| Cedar 项目组测试文档 | |
|---|---|
| 文档名称 | Hash join 测试案例 |
| 作者（测试人员） | 潘宇晨、胡爽 |
| 功能模块 | Hash join |
| 开发人员 | 刘小兵、茅潇潇 |
| 日期 | 20170621 |
| 负责老师 | 测试：张蓉　　开发：张召 |

## 修订记录：

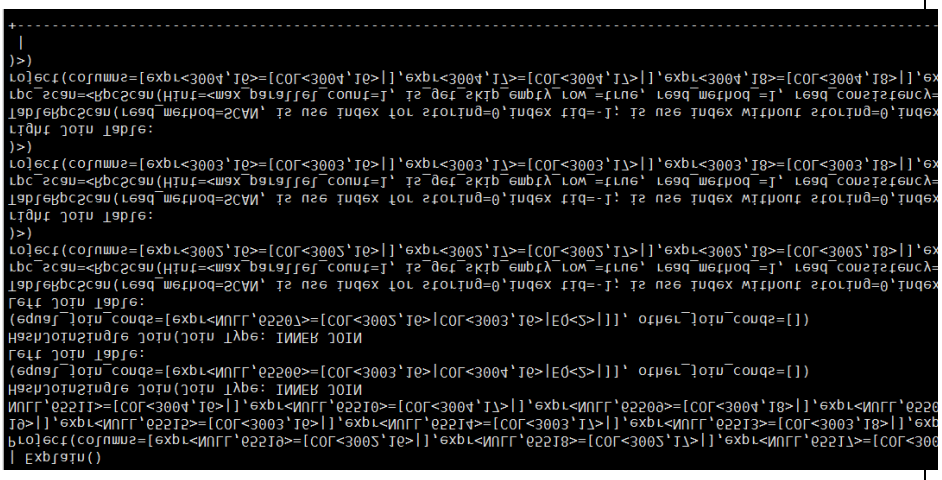
| 日期 | 修改描述 |
|---|---|
| 20170621 | 创建文档 |
| 20170722 | 完善文档 |
| 20170916 | 回归测试 |

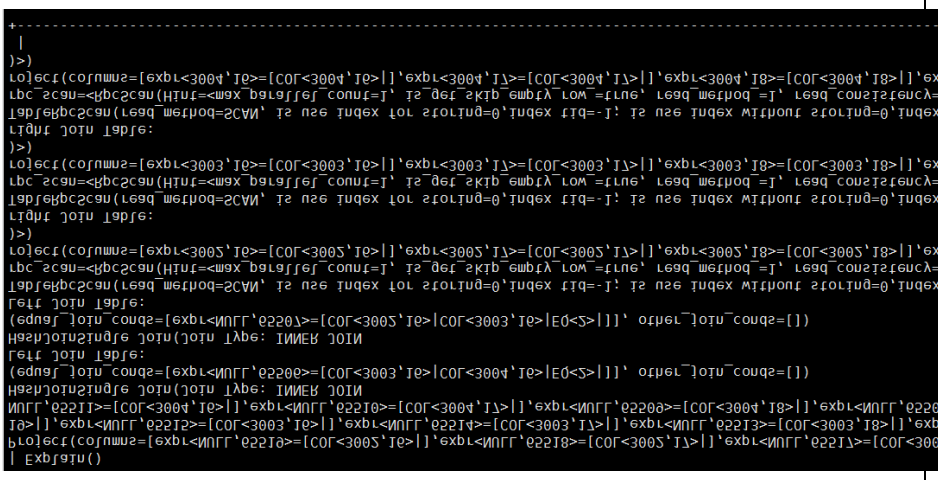
## 硬件配置：

| 机器 IP | 硬件配置 |
|---|---|
| 10.11.1.190-10.11.1.198 | CPU：Intel(R) Xeon(R) CPU E5-2620 * 2，2 * 6 *2 个线程，主频 2000MHz- 2500MHz，L3 缓存 15MB；<br><br>内存：168GB、152GB、158 GB、168GB、80GB、144GB、128 GB、112 GB、144 GB、128 GB；<br><br>网络带宽：1000Mb/s（有少数部分机器之间的网络带宽为 100Mb/s）；<br><br>磁盘 IOPS：76*4=304，磁盘带宽 400MB/s、6500MB/s（读缓存） |

