**IoTDB与InfluxDB数据存取性能测试报告**

文档变更记录

|  |  |  |  |
| --- | --- | --- | --- |
| 版 本 | 日 期 | 描 述 | 作 者 |
| V0.1 | 2017年11月21日 | IoTDB与InfluxDB数据存取性能测试报告v0.1 | 刘睿 |
| V0.2 | 2017年11月23日 | 增加写入测试中的不同编码方式以及度量指标，增加实验机器，增加查询测试 | 刘睿 苏月 |
| V0.3 | 2017年11月27日 | 增加写测试实验结果，异常测试结果说明 | 刘睿 苏月 |
| V0.4 | 2017年12月1日 | 新增MySQL数据库说明和测试结果以及实验参数修改 | 刘睿 苏月 |
| V0.5 | 2017年12月6日 | 新增V1.4.2版InfluxDB写测试结果并绘制写入速率图像，新增查询测试结果 | 刘睿 |
| V0.6 | 2017年12月11日 | 新增InfluxDB写测试、写入速率图像、IoTDB数据点耗存和查询测试结果 | 刘睿 苏月 |
| V0.7 | 2017年12月18日 | 新增InfluxDB查询测试结果 | 刘睿 苏月 |
|  |  |  |  |

文档审核记录

|  |  |  |
| --- | --- | --- |
| 审 核 | 审核日期 | 所属组织/部门 |
|  | 2017年x月x日 | 清华 |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |
|  |  |  |

**目 录**

[1 引言 4](#_Toc501322464)

[1.1测试背景 4](#_Toc501322465)

[1.2测试目的 4](#_Toc501322466)

[1.3测试时间 4](#_Toc501322467)

[2 测试资源和测试环境 5](#_Toc501322468)

[2.1硬件的配置 5](#_Toc501322469)

[2.2软件配置 5](#_Toc501322470)

[3 写入性能测试 6](#_Toc501322471)

[3.1业务背景 6](#_Toc501322472)

[3.2 实验目的 6](#_Toc501322473)

[3.3 实验设计 6](#_Toc501322474)

[3.4实验结果 9](#_Toc501322475)

[3.4.1不同参数变化时InfluxDB与IoTDB的性能对比 9](#_Toc501322476)

[3.4.2 IoTDB系统性能指标与写入速率 14](#_Toc501322477)

[3.4.3 InfluxDB系统性能指标与写入速率 16](#_Toc501322478)

[3.5实验结论 18](#_Toc501322479)

[4 范围查询性能测试 20](#_Toc501322480)

[4.1业务背景 20](#_Toc501322481)

[4.2 实验目的 20](#_Toc501322482)

[4.3 实验设计 20](#_Toc501322483)

[4.4实验结果 23](#_Toc501322484)

[4.2 网络IO使用量： 27](#_Toc501322485)

[4.4.3 CPU占用率： 27](#_Toc501322486)

[4.4.4磁盘IO： 27](#_Toc501322487)

[4.5实验结论 27](#_Toc501322488)

[5 条件查询性能测试 27](#_Toc501322489)

[5.1业务背景 27](#_Toc501322490)

[5.2 实验目的 28](#_Toc501322491)

[5.3 实验设计 28](#_Toc501322492)

[5.5实验结论 36](#_Toc501322493)

[6 精确点查询性能测试 36](#_Toc501322494)

[6.1业务背景 36](#_Toc501322495)

[6.2 实验目的 36](#_Toc501322496)

[6.3 实验设计 37](#_Toc501322497)

[6.5实验结论 42](#_Toc501322498)

[7 最近点查询性能测试 42](#_Toc501322499)

[7.1业务背景 42](#_Toc501322500)

[7.2 实验目的 42](#_Toc501322501)

[7.3 实验设计 43](#_Toc501322502)

[7.5实验结论 46](#_Toc501322503)

[8 聚合查询性能测试 46](#_Toc501322504)

[8.1业务背景 46](#_Toc501322505)

[8.2 实验目的 46](#_Toc501322506)

[8.3 实验设计 47](#_Toc501322507)

[9 GroupBy查询性能测试 53](#_Toc501322508)

[9.1业务背景 53](#_Toc501322509)

[9.2 实验目的 53](#_Toc501322510)

[9.3 实验设计 53](#_Toc501322511)

[10 模糊点查询性能测试 56](#_Toc501322512)

# 1 引言

## 1.1测试背景

本次测试工作的测试对象是清华大学开发的IoTDB，为了进行性能对比同时也将测试相同参数下InfluxDB的性能。测试工作将在清华大学大数据工程实验室完成。用于测试的服务器和硬件测试环境由实验室提供。测试数据集是通过iotdb-benchmark工具随机生成的，数据值的类型包括整数和浮点型。iotdb-benchmark的开发工作以及测试工作由测试小组（刘睿、江天、苏月）负责完成。

## 1.2测试目的

IoTDB作为青海项目传感器数据的历史时序数据库，为传感器时序数据提供高吞吐量的数据写入与数据查询功能，本次实验将对IoTDB的数据写入及数据查询功能的性能与可靠性进行测试。最终目的是根据测试实验结果了解不同参数对系统性能的影响，从而对参数进行调优和最优化配置。测试目标包括两个方面：

* 存储性能测试：

测试IOTDB对大批量工程机械数据写入的性能和稳定性。

* 读取性能测试：

测试IOTDB在已经存储了大规模工况数据的前提下的数据读取性能和聚合计算性能。

## 1.3测试时间

测试时间分为两个阶段：

第一阶段：2017年11月25日至2017年11月27日。主要完成数据写入性能测试和稳定性测试。

第二阶段：2017年12月10日至2017年12月17日。主要完成数据读取性能测试和聚合计算性能测试。（实际时间未定）

# 2 测试资源和测试环境

## 2.1硬件的配置

本次实验使用实验室的7台服务器，配置如下，选择其中3台机器作为IoTDB/InfluxDB服务器，其他机器作为客户机。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 服务器名 | IP | 用户名密码 | CPU | 内存 | 硬盘 | 备注 |
| 6服务器 | 192.168.130.22 | 用户名：liurui  密码：liurui\_2017 | 1\*8核 | 32G | 1.8T | 客户机 |
| 7服务器 | 192.168.130.7 | 用户名：liurui  密码：liurui\_2017 | 1\*8核 | 32G | 1.8T | 服务器 |
| 8服务器 | 192.168.130.8 | 用户名：liurui  密码：liurui | 1\*8核 | 32G | 1.8T | 客户机 |
| 9服务器 | 192.168.130.9 | 用户名：liurui  密码：liurui | 1\*8核 | 32G | 1.8T | 客户机 |
| 10服务器 | 192.168.130.23 | 用户名：liurui  密码：liurui\_2017 | 1\*8核 | 32G | 1.8T | 服务器 |
| 19服务器 | 192.168.130.19 | 用户名：liurui  密码：liurui\_2017 | 1\*8核 | 32G | 1.8T | 服务器 |
| 20服务器 | 192.168.130.20 | 用户名：liurui  密码：liurui\_2017 | 1\*8核 | 32G | 1.8T | 服务器 |
| 21服务器 | 192.168.130.21 | 用户名：liurui  密码：liurui\_2017 | 1\*8核 | 32G | 1.8T | 客户机 |

## 2.2软件配置

|  |  |
| --- | --- |
| 操作系统 | Ubuntu14.04.3 LTS |
| Java版本 | 1.8 |
| InfluxDB版本 | 1.4.2 |
| IoTDB版本 | 0.4.0-SNAPSHOT |
| MySQL服务器端口 | 166.111.141.168:3306 |
| MySQL用户名 | root |
| MySQL密码 | Ise\_Nel\_2017 |

# 3 写入性能测试

## 3.1业务背景

假定青海项目现在对IoTDB的需求是管理1万台设备的传感器数据。其中每台设备有300个传感器，平均数据发送频率为5秒，那么IOTDB要接收的写入请求速率大约为每秒钟60万个数据点。

## 3.2 实验目的

测试IoTDB目前每秒钟可以处理的写入数据点数与系统稳定性，验证系统写入性能能否满足青海项目现在的数据写入需求。最终目的还是根据测试实验结果了解不同参数对系统性能的影响，从而对参数进行调优和最优化配置。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
| GROUP\_NUMBER | 总存储组数 |
| DEVICE\_NUMBER | 总设备数 |
| SENSOR\_NUMBER | 每个设备sensor数 |
| CLIENT\_NUMBER | 客户端线程数（并发度），给每个客户端均匀分配Device |
| CACHE\_NUM | 批量导入数量 |
| LOOP | 循环次数，其他参数不变时，主要通过这个参数改变数据量 |
| ENCODING | 编码方式 |

