本文的组成部分

本文主要有两部分

第一部分为实际场景描述

第二部分为根据实际场景分析出来的测试条件

本文的目的：

1. 说明benchmark测试程序的写入查询负载设计实际的场景来源；
2. 说明测试数据库性能的测试流程以及测试过程中所使用到的查询条件；
3. 说明测试数据库性能的测试结果数据保存的表结构。

第一部分 实际场景描述

基本描述：风力发电站有多个风车，每个风车有多个传感器，传感器采集数据（风机的速度，温度等）以一定的频率保存到数据库中；各个风车都应相应的监控面板，用来监控风车的各个部件的运转情况，监控面板的数据来自于风车上的传感器所采集的值。

实际场景描述

实际场景一：总控部门监控员发现某个风车数据异常报警，打电话通知风车终端负责人员排查核实是否存在异常，风车终端部门人员经核实该终端对应的某个监测点数据的确存在异常，指派相应维修员排查相应监测点故障，并维修，在这个流程中，涉及到各个部门人员进行多次查询。

改业务过程中涉及到的查询

1. 查看一共有多少个风车
2. 查看所有的风车每个风车有多少个传感器
3. 为了实时检测是否传感器异常，如果有异常，对相应的风车报警，查询操作:查看每个风车，对应的每个传感器，一天内，每个小时数据的平均值，最大值，最小值。（设备查看页面，只能查看设备是否正常，无法看到具体哪个传感器异常）
4. 当i号风车存在异常时，为了检测查看i号风车具体是哪个传感器有异常，对相应的传感器报警。查询一天内，每个小时数据的平均值，最大值，最小值设备详情监控页面，可以查看对应设备的运行是否正常）
5. 当发现是具体是哪个传感器异常时，排除具体是哪个分钟段数据异常，查看i号风车m号传感器，1个小时内，每分钟的平均值，最大值，最小值
6. 确定好具体是哪个分钟段数据异常后，查看对应时间点前后15分钟的具体数据，确定具体是从哪个时间点开始数据异常的，查看i号风车m号传感器，15分钟内的所有数据值

在排除传感器数据是否正常时，还有以下几项操作：

1. 查看i号风车，m号传感器， k小时或者 l分钟内的大于某个值的数据
2. 查看i号风车，m号传感器，k小时或者 l分钟内的小于某个值的数据
3. 查看i号风车，m号传感器，k小时或者 l分钟内的等于某个值的数据
4. 查看i号风车，m号传感器，k小时或者 l分钟内的不等于某个值的数据

补充:某个值是表示正常数据的值

实际场景二：需要增加监控面板的情形

由于风车的数量规模扩大，需要增加更多的监控面板；每个风车每个传感器的状况实时监控，增加监控面板的电视墙（也可以是别的东西）。测试最大支持同时运行多少个监控面板。

监控面板包含的数据查询操作

1. 查看i号风车，所有传感器，一天内，每分钟的平均值。
2. 查看i号风车，所有传感器，一天内，每分钟的最大值。
3. 查看i号风车，所有传感器，一天内，每分钟的最小值。
4. 查看i号风车，所有传感器，过去一分钟内的所有数据。

实际场景三：传感器厂商把传感器卖给风车厂商后，需要检测自己出售出去的传感器是否运行正常，传感器坏掉后，根据数据排查传感器的坏掉原因。

具体描述: 传感器提供商查看提供给风力发电站的传感器运行状态，传感器出异常后查询故障原因。

比如厂家提供的一个传感器，一般寿命为1年，而这个传感器用了一个月就不行了，分析这个传感器在使用过程中是否有极端状况。

传感器提供商的技术支持部门，利用风力发电站提供的数据接口，监控自己厂商所生产的传感器是否有运行异常的，没有记录数据的。

涉及到的查询

1. 查看所有风车，k号传感器，k小时或者 l分钟内的的空值数
2. 查询所有风车，k号传感器，k小时或者 l分钟内的最大值，最小值，均值
3. 查看i号风车，k号传感器，k小时或者 l分钟内的小于某个数所有值
4. 查看i号风车，k号传感器，k小时或者 l分钟内的大于某个数所有值

第二部分 测试流程和测试条件

综合第一部分三种实际场景，以及各个场景内的具体进行的操作，整理以下测试方案。

测试流程为

1，数据生成器生成数据；

2，数据生成的数据导入到目标数据库中，同时测试目标数据库的导入性能；

1. 测试目标数据库的没有写入负载下的查询性能并记录结果到关系数据库；
2. 测试目标数据库的有写入负载下的查询性能并记录结果到关系数据库；
3. 测试目标数据库的无查询负载的实时写入性能并记录结果到关系数据库；

6，测试目标数据库的有查询负载的实时写入性能并记录结果到关系数据库；

7，停止测试程序。

解释：

查询负载：数据库在稳定运行时，真实生产场景的查询负载。（目前无法确定，使用的是最大负载的0.2，比如最大支持每秒钟1000个查询，这里的查询负载就是每秒钟200个查询）

