**InfluxDB设备数增长实验报告**

1. **实验目的**

本次实验是为了找出当Influxdb的写入设备数增多之后，系统性能的整体变化，重点观察不同数量设备数写入InfluxDB时，写入速率，查询速率以及写入和查询过程中的内存利用率，CPU利用率，磁盘利用率的变化。

1. **实验设计**

**2.1硬件环境**

实验使用2台服务器

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 服务器名 | IP | 用户名密码 | CPU | 内存 | 硬盘 | 备注 |
| 19服务器 | 192.168.130.19 | 用户名：hadoop  密码：hadoop | 1\*8核 | 32G | 1.8T | 服务器 |
| 17服务器 | 192.168.130.17 | 用户名：hadoop  密码：hadoop | 1\*8核 | 32G | 1.8T | 客户机 |

**2.2软件环境**

|  |  |
| --- | --- |
| 操作系统 | Ubuntu14.04.3 LTS |
| Java版本 | 1.8.0\_101 |
| InfluxDB版本 | 1.4.2 |

**2.3实验设计**

**2.3.1写入测试**

**固定参数表：**

|  |  |
| --- | --- |
| 存储组 | 10 |
| 传感器 | 100 |
| BATCH\_NUM | 1000 |
| LOOP | 5 |
| 客户端数目 | 10 |

说明：写入数据无overflow且每个batch仅插入一个device的数据。

**测试用例表：**

|  |  |  |
| --- | --- | --- |
| 实验编号 | 设备数 | 总点数 |
| 1 | 1000 | 5亿 |
| 2 | 10000 | 50亿 |
| 3 | 100000 | 500亿 |

**2.3.2查询测试**

针对上述3个写入测试的写入数据集，分别进行了2组查询测试，分别为精确点查询和groupBy查询。

其中，精确点查询语句：

Select s\_0 from group\_0 where device=‘d\_0’AND time=1509465600000000000;

GroupBy查询语句：

Select s\_0 from group\_0 where (device=‘d\_0’) AND time > 1509465600000000000 AND time < 1509469200000000000 GROUP BY time(10m)

**固定参数表：**

|  |  |
| --- | --- |
| 每条查询语句包含设备数 | 1 |
| 每条查询语句包含传感器数 | 1 |
| LOOP | 1000 |
| 客户端数目 | 10 |
| GroupBy查询过滤条件 | value>-5 |
| GroupBy时间聚合单位 | 10min |
| GroupBy时间范围 | 1h |

**测试用例表：**

|  |  |  |
| --- | --- | --- |
| 实验编号 | 设备数 | 总点数 |
| 1.1 | 1000 | 5亿 |
| 1.2 | 1000 | 5亿 |
| 2.1 | 10000 | 50亿 |
| 2.2 | 10000 | 50亿 |
| 3.1 | 100000 | 500亿 |
| 3.2 | 100000 | 500亿 |

1. **实验结果**

写入测试

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| 实验ID | 存储组 | 设备 | 传感器 | 总点数 | 时间（s） | 速度（point/s） |
| 1 | 10 | 1000 | 100 | 500000000 | 306.789 | 162.978464w |
| 2 | 10 | 10000 | 100 | 5000000000 | 3423.73 | 146.039553w |
| 3 | 10 | 100000 | 100 | 50000000000 | 40590.987 | 123.180055w |

精确点查询测试：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验ID | 查询次数 | 返回点数 | 时间（s） | 速度（query/s） | 速度（points/s） |
| 1.1 | 10000 | 10000 | 1.334346375 | 7494.306298215913 | 7494.306298215913 |
| 2.1 | 10000 | 10000 | 56.260028 | 177.74608976736167 | 177.74608976736167 |
| 3.1 | 10000 | 10000 | 278.001792 | 35.97099145819583 | 35.97099145819583 |

groupBy查询测试：

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 实验ID | 查询次数 | 返回点数 | 时间（s） | 速度（query/s） | 速度（points/s） |
| 1.2 | 10000 | 69248 | 1.856921125 | 5385.25835 | 37291.83871243999 |
| 2.2 | 10000 | 66203 | 63.003828 | 158.720515 | 1050.7774197349768 |
| 3.2 | 10000 | 68576 | 356.09904 | 28.0820751 | 192.57564818037153 |

1. **实验结论**

结论1：当设备数增大10倍时，写入速度仅下降10%左右。

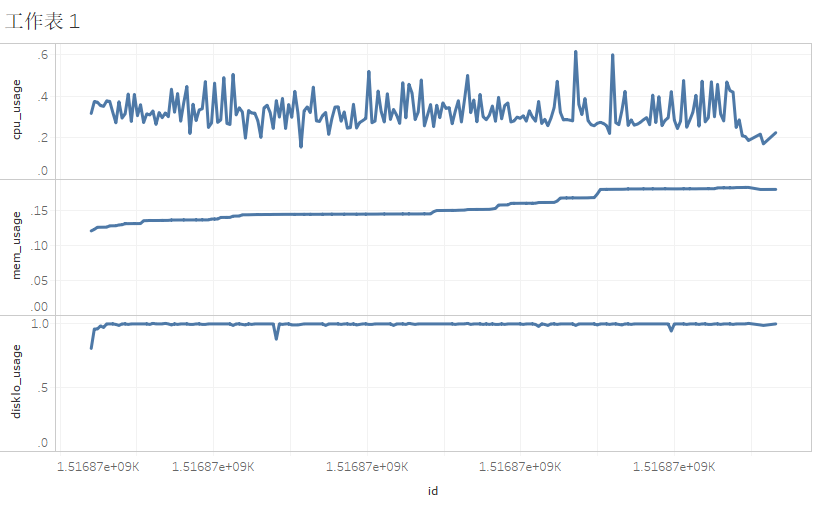
结论2：当设备数从1000增大到10000时，查询性能下降了40倍，而设备数从10000增大到100000时，查询性能下降了5倍左右。