## 3.3 实验设计

#### 3.3.1 度量指标

1. 性能指标
2. 总导入数据点数：数值上等于DEVICE\_NUMBER\* SENSOR\_NUMBER\* CACHE\_NUM\*LOOP。
3. 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。
4. 平均写入速率：数值上等于总导入数据点数/总时间。
5. 客户端批写入速率：每个客户端线程在每个批写入任务中的平均写速率。
6. 单位耗存：使用不同编码方式下平均每个数据点存储所占的字节数。
7. 系统信息指标
8. 客户机及服务器的磁盘IO使用率。
9. 客户机及服务器的CPU使用率。
10. 客户机及服务器的内存使用率。
11. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 3.3.1 实验数据

**1）实验数据说明：**

实验采用iotdb-benchmark工具自动生成的数据，不同传感器有不同的数据分布，数据类型均为DOUBLE类型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。以如下表所设的参数值为例具体地说明产生的数据集：

|  |  |
| --- | --- |
| 参数名 | 值 |
| GROUP\_NUMBER | 2 |
| DEVICE\_NUMBER | 10 |
| SENSOR\_NUMBER | 20 |
| CACHE\_NUM | 5 |
| LOOP | 4 |

总共产生的timeseries个数为10\*20=200，10个device被2个存储组均分:

root.group\_0.d\_0.s\_0 root.group\_0.d\_0.s\_1 … root.group\_0.d\_0.s\_19

root.group\_0.d\_1.s\_0 root.group\_0.d\_1.s\_1 … root.group\_0.d\_1.s\_19

…

root.group\_0.d\_4.s\_0 root.group\_0.d\_4.s\_1 … root.group\_0.d\_4.s\_19

root.group\_1.d\_5.s\_0 root.group\_1.d\_5.s\_1 … root.group\_1.d\_5.s\_19

…

root.group\_1.d\_9.s\_0 root.group\_1.d\_9.s\_1 … root.group\_1.d\_9.s\_19

一个batch中包含5条insert语句类似于：

insert into root.group\_0.d\_0 (timestamp,s\_0,s\_1…s\_19) value(0,2,42.4,…,42)

insert into root.group\_0.d\_0 (timestamp,s\_0,s\_1…s\_19) value(5000,7,31.3,…,12)

…

insert into root.group\_0.d\_0 (timestamp,s\_0,s\_1…s\_19) value(20000,5,91.3,56)

**2）实验参数设置：**

**理论上应选以下参数不同取值的笛卡尔积作为所有可能的实验参数组合：**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 参数名 | 取值 | | | | | | | |
| GROUP\_NUMBER | 1 | 5 | | 10 | 50 | 100 | | 500 |
| DEVICE\_NUMBER | 10 | 100 | | 500 | 1000 | 5000 | | 10000 |
| SENSOR\_NUMBER | 10 | 50 | | 100 | 300 | 500 | | 1000 |
| CLIENT\_NUMBER | 1 | 5 | | 10 | 20 | 50 | | 100 |
| CACHE\_NUM | 1 | 100 | | 500 | 1000 | 5000 | | 10000 |
| LOOP | 10 | 50 | | 100 | 500 | 1000 | | 3000 |
| ENCODING | PLAIN | | RLE | | TS\_2DIFF | | GORILLA | |

实际上为了减小测试空间我们只选择比较有意义的组合参数取值，并在有意义的参数组合中排一个优先级，根据该优先级确定实验顺序。有意义的参数组合详见3.4实验结果。

#### 3.3.2测试步骤：

**第一步，对IoTDB来说需要先创建元数据（Create Schema）：**

从根据config.properties文件中的参数设置创建元数据。创建时以1000条create timeseries 语句为一个batch批量地提交给IoTDB服务以加快元数据的创建过程。

**第二步，每个客户端并发地生成随机数据并产生插入语句，批量地发送给服务器进行写入：**

启动CLIENT\_NUMBER个线程，每个线程对应DEVICE\_NUMBER/CLIENT\_NUMBER个Device。以批处理的方式提交给IoTDB服务器。每个批次中插入语句的条数即CACHE\_NUM。

**第三步，监控服务器系统状态：**

在IoTDB性能测试进行的过程中，利用iotdb-benchmark监控程序记录系统信息，并根据测试实验执行的长短来确定每隔多久获取一次服务器机器的CPU使用率，网络传输，硬盘IO和内存的性能状态，用于服务器性能及系统瓶颈的分析。

**第四步，统计写入性能指标：**

每个线程内置一个成员变量作为计数器来记录当前线程执行插入语句所用的总时间。当所有线程都完成数据的写入后取总时间最大的值作为整个数据集写入所用的总时间。平均写入速率即总写入数据点数除以该总时间。

## 3.4实验结果

### 3.4.1不同参数变化时InfluxDB与IoTDB的性能对比

1. 数据量（LOOP）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| CACHE\_NUM | 1000 |
| GROUP\_NUMBER | 10 |
| DEVICE\_NUMBER | 100 |
| SENSOR\_NUMBER | 100 |
| CLIENT\_NUMBER | 10 |
| ENCODING | PLAIN |

注意：ENCODING参数指的是IoTDB的编码方式，InfluxDB的编码方式为GORILLA

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| LOOP | InfluxDB | IoTDB |
| 10 | 797047.7352 | 710212.1403 |
| 50 | 780346.5050 | 687558.3565 |
| 100 |  | 641843.5801 |
| 500 | 729043.0027 | 633164.8988 |
| 1000 | 1653222.6599 | 670607.8288 |
| 3000 |  | 701597.0641 |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| LOOP | InfluxDB | IoTDB |
| 10 | 125.463 | 140.803 |
| 50 | 640.741 | 727.211 |
| 100 |  | 1558.012 |
| 500 | 6858.306 | 7896.837 |
| 1000 | 6048.792 | 14911.845 |
| 3000 |  | 42759.586 |

1. 批量导入数量（CACHE\_NUM）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| GROUP\_NUMBER | 10 |
| DEVICE\_NUMBER | 100 |
| SENSOR\_NUMBER | 100 |
| CLIENT\_NUMBER | 10 |
| LOOP | 100 |
| ENCODING | PLAIN |

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| CACHE\_NUM | InfluxDB | IoTDB |
| 1 |  | 713775.8743 |
| 10 |  | 686530.2760 |
| 50 |  | 729671.3560 |
| 100 |  | 715589.1087 |
| 500 |  | 654583.9203 |
| 1000 |  | 653850.7559 |
| 5000 |  | 645752.8833 |
| 10000 |  | 710056.7683 |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| CACHE\_NUM | InfluxDB | IoTDB |
| 1 |  | 1.401 |
| 10 |  | 14.566 |
| 50 |  | 68.524 |
| 100 |  | 139.745 |
| 500 |  | 763.844 |
| 1000 |  | 1529.401 |
| 5000 |  | 7742.9 |
| 10000 |  | 14083.381 |

1. 传感器数量（SENSOR\_NUMBER）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| GROUP\_NUMBER | 10 |
| DEVICE\_NUMBER | 100 |
| CACHE\_NUM | 1000 |
| CLIENT\_NUMBER | 10 |
| LOOP | 100 |
| ENCODING | PLAIN |

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| SENSOR\_NUMBER | InfluxDB | IoTDB |
| 10 |  | 569145.4850 |
| 50 |  | 678706.0063 |
| 100 |  | 590161.3028 |
| 300 |  | 653501.9426 |
| 500 |  | 619281.3240 |
| 1000 |  |  |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| SENSOR\_NUMBER | InfluxDB | IoTDB |
| 10 |  | 175.702 |
| 50 |  | 736.696 |
| 100 |  | 1694.452 |
| 300 |  | 4590.652 |
| 500 |  | 8073.875 |
| 1000 |  |  |

注：红色表示该参数在目前的机器配置、IoTDB版本下无法完成测试

1. 总设备数（DEVICE\_NUMBER）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| GROUP\_NUMBER | 10 |
| CACHE\_NUM | 1000 |
| SENSOR\_NUMBER | 100 |
| CLIENT\_NUMBER | 10 |
| LOOP | 100 |
| ENCODING | PLAIN |