写入负载：数据库在稳定运行时，真实生产场景的写入负载。

数据导入过程描述

数据生成器生成过去一个月，一百个设备(风车)，每个设备500个传感器的数据，共18,514,285,714条数据，将数据导入到测试数据库中，测试每秒钟可以导入多少个数据点(data points/sec) 、或者每秒钟导入的数据量(MB/sec)；数据生成过程中会以一定概率漏掉一些数据。

测试性能过程描述

1. 选定要执行的操作；
2. 通过增加每秒钟的执行次数测试每秒钟能执行的最大的请求数/写入数；

查询是通过增加客户端提供每秒钟的执行次数，写入是通过增加风车数增加每秒钟的执行次数；

1. 简单查询
   1. 查询某个风车某个传感器某个15分钟内的数据

select s\_x from point where device=’?’ and time>=beginTime and time<endTime

其中endTime-beginTime=15Min ,beginTime,?为变量；

* 1. 查询某个风车某个传感器在某个15分钟内值大于某个数的数据

select s\_x from point where device=’?’ and time>=beginTime and time<endTime and s\_x>?

其中endTime-beginTime=15minutes,beginTime,?为变量；

* 1. 查询某个风车所有传感器在过去一分钟的值

select s1,s2….. s500 from point where device=’?’ and time>=beginTime and time<endTime

其中endTime-beginTime=1minutes,beginTime,?为变量；

1. 分析查询

2.1 查询所有风车，所有传感器，过去一天内，每个小时的最大值/最小值/平均值

select max/min/mean(s1), max/min/mean (s2)….. max/min/mean(s500) from point where time>=beginTime and time<endTime group by time(1H)

其中endTime-beginTime=1day,beginTime为变量；

2.2 查询某个风车，所有传感器，过去一天内，每分钟的最大值/最小值/平均值

select max/min/mean(s1), max/min/mean (s2)….. max/min/mean(s500) from point where device=? time>=beginTime and time<endTime group by time(1H)

其中endTime-beginTime=1day,beginTime，?为变量；

2.3 查询某个风车，某个传感器，在一个小时内，每分钟的最大值，最小值，平均值

select max/min/mean(s\_x) from point where device=? time>=beginTime and time<endTime group by time(1h)

其中endTime-beginTime=1day, s\_x ,beginTime，?为变量；

1. 简单查询和分析查询混合

将简单查询和分析查询结合真实场景进行混合，测试每秒可以处理的场景数/请求数。

选择1.1和2.1两种查询进行组合。

1. 写入

按照真实的数据写入场景，进行数据写入。

关系数据测试结果数据表描述

1，写入性能表

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段类型 | 描述 |
| id | bigint | 主键id自动生成 |
| db\_type | int | 测试的目标数据库类型 |
| batch\_id | int | 测试批次，计划改为可配置的，目前是系统默认生成的，每次运行程序，会生成一个批次号 |
| operate\_time | timestamp | 写入时间 |
| target\_devices\_per | int | 目标每秒钟写入的设备数 |
| target\_lines\_per | int | 目标每秒中写入的数据点数 |
| actual\_lines\_per | int | 实际每秒中写入的数据点数 |

2，查询性能表

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段类型 | 描述 |
| id | bigint | 主键id自动生成 |
| db\_type | int | 测试的目标数据库类型 |
| batch\_id | int | 测试批次，计划改为可配置的，目前是系统默认生成的，每次运行程序，会生成一个批次号 |
| operate\_time | timestamp | 查询时间 |
| target\_times\_per | int | 目标每秒的查询数 |
| actual\_times\_per | int | 实际每秒的查询数 |
| query\_type | int | 查询类型，简单读，分析读 |
| query\_desc | varchar(255) | 查询描述，描述查询字段和查询条件 |

3，写入负载记录表

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段类型 | 描述 |
| id | bigint | 主键id自动生成 |
| db\_type | int | 测试的目标数据库类型 |
| batch\_id | int | 测试批次，计划改为可配置的，目前是系统默认生成的，每次运行程序，会生成一个批次号 |
| operate\_time | timestamp | 写入时间 |
| target\_devices\_per | int | 目标每秒钟写入的设备数 |
| target\_lines\_per | int | 目标每秒中写入的数据点数 |
| actual\_lines\_per | int | 实际每秒中写入的数据点数 |

4，查询负载记录表

|  |  |  |
| --- | --- | --- |
| 字段名称 | 字段类型 | 描述 |
| id | bigint | 主键id自动生成 |
| db\_type | int | 测试的目标数据库类型 |
| batch\_id | int | 测试批次，计划改为可配置的，目前是系统默认生成的，每次运行程序，会生成一个批次号 |
| operate\_time | timestamp | 查询时间 |
| target\_times\_per | int | 目标每秒的查询数 |
| actual\_times\_per | int | 实际每秒的查询数 |
| query\_type | int | 查询类型，简单读，分析读 |
| query\_desc | varchar(255) | 查询描述，描述查询字段和查询条件 |

上述描述的功能对Influxdb测试的基本版本已经开发完毕。