**全网视频搜索数据库设计文档**

盛大创新院搜索主题院

|  |  |  |  |
| --- | --- | --- | --- |
| **版本号** | **修改日期** | **作者** | **修改内容** |
| 1.0 | 2011-11-22 | 徐从斌 | 全文 |
| 1.1 | 2011-11-24 | 王家军、徐从斌 | Review |
| 1.2 | 2011-12-05 | 王家军 | 增加视频、电影、电视剧、专辑、圈子、统计等数据库设计 |
| 1.3 | 2011-12-07 | 王家军 | 参照12.06号原型更新 |
| 1.4 | 2011-12-08 | 王家军 | 按会议沟通结果更新 |
| 1.5 | 2011-12-09 | 王家军 | 增加酷六排行榜和缤纷来源 |
| 1.6 | 2011-02-15 | 王家军，陈运文 | 增加圈子缩略图字段和说明 |
| 1.7 | 2011-02-17 | 王家军 | 视频表增加站外播放次数和站外评论次数 |
| 1.8 | 2011-02-20 | 王家军 | 视频站外播放次数和评论次数移到统计表中 |
| 1.9 | 2011-02-21 | 王家军 | 视频缩略图从缩略图服务查询去掉thumbnail字段 |
| 2.0 | 2011-03-06 | 王家军 | 索引整理，以及无用表删除 |
| 2.1 | 2011-03-06 | 陈运文，邹溢，姚璐 | 增加个性化推荐模块数据库设计 |
| 2.2 | 2011-03-12 | 王家军 | 圈子表增加Tag、是否认证、是否官方、创建者、创建时间、热门度等属性 |
| 2.3 | 2011-03-13 | 王家军 | 增加候选圈子表，圈子表增加status字段 |
| 2.4 | 2011-03-19 | 王家军 | 圈子统计表中增加圈子类别、状态冗余信息 |
| 2.5 | 2012-04-20 | 徐从斌 | 用户表中增加tags字段 |
| 2.6 | 2012-05-03 | 徐从斌 | 增加用户反馈表 |
| 2.7 | 2012-05-14 | 徐从斌 | 用户表中增加订阅邮件字段 |
| 2.8 | 2012-05-24 | 徐从斌 | 增加友情链接表 |
| 2.9 | 2012-06-18 | 徐从斌 | 增加移动版本表 |
| 3.0 | 2012-07-10 | 徐从斌 | 增加widget表 |

[1. 视频集群 5](#_Toc329710739)

[**1.1.** **视频** 5](#_Toc329710740)

[**1.2.** **电视剧视频关系** 6](#_Toc329710741)

[**1.3.** **专辑视频关系** 6](#_Toc329710742)

[2. 圈子集群 7](#_Toc329710743)

[**2.1.** **圈子** 7](#_Toc329710744)

[**2.2.** **候选圈子** 8](#_Toc329710745)

[**2.3.** **圈子用户关注关系** 9](#_Toc329710746)

[**2.4.** **用户自定义的圈子配置** 9](#_Toc329710747)

[3. 圈子视频关系集群 9](#_Toc329710748)

[**3.1.** **圈子视频关系（算法挖掘）** 10](#_Toc329710749)

[**3.2.** **圈子视频关系（用户添加）** 10](#_Toc329710750)

[**3.3.** **圈内视频更新日志（用户添加）** 10](#_Toc329710751)

[4. 统计集群 11](#_Toc329710752)

[**4.1.** **视频动态** 11](#_Toc329710753)

[**4.2.** **视频统计** 12](#_Toc329710754)

[**4.3.** **圈子动态** 13](#_Toc329710755)

[**4.4.** **圈子统计** 13](#_Toc329710756)

[**4.5.** **圈内视频统计** 15](#_Toc329710757)

[**4.6.** **圈内用户统计** 17](#_Toc329710758)

[**4.7.** **用户动态** 17](#_Toc329710759)

[**4.8.** **需要永久保存的用户动态** 18](#_Toc329710760)

[**4.9.** **用户统计** 19](#_Toc329710761)

[**4.10.** **热搜榜** 19](#_Toc329710762)

[5. 视频全量统计集群 20](#_Toc329710763)

[**5.1.** **视频统计** 20](#_Toc329710764)

[6. Web集群 21](#_Toc329710765)

[**6.1.** **用户** 21](#_Toc329710766)

[**6.1.** **用户关注** 21](#_Toc329710767)

[**6.2.** **第三方帐号连接** 22](#_Toc329710768)

[**6.3.** **视频评论** 22](#_Toc329710769)

[**6.4.** **词表** 23](#_Toc329710770)

[7. Web Redis集群 23](#_Toc329710771)

[**7.1.** **用户视频关系** 23](#_Toc329710772)

[**7.2.** **用户分享圈子** 24](#_Toc329710773)

[**7.3.** **用户个性化推荐** 24](#_Toc329710774)

[**7.4.** **用户反馈** 26](#_Toc329710775)

[**7.5.** **友情链接** 26](#_Toc329710776)

[**7.6.** **移动APP版本列表** 27](#_Toc329710777)

[**7.7.** **Widget表** 28](#_Toc329710778)

[**7.8.** **Widget网站表** 28](#_Toc329710779)