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| DEVICE\_NUMBER | InfluxDB | IoTDB |
| 10 |  | 657246.1387 |
| 100 |  |  |
| 500 |  | 644039.4982 |
| 1000 |  | 643232.4644 |
| 5000 |  | 666515.0212 |
| 10000 |  |  |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| DEVICE\_NUMBER | InfluxDB | IoTDB |
| 10 |  | 152.15 |
| 100 |  |  |
| 500 |  | 7763.499 |
| 1000 |  | 15546.479 |
| 5000 |  | 75017.064 |
| 10000 |  |  |

注：红色表示该参数在目前的机器配置、IoTDB版本下无法完成测试，InfluxDB失败点统计超时，因此未能统计到实际插入点数，该问题尚待解决。

太消耗时间，没有跑

1. 存储组数（GROUP\_NUMBER）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| DEVICE\_NUMBER | 100 |
| CACHE\_NUM | 1000 |
| SENSOR\_NUMBER | 100 |
| CLIENT\_NUMBER | 10 |
| LOOP | 100 |
| ENCODING | PLAIN |

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| GROUP\_NUMBER | InfluxDB | IoTDB |
| 1 |  | 588436.7472 |
| 5 |  | 813835.1983 |
| 10 |  | 796073.7642 |
| 50 |  | 694415.0283 |
| 100 |  | 564409.8925 |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| GROUP\_NUMBER | InfluxDB | IoTDB |
| 1 |  | 1699.418 |
| 5 |  | 1228.75 |
| 10 |  | 1256.165 |
| 50 |  | 1440.061 |
| 100 |  | 1771.762 |

注：红色表示该参数在目前的机器配置、IoTDB版本下无法完成测试

1. 并发数（CLIENT\_NUMBER）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| DEVICE\_NUMBER | 100 |
| CACHE\_NUM | 1000 |
| SENSOR\_NUMBER | 100 |
| GROUP\_NUMBER | 10 |
| LOOP | 100 |
| ENCODING | PLAIN |

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| CLIENT\_NUMBER | InfluxDB | IoTDB |
| 1 |  | 273690.8137 |
| 2 |  | 509157.7106 |
| 4 |  | 817542.4999 |
| 5 |  | 848433.3254 |
| 8 |  | 871066.4815 |
| 10 |  | 773372.4763 |
| 20 |  | 800842.4863 |
| 50 |  | 782344.0593 |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| CLIENT\_NUMBER | InfluxDB | IoTDB |
| 1 |  | 3653.758 |
| 2 |  | 1964.028 |
| 4 |  | 1223.178 |
| 5 |  | 1178.643 |
| 8 |  | 1148.018 |
| 10 |  | 1293.038 |
| 20 |  | 1248.685 |
| 50 |  | 1278.21 |

1. 编码方式（ENCODING）变化

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| DEVICE\_NUMBER | 100 |
| CACHE\_NUM | 1000 |
| SENSOR\_NUMBER | 100 |
| GROUP\_NUMBER | 10 |
| LOOP | 100 |
| CLIENT\_NUMBER | 10 |

平均写入速率Mean Throughput （points/s）

|  |  |  |
| --- | --- | --- |
| ENCODING | InfluxDB | IoTDB |
| PLAIN |  |  |
| RLE |  |  |
| TS\_2DIFF |  |  |
| GORILLA |  |  |

总耗时Cost Time （s）

|  |  |  |
| --- | --- | --- |
| ENCODING | InfluxDB | IoTDB |
| PLAIN |  |  |
| RLE |  |  |
| TS\_2DIFF |  |  |
| GORILLA |  |  |

单位耗存Unit Storage Size （Byte/point）

|  |  |  |
| --- | --- | --- |
| ENCODING | InfluxDB | IoTDB |
| PLAIN |  |  |
| RLE |  |  |
| TS\_2DIFF |  |  |
| GORILLA |  |  |

注：红色表示该参数在目前的机器配置、IoTDB版本下无法完成测试

### 3.4.2 IoTDB系统性能指标与写入速率

我们以IoTDB在如下表所示的参数设置下进行的写入测试为例，绘制了写入过程中每个batch的瞬时写入速率与服务器的CPU利用率、磁盘IO利用率、内存利用率以及网口接收速率随时间变化的图像，如图 1所示。

|  |  |
| --- | --- |
| 参数名 | 值 |
| DEVICE\_NUMBER | 100 |
| CACHE\_NUM | 1000 |
| SENSOR\_NUMBER | 100 |
| GROUP\_NUMBER | 10 |
| LOOP | 100 |
| CLIENT\_NUMBER | 10 |
| ENCOIDNG | PLAIN |

图 1 IoTDB瞬时写入速率与系统信息指标

图 2 IoTDB每个客户端线程的写入速率

### 3.4.3 InfluxDB系统性能指标与写入速率

我们以InfluxDB在如下表所示的参数设置下进行的写入测试为例，绘制了写入过程中每个batch的瞬时写入速率与服务器的CPU利用率、磁盘IO利用率、内存利用率以及网口接收速率随时间变化的图像，如图 3所示。

|  |  |
| --- | --- |
| 参数名 | 值 |
| DEVICE\_NUMBER | 100 |
| CACHE\_NUM | 1000 |
| SENSOR\_NUMBER | 100 |
| GROUP\_NUMBER | 10 |
| LOOP | 100 |
| CLIENT\_NUMBER | 10 |

图 3 InfluxDB瞬时写入速率与系统信息指标



图 4 InfluxDB每个客户端线程的写入速率

## 3.5实验结论

1. 数据规模变化实验中得到的结论
   1. 待填
   2. 待填
2. 批量导入数量变化实验中得到的结论
   1. 待填
   2. 待填
3. 传感器数量变化实验中得到的结论
   1. 待填
4. 总设备数变化实验中得到的结论
   1. 待填
5. 存储组数变化实验中得到的结论
   1. 待填
6. 并发数变化实验中得到的结论
   1. 待填

# 4 范围查询性能测试

## 4.1业务背景

假定上层应用需要从iotDB系统中查询某个设备的某些传感器在某个时间段内的数据进行展示与分析。比如，查询某台设备在某一天内的多个传感器信息。

## 4.2 实验目的

测试IoTDB范围查询目前每秒钟可以得到的结果集数据点数与系统稳定性，验证系统范围查询性能能否满足青海项目现在的数据范围查询需求。探究结果集点数与范围查询效果（用每秒查询到的点数）的关系。根据测试实验结果了解不同参数对系统性能的影响，综合评价系统。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
|  | 有无写入背景（或者查询插入占比）目前不支持 |
| QUERY\_INTERVAL | 范围查询的时间间隔（范围的大小） |
| CLIENT\_NUMBER | 客户端线程数（并发度） |
| QUERY\_SENSOR\_NUM | 一条范围查询语句包含的传感器数目 |
| QUERY\_DEVICE\_NUM | 一条范围查询语句包含的设备数目 |

备注：encoding部分

## 4.3 实验设计

#### 4.3.1 度量指标

1. 性能指标

1) 结果集数据点数 : 查询结果所包含的数据点数。

2) 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。

3)平均查询速率：结果集数据点数/总时间。

4)平均查询性能：总执行查询数/总时间。

2. 系统信息指标

1. 客户机及服务器的磁盘IO使用率。
2. 客户机及服务器的CPU使用率。
3. 客户机及服务器的内存使用率。
4. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 4.3.2 实验数据

实验采用iotdb-benchmark工具自动生成的随机数据。数据类型包括整型和浮点型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。范围查询测试采用插入测试阶段已经插入的数据，参数值如下表所示：（大约3.9G）

|  |  |
| --- | --- |
| 参数名 | 值 |
| 存储组数 | 10 |
| 总设备数 | 100 |
| 每个设备传感器数 | 100 |
| 数据产生间隔/s | 5 |
| 总点数 | 10000\*50000 |

#### 备注1：其中每一个root.performf.grop\_x文件夹下均只有一个tsfile文件

备注2：数据集相当于100个设备，每个设备有100个传感器，在数据产生间隔为5s的情况下，经过69.444444h产生的数据量

#### 4.3.3场景设计

针对可能影响范围查询所需时间的各种因素，设计了各种测试场景用于彼此进行对比试验。

**测试场景1：有写入背景条件下单设备多传感器的性能测试**

有背景写入的条件下，用户查询某台设备的多个传感器在某段时间内的数据。

**测试场景2：无写入背景条件下查询单设备多传感器的性能测试**

无背景写入条件下，用户查询某台设备的某些传感器在某段时间内的数据，测试场景3用来与测试场景1对比，来探究背景写入的有无对于范围查询所需时间的影响。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **测试场景编号** | **查询设备数目** | **查询传感器数目** | **客户端数目** | **时间长度** | **背景写入** | **查询速度（points/s）** |
| 1 | 1 | 1 | 4 | 0.5h | 有 |  |
| 2 | 1 | 实验变量 | 实验变量 | 实验变量 | 无 |  |