结论2:在写入过程中InfluxDB的CPU利用率几乎在0.2~0.6之间，但会有一些周期性的高峰。

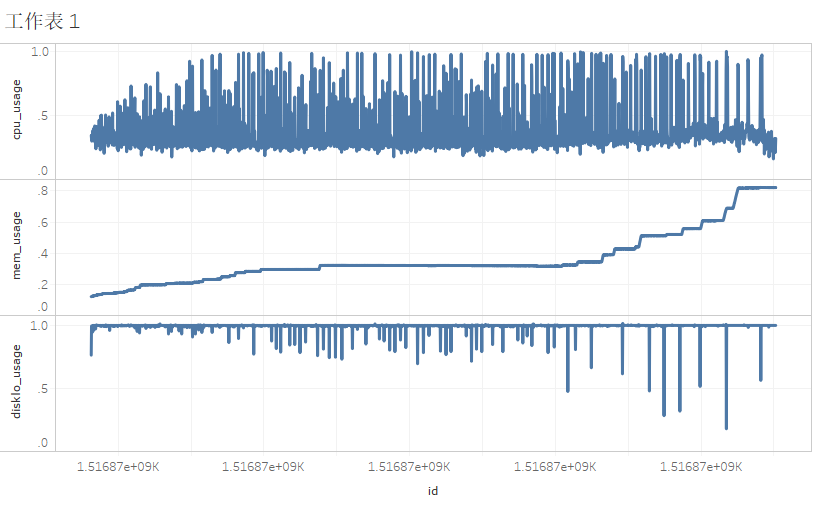
结论3：InfluxDB的内存利用率随着写入时间变长有增长趋势，在大数据量的时候尤其明显。

结论4：InfluxDB的磁盘利用率非常高，写入实验中都接近1。

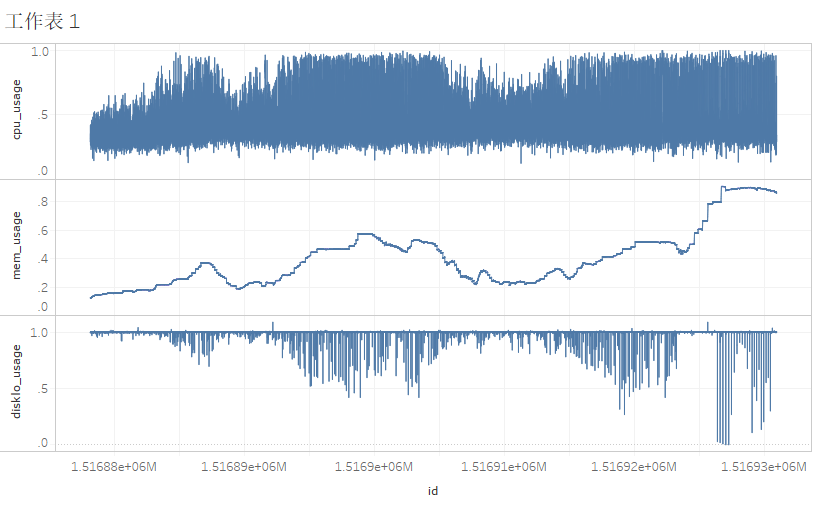
1. **附录**

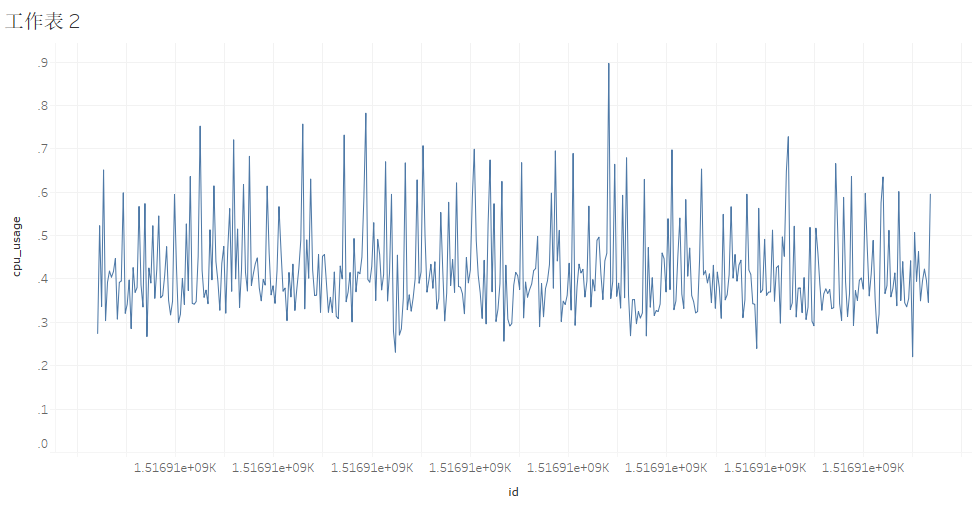


实验1对应CPU，内存，磁盘利用率



实验2对应CPU，内存，磁盘利用率

实验3对应CPU，内存，磁盘利用率



实验3对应一段时间内CPU变化

注：之前写入测试结果为90w Points/s为BATCH\_NUM=500时的结果；