网站数据量预估如下。

用户：注册用户1000万，每日活跃用户数为10万。

# Web集群

访问地址：'127.0.0.1:27017'。

* 1. **用户**

保存用户信息。

imeiwei.user：

|  |  |  |  |
| --- | --- | --- | --- |
| 字段 | 类型 | 是否必需 | 说明 |
| \_id | integer | 是 | 用户ID，10000起 |
| email | string | 是 | 邮箱 |
| is\_email\_verified | integer | 是 | 邮箱是否已验证，0否，1是 |
| password | string | 是 | 密码，明文的md5值 |
| nick | string | 是 | 昵称，限20个字 |
| intro | string | 是 | 自我介绍，限100个字 |
| avatar | object | 是 | 头像 |
| medal | array | 否 | 已获得的勋章列表, 1为关注圈子，2为邀请好友，3为成功创建圈子 |
| tags | array | 否 | 用户的兴趣Tag |
| accept\_subscribe\_email | integer | 是 | 是否接受邮件订阅，0否，1是 |
| last\_login\_time | date | 否 | 上次登录日期 |
| last\_login\_ip | string | 否 | 上次登录IP |
| create\_time | date | 是 | 注册时间 |
| update\_time | date | 是 | 更新时间 |
| 类型：MongoDB  主键：\_id  索引：nick（唯一）  规模：1000万条，1GB  写入：  前端：所有字段 | | | |

* 1. **用户收藏的菜单**

存储用户关注关系，每个用户最多允许关注2000个用户。

imeiwei.user\_following：

|  |  |  |  |
| --- | --- | --- | --- |
| 字段 | 类型 | 是否必需 | 说明 |
| \_id | object id | 是 | 自增长ID |
| user | integer | 是 | 用户ID |
| following | integer | 是 | 被关注的用户ID |
| bidirectional | boolean | 否 | 是否互相关注，true为是，false或没有该字段为否 |
| hidden | boolean | 否 | 是否悄悄关注，true为是，false或没有该字段为否 |
| create\_time | date | 是 | 关注发生时间 |
| 类型：MongoDB  主键：\_id  索引：user+following（唯一）、following、user+create\_time、following+create\_time  规模：1亿条，1GB  写入：  前端：所有字段 | | | |

* 1. **第三方帐号连接**

保存第三方帐号跟本站帐号的对应关系。

imeiwei.user\_connect：

|  |  |  |  |
| --- | --- | --- | --- |
| 字段 | 类型 | 是否必需 | 说明 |
| \_id | integer | 是 | 自增长ID |
| third\_party | integer | 是 | 第三方，1 – 盛大，2 – 新浪，3 – 腾讯，4 – 人人 |
| user\_id | integer | 是 | 用户ID |
| connect\_id | string | 是 | 第三方帐号ID |
| access\_token | string | 是 | 授权信息，JSON格式 |
| create\_time | date | 是 | 创建时间 |
| update\_time | date | 否 | 更新日期 |
| 类型：MongoDB  主键：\_id  索引： user\_id+ third\_party（唯一）、third\_party +connect\_id（唯一）  规模：400万条，4GB  写入：  前端：所有字段 | | | |

# Web Redis集群

访问地址：10.156.24.38:6379（Master）、10.156.24.39:6379（Slave）

* 1. **用户视频关系**

Redis DB号：1。

用户的各种视频列表，包括看过的、推过的、评论过的、以后观看等，保留最近添加的1000个视频。

|  |  |
| --- | --- |
| 类型 | Redis |
| Key | user.video:watched:<UserId> 看过的  user.video:liked:<UserId> 推过的  user.video:commented:<UserId> 评论过的  user.video:watch.later:<UserId> 以后观看  user.video:shared:<UserId> 分享过的  user.video:mooded:<UserId> 心情过的 |
| Value | Sorted Set，Set里的member，对于评论过的为JSON，其它为视频ID，按添加时间倒序，score为添加时间  评论过的，JSON格式如下：  {  video\_id: // 视频ID  content: // 评论内容  time: // 评论时间  } |
| 规模 | 4000万个Key，每个Key 1000条，共400GB |
| 写入 | 前端 |

* 1. **用户反馈**

video\_search.feedback：

|  |  |  |  |
| --- | --- | --- | --- |
| 字段 | 类型 | 是否必需 | 说明 |
| \_id | string | 是 | 自增 |
| user\_id | int | 否 | 用户ID，非必填 |
| answers | object | 否 | 选择题答案 |
| content | string | 是 | 反馈内容 |
| client\_ip | string | 是 | 客户端IP |
| create\_time | date | 是 | 加入日期 |
| 类型：MongoDB  主键：\_id  索引：\_id（唯一）  Shard Key：  规模：10万条，100MB  写入：  web | | | |

* 1. **移动APP版本列表**

video\_search.app\_version：

|  |  |  |  |
| --- | --- | --- | --- |
| 字段 | 类型 | 是否必需 | 说明 |
| \_id | string | 是 | 自增 |
| version | Float | 是 | 版本号 |
| version\_name | string | 是 | 版本名称 |
| source | string | 是 | 更新文件，存入fastdfs |
| size | float | 是 | 文件大小 |
| desc | string | 是 | 版本描述 |
| is\_force | string | 是 | 是否对低版本强制更新 |
| force\_list | array | 否 | 需要强制更新的版本列表 |
| user\_id | string | 是 | 管理员ID |
| create\_time | date | 是 | 加入日期 |
| update\_time | date | 是 | 修改日期 |
| 类型：MongoDB  主键：\_id  索引：\_id（唯一）  Shard Key：  规模：1万条，10MB  写入：  web | | | |