#### 4.3.4测试步骤：

**第一步，每个客户端并发地生成查询语句并发送给服务器：**

启动CLIENT\_NUMBER个线程，每个线程执行LOOP次查询。其中每次执行查询的QUERY\_DEVICE\_NUM个设备以及QUERY\_SENSOR\_NUM个传感器是随机生成的，生成的查询语句中所有设备要查询的传感器是相同的。第i次查询的开始时间生成规则如下：数据库中所有数据的开始时间 + i \* startTimeInterval；其中startTimeInterval =（数据库中所有数据的时间跨度）/LOOP ；i为0,1,2,...LOOP-1;其中每次查询的开始时间和结束时间的时间间隔有由配置参数决定。

**第二步，监控服务器系统状态：**

在IoTDB性能测试进行的过程中，利用iotdb-benchmark监控程序记录系统信息，并根据测试实验执行的长短来确定每隔多久获取一次服务器机器的CPU使用率，网络传输，硬盘IO和内存的性能状态，用于服务器性能及系统瓶颈的分析。

**第三步，统计查询性能指标：**

每个客户端线程均有2个的成员变量，分别存储执行查询语句所花费的总时间以及返回查询结果的总点数。当所有线程都完成查询后取总时间最大的值作为整个数据集的查询总时间，所有线程查询结果点数之和作为总查询点数。平均查询速率即总查询点数除以总时间。所有线程一共处理的查询语句数目等于LOOP\*CLIENT\_NUMBER，所以处理的查询语句的速度为LOOP\*CLIENT\_NUMBER/总时间；

伪代码：

**while** (i < config.LOOP) {

Collections.*shuffle*(clientDevicesIndex);

**for** (**int** m = 0; m < config.QUERY\_DIVICE\_NUM; m++){

queryDevicesIndex.add(clientDevicesIndex.get(m));

}

database.executeOneQuery(queryDevicesIndex,

i, startTimeInterval \* i + Constants.***START\_TIMESTAMP***, **this**, *errorCount*);

i++;

queryDevicesIndex.clear();

}

备注1：在executeOneQuery中调用生成sql语句的函数，在该函数中生成乱序个config.QUERY\_SENSOR\_NUM个传感器。

备注2：当查询传感器数目与查询设备数目均为1，时间间隔为0.5h时，查询语句为： SELECT s\_30 FROM root.performf.group\_4.d\_43 WHERE time > 2010-01-01 12:00:00 AND time < 2010-01-01 12:30:00

## 4.4实验结果

#### 4.4.1 不同参数变化时InfluxDB与IoTDB的性能对比

1. 测试场景1——探究不同写入背景对性能影响（暂时未做）

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 2 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| QUERY\_INTERVAL | 20\*1000ms |
| LOOP | 3000 |
| IOTDB\_VERSION | 0.3.0 |

I

1. 测试场景2——探究无写入背景不同设备数对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| 写入背景 | 无 |
| CLIENT\_NUMBER | 4 |
| QUERY\_INTERVAL | 0.5h |
| LOOP | 5000 |

服务器：10号机器（192.168.130.23）

客户端：21号机器

IOTDB版本：0.3.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DIVICE\_NUM（查询语句中设备个数） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB范围查询不同设备数实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DIVICE\_NUM（查询语句中设备个数） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB范围查询不同设备数实验结果

1. 测试场景2——探究在无写入背景的情况下QUERY\_INTERVAL（时间间隔）对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| 写入背景 | 无 |
| CLIENT\_NUMBER | 4 |
| QUERY\_DEVICE\_NUM | 1 |
| LOOP | 5000 |

服务器：10号机器（192.168.130.23）

客户端：21号机器

IOTDB版本：0.3.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | 时间间隔 |
|  |  |  |  |  | 20000 |
|  |  |  |  |  | 60000 |
|  |  |  |  |  | 1800000 |
|  |  |  |  |  | 3600000 |
|  |  |  |  |  | 7200000 |

IoTDB范围查询不同时间间隔实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | 时间间隔 |
|  |  |  |  |  | 20000 |
|  |  |  |  |  | 60000 |
|  |  |  |  |  | 1800000 |
|  |  |  |  |  | 3600000 |
|  |  |  |  |  | 7200000 |

InfluxDB范围查询不同时间间隔实验结果

1. 测试场景2——探究在无写入背景的情况下CLIENT\_NUMBER（客户端个数）对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| 写入背景 | 无 |
| QUERY\_INTERVAL | 0.5h |
| QUERY\_DEVICE\_NUM | 1 |
| LOOP | 5000 |

服务器：10号机器（192.168.130.23）

客户端：21号机器

IOTDB版本：0.3

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | clientNum（客户端数目） | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s |
|  |  |  | 1 |  |  |
|  |  |  | 2 |  |  |
|  |  |  | 4 |  |  |
|  |  |  | 6 |  |  |
|  |  |  | 8 |  |  |

IoTDB范围查询不同客户端个数实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | clientNum（客户端数目） | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s |
|  |  |  | 1 |  |  |
|  |  |  | 2 |  |  |
|  |  |  | 4 |  |  |
|  |  |  | 6 |  |  |
|  |  |  | 8 |  |  |

InflxDB范围查询不同客户端个数实验结果

1. 测试场景2——探究在无写入背景的情况下QUERY\_SENSOR\_NUM对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_INTERVAL | 0.5h |
| 写入背景 | 无 |
| CLIENT\_NUMBER | 4 |
| QUERY\_DEVICE\_NUM | 1 |
| LOOP | 5000 |

服务器：10号机器（192.168.130.23）

客户端：21号机器

IOTDB版本：0.3.1

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_SENSOR\_NUM（查询语句中传感器数目） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB范围查询不同传感器个数实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_SENSOR\_NUM（查询语句中传感器数目） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB范围查询不同传感器个数实验结果

### 4.2 网络IO使用量：

### 4.4.3 CPU占用率：

### 4.4.4磁盘IO：

## 4.5实验结论

场景1和场景2的对应实验结果来探究写入背景影响；

场景2内部探究QUERY\_INTERVAL、CLIENT\_NUMBER、QUERY\_SENSOR\_NUM、LOOP对速率的影响；

**结论1：**根据实验B、E，在无写入背景情况下，单设备多传感器的查询时间与查询传感器数目成线性关系，多设备单传感器的查询时间与查询传感器数目成线性关系。（查询时间与查询的传感器是否跨存储组无关，所需时间与查询的timeseries数成正比）

**结论2：**根据实验D，在无写入背景情况下，单设备单传感器的查询时间并没有与查询线程数目呈线性关系，而是逐渐上升至一个稳定值（大概当有2个查询线程时查询速率就不在增加）

**结论3：**根据实验C，在无写入背景情况下，单设备单传感器的查询时间几乎不随查询的时间间隔变化而变化。

# 5 条件查询性能测试

## 5.1业务背景

上层应用需要从iotDB系统中查询某些设备的某些传感器在某个时间段内满足某些过滤条件的数据进行展示与分析。比如，查询某几台车在某一天内的发动机转速大于50的数据用于分析。

## 5.2 实验目的

测试IoTDB条件查询目前每秒钟可以得到的结果集数据点数与系统稳定性，验证系统范围查询性能能否满足青海项目现在的数据范围查询需求。根据测试实验结果了解不同参数对系统性能的影响，综合评价系统。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
|  | 有无写入背景（或者查询插入占比）目前不支持 |
| QUERY\_INTERVAL | 条件查询的时间间隔（范围的大小） |
| CLIENT\_NUMBER | 客户端线程数（并发度），给每个客户端均匀分配Device |
| QUERY\_SENSOR\_NUM | 一条查询语句包含的传感器数目 |
| QUERY\_DEVICE\_NUM | 一条查询语句包含的设备数目 |
| QUERY\_LOWER\_LIMIT | 不同的过滤条件对于条件查询所需时间的影响。 |

## 5.3 实验设计

#### 5.3.1 度量指标

1. 性能指标

1) 结果数据点数 : 查询结果所包含的数据点数。

2) 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。

3)范围数据点数：在该查询所对应的时间范围内的数据点数

4)平均查询速率：结果数据点数/总时间

5)平均处理范围数据点速率：范围数据点数/总时间

6)平均处理查询语句速率：总查询数/总时间

2. 系统信息指标

1. 客户机及服务器的磁盘IO使用率。
2. 客户机及服务器的CPU使用率。
3. 客户机及服务器的内存使用率。
4. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 5.3.2 实验数据

