BlogChina Hbase表模式设计

1. RowKey设计：

遵循定长，唯一的原则。因为所以ID都不是定长的，所以Author表使用MD5(AuthorID)，Blog

表使用MD5(AuthorID)\_MD5(BlogID)，Comment使用MD5(BlogID)\_MD5(Comment\_ID)。

1. 表模式设计：

SpiderInfo表：

|  |  |  |
| --- | --- | --- |
| Rowkey | spiderName | time |
|  | blogchina | Starttime |
|  |  | endtime |

Author表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RowKey | 列族 | | | |
| MD5(AuthorID) | **~~SpiderInfo~~** | **authorInfo** | **focuse** | **Fans** |
| ~~Time~~ | Author\_ID | focuseNum | fansNum |
|  | Author\_name | 0 | 0 |
| Author\_blog\_name | 1 | 1 |
| introduce | … | … |
| Image\_url | focuseNum-1 | fansNum-1 |
| b\_image |  |  |
| article\_num |
| read\_num |
| fans\_num |
| focuse\_num |
| all\_article\_url |

Blog表：

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| RowKey | 列族 | | | |
| MD5(AuthorID)\_MD5(BlogID) | **~~SpiderInfo~~** | **bloginfo** | **Pctures** | **comments** |
| ~~time~~ | blog\_id | PictureNum | Comment\_num |
|  | author\_id | 0 | 0 |
| title | 1 | 1 |
| sub\_title | … | … |
| publish\_time | PictureNum-1 | Commemt\_num-1 |
| category |  |  |
| picture\_urls |
| read\_num |
| comment\_num |
| hand\_up\_num |
| hand\_down\_num |
| Content |
| url |

Comment表：

|  |  |  |  |
| --- | --- | --- | --- |
| RowKey | 列族 | | |
| MD5(BlogID)\_MD5(Comment\_ID) | **~~Spider\_info~~** | **Comment\_info** | **Praise\_info** |
| ~~Time~~ | Comment\_id | Praise\_num |
|  | comment\_user\_id | 0 |
| comment\_blog\_id | 1 |
| comment\_time | … |
| comment\_content | Praise\_num - 1 |
| praise\_num |  |
| replay\_id |
| ip |
| last\_ip |

create 'spider\_info', {NAME => 'spider\_name', VERSIONS => 1}, {NAME => 'time', VERSIONS => 2}

create 'blogchina\_author', {NAME => 'author\_info', VERSIONS => 5}, {NAME => 'focuse', VERSIONS => 10}, {NAME => 'fans', VERSIONS => 10}

create 'blogchina\_blog', 'blog\_info', 'pictures', {NAME => 'comments', VERSIONS => 10}

create 'blogchina\_comment', 'comment\_info', 'praise\_info'