## 功能测试案例：

| 编号 | 1 | 配置 | 197，单集群，RS+UPS+CS+MS |
|---|---|---|---|
| 测试目的 | 测试基本 JOIN 功能（where/from 条件-两表连接） | | |
| 测试输入 | create table t1(k int primary key, c1 int,c2 varchar(20),c3 int);<br>insert into t1<br>values(1,1,'aaaa',1),(2,2,'bbbb',1),(3,3,'cccc',1),(4,4,'eefd',1),(5,5,'ddee',1),(6,6,'bbffee',1),(7,6,'bbffee',1);<br>create table t2(k int primary key, c1 int,c2 varchar(20),c3 int);<br>insert into t2<br>values(1,1,'aaaa',1),(2,7,'bbbb',1),(3,2,'cccc',1),(4,8,'eefd',1),(5,1,'ddee',1),(6,16,'bbffee',1); | | |

| | | | |
|---|---|---|---|
| | explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.k=t2.k;<br>explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.k=t2.k;<br>select * from t1, t2 where t1.k=t2.k;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k; | | |
| 测试结果 | PASS | | |
| 编号 | 2 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（where 条件-多表连接） | | |
| 测试输入 | explain select /*+JOIN(hash_join, MERGE_JOIN)*/ * from t1, t2, t3 where t1.k=t2.k and t2.k=t3.k;<br>ERROR<br>explain select /*+JOIN(hash_join, hash_join, MERGE_JOIN)*/ * from t1, t2, t3, t4 where t2.c3 >0 and t1.k=t2.k and t1.k=t3.k and t3.c1 < 7 and t2.k=t4.k;<br>ERROR<br>explain select /*+JOIN(hash_join, MERGE_JOIN)*/ * from t1 join t2 on t1.k=t2.k, t3 where t2.k=t3.k;<br>ERROR<br>explain select /*+JOIN(hash_join, hash_join, MERGE_JOIN)*/ * from t1 join t2 on t1.k=t2.k, t3 , t4 where t2.c3 >0 and t1.k=t3.k and t3.c1 < 7 and t2.k=t4.k;<br>ERROR | | |
| 测试结果 | ERROR<br>  | | |
| 编号 | 3 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（where 条件-多表连接-hint 内数目小于连接数目） | | |
| 测试输入 | explain select /*+JOIN(hash_join, MERGE_JOIN)*/ * from t1, t2, t3, t4 where t2.c3 >0 and t1.k=t2.k and t1.k=t3.k and t3.c1 < 7 and t2.k=t4.k;<br>explain select /*+JOIN(hash_join, MERGE_JOIN)*/ * from t1 join t2 on t1.k=t2.k, t3 , t4 where t2.c3 >0 and t1.k=t3.k and t3.c1 < 7 and | | |

| | | | |
|---|---|---|---|
| | t2.k=t4.k; | | |
| 测试结果 | ERROR<br>  | | |
| 编号 | 4 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（where/from 条件-条件过滤） | | |
| 测试输入 | 1）like<br>explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t2.c2 like 'bb%' and t1.c1 > 5;<br>explain select /*+JOIN(hash_join)*/ * from t1,t2 where t2.c2 like 'bb%' and t1.k=t2.k and t1.c1 > 5;<br><br>select /*+JOIN(hash_join)*/ * from t1,t2 where t2.c2 like 'bb%' and t1.k=t2.k and t1.c1 >5;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t2.c2 like 'bb%' and t1.c1 > 5;<br>select * from t1,t2 where t2.c2 like 'bb%' and t1.k=t2.k and t1.c1 > 5;<br>PASS<br>2) between, in, not like<br>select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and t1.c1 between 1 and 8 and t2.c1 in (3, 7, 8) and t1.c1 not in (4,3) and t2.c2 not like 'ff%';<br>select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and t1.c1 between 1 and 8 and t2.c1 in (3, 7, 8) and t1.c1 not in (4,3) and t2.c2 not like 'ff%' order by t1.k;<br>select * from t1 , t2 where t1.k=t2.k and t1.c1 between 1 and 8 and t2.c1 in (3, 7, 8) and t1.c1 not in (4,3) and t2.c2 not like 'ff%';<br>PASS<br>3) %, not like, in , !=<br>select /*+JOIN(hash_join)*/ * from t1 ,t2, t3 where t1.k=t2.k and t1.c1 = t3.c1 and t1.c2 not like 'g%' and t3.c1 in (1, 2, 3, 4, 5, 6, 7, 8, 9) and t3.c2 != 'aa' and t2.k % 2=0;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k, t3 where | | | |