实验采用iotdb-benchmark工具自动生成的随机数据。数据类型包括整型和浮点型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。范围查询测试采用插入测试阶段已经插入的数据，参数值如下表所示：（大约3.9G）

|  |  |
| --- | --- |
| 参数名 | 值 |
| 存储组数 | 10 |
| 总设备数 | 100 |
| 每个设备传感器数 | 100 |
| 数据产生间隔/s | 5 |
| 总点数 | 10000\*50000 |

#### 备注1：其中每一个root.performf.grop\_x文件夹下均只有一个tsfile文件

备注2：数据集相当于100个设备，每个设备有100个传感器，在数据产生间隔为5s的情况下，经过69.444444h产生的数据量

当查询传感器数目与查询设备数目均为1，时间间隔为0.5h时，查询语句为：SELECT s\_39 FROM root.performf.group\_2.d\_29 WHERE time > 2010-01-01 12:00:00 AND time < 2010-01-01 12:30:00 AND root.performf.group\_2.d\_29.s\_39 > 0.0

#### 5.3.3场景设计

针对可能影响条件查询所需时间的返回结果数据点数，背景写入，过滤条件等因素设计多组测试场景彼此对照。

**测试场景1：有写入背景单设备多传感器大于X的查询（暂未实现）**

有背景写入的条件下，用户查询某台设备的多个传感器在某段时间内的满足传感器数值大于X的数据。

**测试场景2：无写入背景多设备单传感器大于0的查询**

无背景写入的条件下，用户查询某些设备的某个传感器在某段时间内的数据大于0的数据。

测试场景2是为了探究在无背景写入有过滤条件时，条件查询所需时间与返回结果数据量的关系。

**测试场景3：无写入背景单设备多传感器大于0的查询**

无背景写入条件下，用户查询某台设备的某些传感器在某段时间内的满足传感器数值大于0的数据，测试场景3用来与测试场景1对比，来探究背景写入的有无对于条件查询所需时间的影响。

**测试场景4：探究无写入背景单设备单传感器大于0的查询效果与客户端个数的关系**

验证在查询客户端个数变化时，查询效率是否会线性增加。

**测试场景5：无写入背景多设备单传感器大于-10的查询**

无背景写入的条件下，用户查询某些设备的某个传感器在某段时间内的数据大于等于0的数据。

测试场景2用来与测试场景1进行对比，来探究同样的查询范围下，不同的过滤条件对于条件

**测试场景6：探究无写入背景单设备单传感器大于0的查询效果与条件查询时间范围的关系**

探究查询效果与查询的时间段的大小的关系。

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **测试场景编号** | **查询设备数目** | **查询传感器数目** | **客户端数目** | **时间长度** | **背景写入** | **过滤条件** |
| 1 | 1 | 变量 | 4 | 0.5h | 有 | value>0 |
| 2 | 变量 | 1 | 4 | 0.5h | 无 | value>0 |
| 3 | 1 | 变量 | 4 | 0.5h | 无 | value>0 |
| 4 | 1 | 1 | 变量 | 0.5h | 无 | value>0 |
| 5 | 1 | 变量 | 4 | 0.5h | 无 | value>-10 |
| 6 | 1 | 1 | 4 | 变量 | 有 | value>0 |

A)测试场景1——探究不同写入背景对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | XX |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | XX |
| QUERY\_INTERVAL | XX |
| LOOP | XX |

平均查询速率Mean Throughput （points／s）

|  |  |  |
| --- | --- | --- |
| 读写占比 | InfluxDB | IoTDB |
| 20% |  |  |
| 40% |  |  |
| 60% |  |  |
| 80% |  |  |
| 100% |  |  |

B)测试场景2——探究不同设备数对无写入背景多设备单传感器大于0的查询性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| 写入背景 | 无 |
| CLIENT\_NUMBER | 4 |
| QUERY\_INTERVAL | 0.5h |
| LOOP | 5000 |
| QUERY\_LOWER\_LIMIT | 0（过滤条件values>0） |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DIVICE\_NUM（查询语句中设备个数） | errorNum（出错的查询个数） |
|  |  |  |  |  | 1 |  |
|  |  |  |  |  | 5 |  |
|  |  |  |  |  | 10 |  |
|  |  |  |  |  | 20 |  |
|  |  |  |  |  | 50 |  |

红色为查询过程中有出错的查询(这可能是由于jdbc版本太低，会在jdbc0.3.1版本进行重复试验)

IotDB条件查询不同设备个数实验结果

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DIVICE\_NUM（查询语句中设备个数） | errorNum（出错的查询个数） |
| 20000 |  |  |  |  | 1 | 0 |
| 20000 |  |  |  |  | 5 | 0 |
| 20000 |  |  |  |  | 10 | 0 |
| 20000 |  |  |  |  | 20 | 0 |
| 20000 |  |  |  |  | 50 | 0 |

InfluxDB条件查询不同设备个数实验结果

C）测试场景3——探究在无写入背景的情况下QUERY\_SENSOR\_NUM对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| QUERY\_INTERVAL | 0.5h |
| 写入背景 | 无 |
| QUERY\_LOWER\_LIMIT | 0（过滤条件values>0） |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | errorNum（出错的查询个数） | QUERY\_SENSOR\_NUM |
|  |  |  |  |  |  | 1 |
|  |  |  |  |  |  | 5 |
|  |  |  |  |  |  | 10 |
|  |  |  |  |  |  | 20 |
|  |  |  |  |  |  | 50 |

红色为查询过程中有出错的查询(这可能是由于jdbc版本太低，会在jdbc0.3.1版本进行重复试验)

IotDB条件查询单设备不同传感器个数实验结果

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | errorNum（出错的查询个数） | QUERY\_SENSOR\_NUM | |
|  |  |  |  |  | 0 | 1 |
|  |  |  |  |  | 0 | 5 |
|  |  |  |  |  | 0 | 10 |
|  |  |  |  |  | 0 | 20 |
|  |  |  |  |  | 0 | 50 |

InfluxDB条件查询单设备不同传感器个数实验结果

D)测试场景4——探究在无写入背景的情况下CLIENT\_NUMBER（客户端个数）对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| QUERY\_DEVICE\_NUMBER | 1 |
| QUERY\_INTERVAL | 0.5h |
| LOOP | 5000 |
| 写入背景 | 无 |
| QUERY\_LOWER\_LIMIT | 0（过滤条件values>0） |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | clientNum | （每秒执行的查询数目）query/s | rate |
|  |  |  | 1 |  |  |
|  |  |  | 2 |  |  |
|  |  |  | 4 |  |  |
|  |  |  | 6 |  |  |
|  |  |  | 8 |  |  |

IotDB条件查询单设备单传感器不同客户端个数实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | clientNum | （每秒执行的查询数目）query/s | rate |
|  |  |  | 1 |  |  |
|  |  |  | 2 |  |  |
|  |  |  | 4 |  |  |
|  |  |  | 6 |  |  |
|  |  |  | 8 |  |  |

InfluxDB条件查询单设备单传感器不同客户端个数实验结果

E)测试场景5——探究在无写入背景的情况下QUERY\_INTERVAL（时间间隔）对性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| 写入背景 | 无 |
| CLIENT\_NUMBER | 4 |
| QUERY\_DIVICE\_NUM | 1 |
| LOOP | 5000 |
| QUERY\_LOWER\_LIMIT | 0（过滤条件values>0） |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_INTERVAL |
|  |  |  |  |  | 20000 |
|  |  |  |  |  | 60000 |
|  |  |  |  |  | 1800000 |
|  |  |  |  |  | 3600000 |
|  |  |  |  |  | 7200000 |

IotDB条件查询单设备单传感器不同时间间隔实验结果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_INTERVAL |
|  |  |  |  | 20000 |
|  |  |  |  | 60000 |
|  |  |  |  | 1800000 |
|  |  |  |  | 3600000 |
|  |  |  |  | 7200000 |

InfluxDB条件查询单设备单传感器不同时间间隔实验结果

F)测试场景6——探究不同设备数对无写入背景多设备单传感器大于-10的查询性能影响

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| QUERY\_SENSOR\_NUM | 1 |
| 写入背景 | 无 |
| CLIENT\_NUMBER | 4 |
| QUERY\_INTERVAL | 0.5h |
| LOOP | 5000 |
| QUERY\_LOWER\_LIMIT | -10（过滤条件values>-10） |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | errorNum（出错的查询个数） | QUERY\_DIVICE\_NUM |
|  |  |  |  |  |  | 1 |
|  |  |  |  |  |  | 5 |
|  |  |  |  |  |  | 10 |
|  |  |  |  |  |  | 20 |
|  |  |  |  |  |  | 50 |

