1，通过get方式，指定rowkey获取唯一一条记录

2，通过scan方式，设置起始行和结束行参数进行范围匹配

3，全表扫描，即直接扫描整张表中所有行记录

表 1. HBase 与 RDBMS 的区别

|  | **HBase** | **RDBMS** |
| --- | --- | --- |
| 硬件架构 | 类似于 Hadoop 的分布式集群，硬件成本低廉 | 传统的多核系统，硬件成本昂贵 |
| 容错性 | 由软件架构实现，由于由多个节点组成，所以不担心一点或几点宕机 | 一般需要额外硬件设备实现 HA 机制 |
| 数据库大小 | PB | GB、TB |
| 数据排布方式 | 稀疏的、分布的多维的 Map | 以行和列组织 |
| 数据类型 | Bytes | 丰富的数据类型 |
| 事物支持 | ACID 只支持单个 Row 级别 | 全面的 ACID 支持，对 Row 和表 |
| 查询语言 | 只支持 Java API （除非与其他框架一起使用，如 Phoenix、Hive） | SQL |
| 索引 | 只支持 Row-key，除非与其他技术一起应用，如 Phoenix、Hive | 支持 |
| 吞吐量 | 百万查询/每秒 | 数千查询/每秒 |

表 2. 数据在 RDBMS 中的排布示例

| **ID** | **姓** | **名** | **密码** | **时间戳** |
| --- | --- | --- | --- | --- |
| 1 | 张 | 三 | 111 | 20160719 |
| 2 | 李 | 四 | 222 | 20160720 |

表 3. 数据在 HBase 中的排布（逻辑上）

| **Row-Key** | **Value（CF、Qualifier、Version）** |
| --- | --- |
| 1 | info{'姓': '张'，'名':'三'} pwd{'密码': '111'} |
| 2 | Info{'姓': '李'，'名':'四'} pwd{'密码': '222'} |

表 4. 数据在 HBase 中的排布

| **Row-Key** | **CF:Column-Key** | **时间戳** | **Cell Value** |
| --- | --- | --- | --- |
| 1 | info:fn | 123456789 | 三 |
| 1 | info:ln | 123456789 | 张 |
| 2 | info:fn | 123456789 | 四 |
| 2 | info:ln | 123456789 | 李 |

问题：java.io.IOException: Could not locate executable null\bin\winutils.exe in the Hadoop binaries.

解决问题：

-Dhadoop.home.dir=D:\virtual-hadoop

hbase shell

|  |
| --- |
| scan 'user\_category\_push\_data', FILTER => "PrefixFilter('12345677')"  scan 'user\_category\_push\_data', {STARTROW=>'12345677@5',STOPROW=>'12345677@4', FILTER => "PrefixFilter ('12345677')"} |

hbase是k，value存储，不适合做分页，适合查询查询单挑数据

hbasejdbc连接，连接问题很多都是跟项目的其他jar包冲突