# H2,sqlite

1. 内存模式只支持单个连接，重新建立连接将查询不到上个连接的内存数据，所以内存模式基本不可用

# Derby

特点：

1. 有内嵌和独立部署两种模式。
2. 在分布式环境下，多台机器如果使用内嵌模式，每台机器都要加载全量数据，内存资源占用较多，每台机器的数据同步也是问题，所以只能使用独立部署的方式。
3. 独立部署模式下，支持多连接访问，不存在sqlite和h2内存模式只支持单个连接的问题。

问题：

1. 数据插入速度很慢，实测插入8W条数据用时在4分钟左右。
2. 数据插入过程中，查询速度很慢，是正常查询速度的十倍左右。
3. 聚合条件查询，返回结果数据量较大时，性能优势并不明显；返回结果数据量较小时，性能优势明显。查询结果见附录
4. 不支持中文别名
5. 内存占用大

# Mongodb

特点：

1. 可单台和集群部署
2. 首次查询时，内存由硬盘加载至内存，后续查询直接使用内存数据；
3. Mysql数据导出成csv文件，可通过mongodb命令直接导入，230W数据导入速度在三分钟左右，速度很快。

问题：

1. 不支持sql语法，换mongodb的话，nbd sql引擎和图表数据生成的部分改动较大，配置方式也要发生修改。
2. 聚合查询时，返回结果数据量较大时，性能优势并不明显；返回结果数据量较小时，性能优势明显。
3. 内存占用大

# 附录：查询时间实测

### SQL1（2万条数据聚合）：

select sum(pepole\_nums) as aabb,start\_province\_name as bbaa from ol\_itf\_nbd\_tourist\_info where (start\_date>='2017-04-16' and start\_date<='2017-05-16') and (target\_province\_name in('西藏','河北','新疆','台湾','四川','陕西','山西','上海','山东','青海','澳门','宁夏','内蒙古','吉林','江西','安徽','江苏','北京','湖南','重庆','福建','湖北','甘肃','香港','河南','广东','浙江','广西','云南','黑龙江','贵州','海南','辽宁')) group by start\_province\_name order by sum(pepole\_nums) desc

### SQL2（142万条数据聚合）：

select sum(pepole\_nums) as aabb,start\_province\_name as bbaa from ol\_itf\_nbd\_tourist\_info where (start\_date>='2016-04-16' and start\_date<='2017-05-16') and (target\_province\_name in('西藏','河北','新疆','台湾','四川','陕西','山西','上海','山东','青海','澳门','宁夏','内蒙古','吉林','江西','安徽','江苏','北京','湖南','重庆','福建','湖北','甘肃','香港','河南','广东','浙江','广西','云南','黑龙江','贵州','海南','辽宁')) group by start\_province\_name order by sum(pepole\_nums) desc

### SQL3（210万条数据聚合）：

select sum(pepole\_nums) as aabb,start\_province\_name as bbaa from ol\_itf\_nbd\_tourist\_info where (start\_date>='2010-04-16' and start\_date<='2017-05-16') and (target\_province\_name in('西藏','河北','新疆','台湾','四川','陕西','山西','上海','山东','青海','澳门','宁夏','内蒙古','吉林','江西','安徽','江苏','北京','湖南','重庆','福建','湖北','甘肃','香港','河南','广东','浙江','广西','云南','黑龙江','贵州','海南','辽宁')) group by start\_province\_name order by sum(pepole\_nums) desc

|  |  |  |  |
| --- | --- | --- | --- |
|  | Mysql（远程开发库） | Derby | Mongodb |
| 数据总量 | 2782699 | 2782699 | 2782699 |
| 是否使用索引 | 是 | 是 | 是 |
| 占用内存 |  | 1.6G | 1.2G |
| SQL1执行平均耗时 | 3s | 0.9s | 0.3s |
| SQL2执行平均耗时 | 3.3s | 1.7s | 3.3s |
| SQL3执行平均耗时 | 3.8s | 2.2s | 4.8s |

# 进程内存占用截图

***Derby：***

***Mongo：***