IotDB条件查询多设备单传感器values>-10实验结果（在iotdb0.3下实验）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_DIVICE\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB条件查询多设备单传感器values>-10实验结果

#### 5.3.4测试步骤：

**每个客户端并发地生成查询语句并发送给服务器：**

每个客户端线程均有2个的成员变量，分别存储执行查询语句所花费的总时间以及返回查询结果的总点数。当所有线程都完成查询后取总时间最大的值作为整个数据集的查询总时间，所有线程查询结果点数之和作为总查询点数。平均查询速率即总查询点数除以总时间。所有线程一共处理的查询语句数目等于LOOP\*CLIENT\_NUMBER，所以处理的查询语句的速度为LOOP\*CLIENT\_NUMBER/总时间；

伪代码：

**while** (i < config.LOOP) {

Collections.*shuffle*(clientDevicesIndex);

**for** (**int** m = 0; m < config.QUERY\_DIVICE\_NUM; m++){

queryDevicesIndex.add(clientDevicesIndex.get(m));

}

database.executeOneQuery(queryDevicesIndex,

i, startTimeInterval \* i + Constants.***START\_TIMESTAMP***, **this**, *errorCount*);

i++;

queryDevicesIndex.clear();

}

备注1：在executeOneQuery中调用生成sql语句的函数，在该函数中生成乱序个config.QUERY\_SENSOR\_NUM个传感器。

## 5.5实验结论

**结论1：**根据实验2和3，无写入背景多设备单传感器大于0的查询性能与设备数成线性关系，多设备单传感器的查询时间与查询传感器数目成线性关系。（查询时间与查询的传感器是否跨存储组无关，所需时间与查询的timeseries数成正比）但是在多设备或者多传感器下均有些查询会抛异常，无法正常执行。

**结论2：**根据实验4，在无写入背景情况下，单设备单传感器的查询时间并没有与查询线程数目呈线性关系，而是逐渐上升至一个稳定值（大概当有2个查询线程时查询速率就不在增加）

**结论3：**根据实验5，在无写入背景情况下，单设备单传感器的查询时间几乎不随查询的时间间隔变化而变化。

# 6 精确点查询性能测试

## 6.1业务背景

业务需要精确点查询来查看设备的某些传感器在某个时间点发送的数据。如查看某台设备在去年同期的传感器数据。

## 6.2 实验目的

测试IoTDB精确点查询目前每秒钟可以得到的结果集数据点数与系统稳定性，验证系统范围查询性能能否满足青海项目现在的数据范围查询需求。探究结果集点数与精确点查询效果的关系以及在查询结果是否为空情况下的性能对比？。探究每次精确点查询时间分布情况。根据测试实验结果了解不同参数对系统性能的影响，综合评价系统。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
|  | 有无写入背景（或者查询插入占比）目前不支持 |
| QUERY\_SENSOR\_NUM | 一条查询语句包含的传感器数目 |
| IS\_EMPTY\_PRECISE\_POINT\_QUERY | 精确点查询结果是否为空 |
| QUERY\_DEVICE\_NUM | 一条查询语句包含的设备数目，探究不同存储组 |

## 6.3 实验设计

#### 6.3.1 度量指标

1. 性能指标

1) 查询数据点数 : 查询所包含的数据点数。

2) 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。

4)平均查询速率：查询语句条数/总时间

2. 系统信息指标

1. 客户机及服务器的磁盘IO使用率。
2. 客户机及服务器的CPU使用率。
3. 客户机及服务器的内存使用率。
4. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 6.3.2 实验数据

实验采用iotdb-benchmark工具自动生成的随机数据。数据类型包括整型和浮点型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。范围查询测试采用插入测试阶段已经插入的数据，参数值如下表所示：（大约3.9G）

|  |  |
| --- | --- |
| 参数名 | 值 |
| 存储组数 | 10 |
| 总设备数 | 100 |
| 每个设备传感器数 | 100 |
| 数据产生间隔/s | 5 |
| 总点数 | 10000\*50000 |

#### 备注1：其中每一个root.performf.grop\_x文件夹下均只有一个tsfile文件

备注2：数据集相当于100个设备，每个设备有100个传感器，在数据产生间隔为5s的情况下，经过69.444444h产生的数据量

当查询传感器数目与查询设备数目均为1时，查询语句为： SELECT s\_57 FROM root.performf.group\_4.d\_49 WHERE time = 2010-01-01 12:00:00

#### 6.3.3场景设计

针对可能影响精确点查询所需时间与查询客户端个数、查询传感器个数、背景写入、是否为空值查询以及数据库中现有数据的大小等因素设计多组测试场景彼此对照。

**测试场景1：有写入背景条件下查询成功的性能测试**

在有背景写入的条件下，用户查询某台设备的多种传感器在某个时间点的数据，且查询结果均不为空。本测试场景用来探究有背景写入，查询不为空的情况下传感器数目对于精确点查询所需时间的影响。

* 背景写入设置：X个线程，批处理大小为X的背景写入程序，写入速率大约为X点每秒。
* 查询不为空数据保证：按照与生成插入数据相同的方式生成有效的查询数据，保证生成的查询数据存在于数据库中。

**测试场景2：无写入背景条件下单设备多传感器查询成功的性能测试**

在无背景写入的条件下，用户查询某台设备的多种传感器在某个时间点的数据，且查询结果均不为空。本测试场景可以与测试场景1进行对比，探究背景写入的有无对于查询不为空的精确点查询所需时间的影响。所传入的时间戳需要保证对应的设备在相应的时间点存在数据点。

**测试场景3：无写入背景条件下查询不成功的性能测试**

在无背景写入的条件下，用户查询某台设备的多种传感器在某个时间点的数据，但所查询时间点处不存在数据。本测试场景可以与测试场景2进行对比，探究无背景写入情况下，查询结果是否存在对于精确点查询所需时间的影响。所传入的时间戳需要保证对应的设备在相应的时间点不存在数据点。

* 查询为空数据保证：所生成的时间戳需要保证对应的设备在相应的时间点存在数据点，（目前保证生成的时间戳在最大与最小时间之间且不存在）。

**测试场景4：无写入背景条件下多设备单传感器查询成功的性能测试**

在无背景写入的条件下，用户查询某些设备的一种传感器在某个时间点的数据，且查询结果均不为空。本测试场景可以与测试场景2进行对比，探究背景查询timeseries是否跨存储组对于查询不为空的精确点查询所需时间的影响。所传入的时间戳需要保证对应的设备在相应的时间点存在数据点。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **测试场景编号** | **查询设备数目** | **查询传感器数目** | **背景写入** | **客户端数目** | **查询为空** |
| 1 | 1 | 变量 | 有 | 4 | 否 |
| 2 | 1 | 变量 | 无 | 4 | 否 |
| 3 | 1 | 变量 | 无 | 4 | 是 |
| 4 | 变量 | 1 | 无 | 4 | 否 |

**A）测试场景1--**

**B）测试场景2--无写入背景条件下查询成功的性能测试--单设备多传感器**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 查询结果是否为空 | 否 |
| 写入背景 | 无 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB精确点查询单设备多传感器查询成功实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB精确点查询单设备多传感器查询成功实验结果

**C）测试场景3--无写入背景条件下查询不成功的性能测试--单设备多传感器**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 查询结果是否为空 | 是 |
| 写入背景 | 无 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_SENSOR\_NUM（查询语句中传感器数目） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB精确点查询单设备多传感器查询不成功实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_SENSOR\_NUM（查询语句中传感器数目） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB精确点查询单设备多传感器查询不成功实验结果

**D）测试场景4--无写入背景条件下查询成功的性能测试--多设备单传感器**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_SENSOR\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 查询结果是否为空 | 否 |
| 写入背景 | 无 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DEVICE\_NUM（查询语句中设备个数） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB精确点查询多设备单传感器查询成功实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DEVICE\_NUM（查询语句中设备个数） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB精确点查询多设备单传感器查询成功实验结果

**6.3.4每个客户端并发地生成查询语句并发送给服务器：**

每个客户端线程均有2个的成员变量，分别存储执行查询语句所花费的总时间以及返回查询结果的总点数。当所有线程都完成查询后取总时间最大的值作为整个数据集的查询总时间，所有线程查询结果点数之和作为总查询点数。平均查询速率即总查询点数除以总时间。所有线程一共处理的查询语句数目等于LOOP\*CLIENT\_NUMBER，所以处理的查询语句的速度为LOOP\*CLIENT\_NUMBER/总时间；

