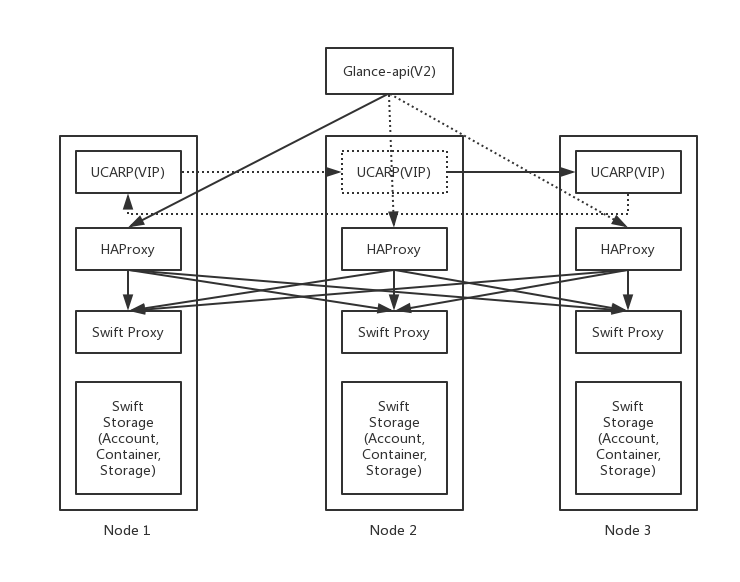
OpenStack Swift，即OpenStack Object Storage是OpenStack开源云计算项目的子项目之一。Swift的目的是使用普通硬件来构建冗余的、可扩展的分布式对象存储集群，存储容量可达PB级。Swift并不是文件系统或者实时的数据存储系统，它是对象存储，用于永久类型的静态数据的长期存储，这些数据可以检索、调整，必要时进行更新。

Swift中各节点可以完全对等，采用完全对称的系统架构，能极大地降低系统维护成本。

数据存储容量与Swift性能都具有很好的可扩展性。Swift的元数据存储是完全均匀随机分布的，并且与对象文件存储一样，元数据也会存储多份，因此消除了单点故障。

### 具体部署方案

根据数据规模估计，我们在3个节点上部署swift，对于非结构化数据进行存储。对于三个节点，我们使用相同的配置，部署图如下：



**图5.2 Swift部署图**

1. 数据库安全措施

数据库的特点之一是由数据库管理系统提供统一的数据保护功能来保证数据的安全可靠和正确有效。数据库的安全性是指保护数据库以防止不合法使用所造成的数据泄露，更改或破坏。系统安全保护措施是否有效是数据库系统的主要技术指标之一。

* 1. 数据库安全问题
     1. 非授权用户对数据库的恶意存取和破坏

一些黑客在用户存取数据库时盗取用户名和用户口令，然后假冒合法用户偷取，修改甚至破坏用户数据，因此必须阻止有损数据库安全的非法操作，以保证数据免受未经授权的访问和破坏。

* + 1. 安全环境的脆弱性

数据库的安全性与计算机系统的安全性，包括计算机硬件，操作系统，网络系统等的安全性是紧密联系的。操作系统安全的脆弱，网络协议安全保障的不足等都会造成数据库安全性的破坏。

* + 1. 数据过载

数据量的突然增多可能导致资源存储空间不足，引发数据损坏。因此要做好预算，留下充足的弹性存储空间。

* + 1. 物理损坏

存储设备可能受到来自自然的物理损害，例如火灾，设备老化等。

* 1. 数据库安全措施
     1. 用户身份鉴别

每次用户要求进入系统时，由系统进行核对，通过鉴定后才提供使用数据库管理系统的权限。

* + 1. 存取控制

确保只授权给有资格的员工访问数据库的权限，同时令所有未被授权的人员无法结接近数据，通过定义用户权限和合法权限检查两部分实现。

* + 1. 数据加密

对数据库重要数据进行存储加密，防止被恶意盗用；同时为保证数据不被恶意截取，保证客户和服务器的安全数据交换，对传输数据使用传输加密。

* + 1. 数据备份

定期将数据上传到网络服务器备份，防止由于突发事件（或者操作者失误）导致的设备损坏，同时也有利于数据维护。