伪代码：

**while** (i < config.LOOP) {

Collections.*shuffle*(clientDevicesIndex);

**for** (**int** m = 0; m < config.QUERY\_DIVICE\_NUM; m++){

queryDevicesIndex.add(clientDevicesIndex.get(m));

}

database.executeOneQuery(queryDevicesIndex,

i, startTimeInterval \* i + Constants.***START\_TIMESTAMP***, **this**, *errorCount*);

i++;

queryDevicesIndex.clear();

}

备注1：在executeOneQuery中调用生成sql语句的函数，在该函数中生成乱序个config.QUERY\_SENSOR\_NUM个传感器。

## 6.5实验结论

**结论1：**根据实验2，在无写入背景情况下，单设备多传感器的查询时间与查询传感器数目成线性关系。（所需时间与结果集点数成正比）

**结论2：**根据实验3，在无写入背景情况下，在无写入背景情况下，单设备多传感器的查询时间与查询传感器数目成线性关系。并且根据与实验二的对比，发现精确点查询查询结果是否为空对查询效率影响不大

**结论3：**根据实验4，多设备单传感器的查询时间与查询传感器数目成线性关系。（所需时间与结果集点数成正比）

# 7 最近点查询性能测试

## 7.1业务背景

业务需要最近点查询来查看设备的某些传感器最新的数据。如查看房间现在的空调度数。

## 7.2 实验目的

测试IoTDB最近点查询在不同参数下的平均速度。根据测试实验结果了解不同参数对系统性能的影响，综合评价系统。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
|  | 有无写入背景（或者查询插入占比）目前不支持 |
| QUERY\_SENSOR\_NUM | 一条查询语句包含的传感器数目 |
| QUERY\_DEVICE\_NUM | 一条查询语句包含的设备数目，探究不同存储组 |

## 7.3 实验设计

#### 7.3.1 度量指标

1. 性能指标

1) 待查询数据点数 : 查询传感器数目\*查询设备数

2) 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。

3)平均查询速率：待查询数据点数/总时间

2. 系统信息指标

1. 客户机及服务器的磁盘IO使用率。
2. 客户机及服务器的CPU使用率。
3. 客户机及服务器的内存使用率。
4. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 7.3.2 实验数据

实验采用iotdb-benchmark工具自动生成的随机数据。数据类型包括整型和浮点型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。范围查询测试采用插入测试阶段已经插入的数据，参数值如下表所示：（大约3.9G）

|  |  |
| --- | --- |
| 参数名 | 值 |
| 存储组数 | 10 |
| 总设备数 | 100 |
| 每个设备传感器数 | 100 |
| 数据产生间隔/s | 5 |
| 总点数 | 10000\*50000 |

#### 备注1：其中每一个root.performf.grop\_x文件夹下均只有一个tsfile文件

备注2：数据集相当于100个设备，每个设备有100个传感器，在数据产生间隔为5s的情况下，经过69.444444h产生的数据量

当查询传感器数目与查询设备数目均为1时，查询语句为： SELECT max\_time(s\_76) FROM root.performf.group\_3.d\_31

#### 7.3.3场景设计

针对可能影响精确点查询所需时间查询客户端个数、查询传感器个数、背景写入、是否为空值查询以及数据库中现有数据的大小等因素设计多组测试场景彼此对照。

**测试场景1：有写入背景条件下单设备多传感器的性能测试**

有背景写入条件下，用户查询某台设备的多个传感器的最新数据。测试场景1用来探究有背景写入情况下传感器数目对于最近点查询所需时间的影响。

**测试场景2：无写入背景条件下多设备单传感器的性能测试**

有背景写入条件下，用户查询多台设备的某个传感器的最新数据，测试场景2用来探究无背景写入情况下设备数目对于最近点查询所需时间的影响。

**测试场景3：无写入背景条件下单设备多传感器的性能测试**

无背景写入条件下，用户查询某台设备的多个传感器的最新数据。测试场景3用来与测试场景1进行对比，探究背景写入的有无对于最近点查询所需时间的影响。

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **测试场景编号** | **查询设备数目** | **查询传感器数目** | **背景写入** | **客户端数目** |
| 1 | 1 | 变量 | 有 | 4 |
| 2 | 变量 | 1 | 无 | 4 |
| 3 | 1 | 变量 | 无 | 4 |

**A）测试场景1--有写入背景条件下单设备多传感器的性能测试**

**B）测试场景2--无写入背景条件下多设备单传感器的性能测试**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_SENSOR\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 写入背景 | 无 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DEVICE\_NUM（查询语句中设备个数） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB最近点查询多设备单传感器实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_DEVICE\_NUM（查询语句中设备个数） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB最近点查询多设备单传感器实验结果

**C）测试场景3--无写入背景条件下单设备多传感器的性能测试**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 写入背景 | 无 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_SENSOR\_NUM（查询语句中传感器数目） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB最近点查询单设备多传感器实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum（所有线程一共进行的查询总数） | point（结果集点数） | time（最后一个查询线程结束时间）s | （每秒执行的查询数目）query/s | （每秒返回的结果集点数）point/s | QUERY\_SENSOR\_NUM（查询语句中传感器数目） |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB最近点查询单设备多传感器实验结果

## 7.5实验结论

**结论1：**根据实验2，在无写入背景情况下多设备单传感器的查询时间在查询的传感器数目较少时随着查询设备数目（timeseries）增大而增大，而后逐渐稳定（无明显线性关系）（发现在查询timeseries数相同时多设备单传感器的查询时间要大于单设备多传感器的时间）

**结论2：**根据实验3，在无写入背景情况下，在无写入背景情况下，单设备多传感器的查询时间在查询的传感器数目较少时随着传感器数目增大而增大，而后逐渐稳定（无明显线性关系）

# 8 聚合查询性能测试

## 8.1业务背景

业务需要聚合查询来查看设备的某些统计信息。如查看一段时间内的cpu平均利用率等。

## 8.2 实验目的

测试IoTDB聚合查询在不同参数下的不同聚合函数的平均速度。根据测试实验结果了解不同参数对系统性能的影响，综合评价系统。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
|  | 有无写入背景（或者查询插入占比）目前不支持 |
| CLIENT\_NUMBER | 客户端线程数（并发度），给每个客户端均匀分配Device |
| QUERY\_SENSOR\_NUM | 一条查询语句包含的传感器数目 |
| QUERY\_DEVICE\_NUM | 一条查询语句包含的设备数目，探究不同存储组 |
| QUERY\_AGGREGATE\_FUN | 待测试的聚合函数种类 |

## 8.3 实验设计

#### 8.3.1 度量指标

1. 性能指标

1) 查询数据点数 : 查询所包含的数据点数。

2) 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。

4)平均查询速率：查询数据点数/总时间

2. 系统信息指标

1. 客户机及服务器的磁盘IO使用率。
2. 客户机及服务器的CPU使用率。
3. 客户机及服务器的内存使用率。
4. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 8.3.2 实验数据

实验采用iotdb-benchmark工具自动生成的随机数据。数据类型包括整型和浮点型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。范围查询测试采用插入测试阶段已经插入的数据，参数值如下表所示：（大约3.9G）

|  |  |
| --- | --- |
| 参数名 | 值 |
| 存储组数 | 10 |
| 总设备数 | 100 |
| 每个设备传感器数 | 100 |
| 数据产生间隔/s | 5 |
| 总点数 | 10000\*50000 |

#### 备注1：其中每一个root.performf.grop\_x文件夹下均只有一个tsfile文件

备注2：数据集相当于100个设备，每个设备有100个传感器，在数据产生间隔为5s的情况下，经过69.444444h产生的数据量

当查询传感器数目与查询设备数目均为1且时间间隔为0.5h时，查询语句为： SELECT max\_value(s\_76) FROM root.performf.group\_3.d\_31 WHERE time > 2010-01-01 12:00:00 AND time < 2010-01-01 12:30:00

#### 8.3.3场景设计

针对可能影响聚合查询所需时间的返回结果数据点数，背景写入，过滤条件，聚合类型，客户端个数等因素设计多组测试场景彼此对照。

**测试场景1：无背景写入单设备多传感器的COUNT聚合**

在无背景写入的条件下，用户查询某台设备的某些传感器在某段时间内的COUNT结果。

**测试场景2：无背景写入多设备单传感器的COUNT聚合**

在无背景写入的条件下，用户查询某些设备的某中传感器在某段时间内的COUNT结果。

测试场景1,2均是为了探究无背景写入条件下，返回结果数据点数与聚合查询所需时间的关系。测试场景1,2的结果将统一作图。

**测试场景3：无背景写入单设备单传感器的MAX\_VALUE不同时间范围的聚合**

在无背景写入的条件下，用户查询某些设备的某中传感器在某段时间内满足过滤条件>=0的数据的COUNT结果。用来与测试场景1对照，探究背景写入的有无对于聚合查询所需时间的影响。

**测试场景4：无背景写入单设备多传感器的MAX\_VALUE聚合**

无背景写入条件下，用户查询某些设备的某中传感器在某段时间内满足过滤条件>=50的数据的COUNT结果。用来与测试场景1对照，探究过滤条件对于聚合查询所需时间的影响。

**测试场景5~8：无背景写入单设备多传感器的不同种类的聚合操作**

在无背景写入的条件下，用户查询某台设备的某些传感器在某段时间内的MIN\_TIME,MIN\_VALUE,SUM,AVG结果。用来探究不同聚合类型对聚合查询所需时间的影响。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **测试场景编号** | **查询设备数目** | **查询传感器数目** | **时间长度** | **背景写入** | **聚合操作** |
| 1 | 1 | 实验变量 | 0.5h | 无 | COUNT |
| 2 | 实验变量 | 1 | 0.5h | 无 | COUNT |
| 3 | 1 | 1 | 实验变量 | 无 | MAX\_VALUE |
| 4 | 1 | 实验变量 | 0.5h | 无 | MAX\_VALUE |
| 5 | 1 | 实验变量 | 0.5h | 无 | MIN\_TIME |
| 6 | 1 | 实验变量 | 0.5h | 无 | MIN\_VALUE |
| 7 | 1 | 实验变量 | 0.5h | 无 | SUM |
| 8 | 1 | 实验变量 | 0.5h | 无 | AVG |

**A）测试场景1：无背景写入单设备多传感器的COUNT聚合**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | COUNT |
| 写入背景 | 无 |
| 时间间隔 | 0.5h |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB聚合函数Count查询单设备多传感器实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB聚合函数Count查询单设备多传感器实验结果

**B）测试场景2：无背景写入多设备单传感器的COUNT聚合**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_SENSOR\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | COUNT |
| 写入背景 | 无 |
| 时间间隔 | 0.5h |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_DIVICE\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB聚合函数Count查询多设备单传感器实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_DIVICE\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB聚合函数Count查询多设备单传感器实验结果

**C）测试场景3：无背景写入单设备单传感器的MAX\_VALUE聚合在不同时间范围查询性能**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| QUERY\_SENSOR\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | MAX\_VALUE/MAX |
| 写入背景 | 无 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_INTERVAL |
|  |  |  |  |  | 20000 |
|  |  |  |  |  | 60000 |
|  |  |  |  |  | 1800000 |
|  |  |  |  |  | 3600000 |
|  |  |  |  |  | 7200000 |

IoTDB聚合函数MAX\_VALUE查询单设备单传感器不同时间间隔实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_INTERVAL |
|  |  |  |  |  | 20000 |
|  |  |  |  |  | 60000 |
|  |  |  |  |  | 1800000 |
|  |  |  |  |  | 3600000 |
|  |  |  |  |  | 7200000 |

InfluxDB聚合函数MAX查询单设备单传感器不同时间间隔实验结果

**D）测试场景4：无背景写入单设备多传感器的MAX\_VALUE聚合查询性能**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | MAX\_VALUE/MAX |
| 时间间隔 | 0.5h |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB聚合函数MAX\_VALUE查询单设备多传感器实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB聚合函数MAX查询单设备多传感器实验结果

**E）测试场景5：无背景写入单设备多传感器的MIN\_TIME聚合查询性能**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | MIN\_TIME/FIRST |
| 时间间隔 | 0.5h |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB聚合函数MIN\_TIME查询单设备多传感器实验结果

InfluxDB聚合函数FIRST查询单设备多传感器实验结果

**F）测试场景4：无背景写入单设备多传感器的MIN\_VALUE聚合查询性能**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | MIN\_VALUE/MIN |
| 时间间隔 | 0.5h |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

IoTDB聚合函数MIN\_VALUE查询单设备多传感器实验结果

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | rate | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB聚合函数MIN查询单设备多传感器实验结果

# GroupBy查询性能测试

## 9.1业务背景

业务需要聚合查询来查看设备的某些统计信息。如查看一段时间内的cpu平均利用率等。

## 9.2 实验目的

测试IoTDB聚合查询在不同参数下的不同聚合函数的平均速度。根据测试实验结果了解不同参数对系统性能的影响，综合评价系统。

重点考察以下参数和设置对系统性能的影响：

|  |  |
| --- | --- |
| 参数名 | 说明 |
|  | 有无写入背景（或者查询插入占比）目前不支持 |
| CLIENT\_NUMBER | 客户端线程数（并发度），给每个客户端均匀分配Device |
| QUERY\_SENSOR\_NUM | 一条查询语句包含的传感器数目 |
| LOOP | 循环次数，其他参数不变时，主要通过这个参数改变数据量 |
| QUERY\_DEVICE\_NUM | 一条查询语句包含的设备数目，探究不同存储组 |
| QUERY\_AGGREGATE\_FUN | 待测试的聚合函数种类 |

## 9.3 实验设计

#### 9.3.1 度量指标

1. 性能指标

1) 查询数据点数 : 查询所包含的数据点数。

2) 总时间：多个客户端并发所花费的写入总时间的最大值，即最后一个结束的客户端的写入总时间。

4)平均查询速率：查询数据点数/总时间

2. 系统信息指标

1. 客户机及服务器的磁盘IO使用率。
2. 客户机及服务器的CPU使用率。
3. 客户机及服务器的内存使用率。
4. 客户机及服务器的以太网口吞吐率(KB/s)。

#### 9.3.2 实验数据

实验采用iotdb-benchmark工具自动生成的随机数据。数据类型包括整型和浮点型。总体数据量大小受四个参数的影响：DEVICE\_NUMBER，SENSOR\_NUMBER，CACHE\_NUM，LOOP。范围查询测试采用插入测试阶段已经插入的数据，参数值如下表所示：（大约3.9G）

|  |  |
| --- | --- |
| 参数名 | 值 |
| 存储组数 | 10 |
| 总设备数 | 100 |
| 每个设备传感器数 | 100 |
| 数据产生间隔/s | 5 |
| 总点数 | 10000\*50000 |

#### 备注1：其中每一个root.performf.grop\_x文件夹下均只有一个tsfile文件

备注2：数据集相当于100个设备，每个设备有100个传感器，在数据产生间隔为5s的情况下，经过69.444444h产生的数据量

当查询传感器数目与查询设备数目均为1且时间间隔为0.5h时，查询语句为：SELECT max\_value(s\_81) FROM root.performf.group\_9.d\_92 WHERE root.performf.group\_9.d\_92.s\_81 > 0.0 GROUP BY(600000ms, 1262275200000,[2010-01-01 12:00:00,2010-01-01 13:00:00])

#### 9.3.3场景设计

针对可能影响groupBy查询所需时间的返回结果数据点数，背景写入，过滤条件，聚合类型，客户端个数等因素设计多组测试场景彼此对照。

**测试场景1：单设备多传感器值Max聚合函数**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_DEVICE\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | MIN\_VALUE/MIN |
| 时间间隔 | 2min |
| 时间分组大小 | 25s |
| 过滤条件 | Values>0 |

IoTDB单设备多传感器MAX函数groupBY查询

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | Rate（每秒返回点数） | QUERY\_SENSOR\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB单设备多传感器MAX函数groupBY查询

**测试场景2：多设备单传感器值Max聚合函数**

固定参数表

|  |  |
| --- | --- |
| 参数名 | 值 |
| LOOP | 5000 |
| QUERY\_SENSOR\_NUMBER | 1 |
| CLIENT\_NUMBER | 4 |
| 聚合函数 | MIN\_VALUE/MIN |
| 时间间隔 | 2min |
| 时间分组大小 | 25s |
| 过滤条件 | Values>0 |

IoTDB多设备单传感器MAX函数groupBY查询

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| queryNum | point | time | （每秒执行的查询数目）query/s | Rate（每秒返回点数） | QUERY\_DIVICE\_NUM |
|  |  |  |  |  | 1 |
|  |  |  |  |  | 5 |
|  |  |  |  |  | 10 |
|  |  |  |  |  | 20 |
|  |  |  |  |  | 50 |

InfluxDB多设备单传感器MAX函数groupBY查询

# 模糊点查询性能测试