t1.c1 = t3.c1 and t1.c2 not like 'g%' and t3.c1 in (1, 2, 3, 4, 5, 6, 7, 8, 9) and t3.c2 != 'aa' and t2.k % 2=0;
select * from t1 ,t2, t3 where t1.k=t2.k and t1.c1 = t3.c1 and t1.c2 not like 'g%' and t3.c1 in (1, 2, 3, 4, 5, 6, 7, 8, 9) and t3.c2 != 'aa' and t2.k % 2=0;
PASS

4) group order sum
explain select /*+JOIN(hash_join, hash_join, hash_join, hash_join, hash_join)*/ sum(t2.k) from t1 left join t2 on t1.k=t2.k inner join t3 on t2.c1=t3.k, t4, t5, t6 where t3.c1 <= 10 and t3.k=t4.k and t5.c1 >= 1 and t4.k = t5.c1 and t5.c2 = t6.c1 group by t2.c2 order by t4.c1 desc;
select /*+JOIN(hash_join, hash_join, hash_join, hash_join, hash_join)*/ sum(t2.k) from t1 left join t2 on t1.k=t2.k inner join t3 on t2.c1=t3.k, t4, t5, t6 where t3.c1 <= 10 and t3.k=t4.k and t5.c1 >= 1 and t4.k = t5.k and t5.k = t6.k group by t2.c2 order by t4.c1 desc;
select sum(t2.k) from t1 left join t2 on t1.k=t2.k inner join t3 on t2.c1=t3.k, t4, t5, t6 where t3.c1 <= 10 and t3.k=t4.k and t5.c1 >= 1 and t4.k = t5.k and t5.k = t6.k group by t2.c2 order by t4.c1 desc;

5)> < =
Explain select /*+JOIN(hash_join)*/ * from t1 left join t2 on t1.k=t2.k where (t1.c1 <> 4 and t2.c1 <= 7) or (t2.c1 = 9) or (t2.c3 = 8);
select /*+JOIN(hash_join)*/ * from t1 left join t2 on t1.k=t2.k where (t1.c1 <> 4 and t2.c1 <= 7) or (t2.c1 = 9) or (t2.c3 = 8);
select * from t1 left join t2 on t1.k=t2.k where (t1.c1 <> 4 and t2.c1 <= 7) or (t2.c1 = 9) or (t2.c3 = 8);

6)mod
Explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t2.c1 mod 2 = 0 order by t1.k;
select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t2.c1 mod 2 = 0 order by t1.k;
select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and t2.c1 mod 2 = 0 order by t1.k;
select * from t1 join t2 on t1.k=t2.k where t2.c1 mod 2 = 0 order by t1.k;

7)not
Explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where not t1.c1 > 9;
select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where not t1.c1 > 9;
select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and not

| | |
|---|---|
| | t1.c1 > 9;<br>select * from t1 join t2 on t1.k=t2.k where not t1.c1 > 9;<br><br>8) sum<br>Explain select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and t1.c1 > 5 group by t2.c2 having sum(t2.c1) > 6;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t1.c1 > 5 group by t2.c2 having sum(t2.c1) > 6;<br>select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and t1.c1 > 5 group by t2.c2 having sum(t2.c1) > 6;<br>select * from t1 join t2 on t1.k=t2.k where t1.c1 > 5 group by t2.c2 having sum(t2.c1) > 6;<br><br>9) max<br>Explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t1.c1 > 5 group by t2.c2 having max(t2.c1) > 6;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where t1.c1 > 5 group by t2.c2 having max(t2.c1) > 6;<br>select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and t1.c1 > 5 group by t2.c2 having max(t2.c1) > 6;<br>select * from t1 join t2 on t1.k=t2.k where t1.c1 > 5 group by t2.c2 having max(t2.c1) > 6;<br><br>10) coalesce<br>Explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where coalesce(t1.c1, t1.k) > 3 order by t1.c1;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k where coalesce(t1.c1, t1.k) > 3 order by t1.c1;<br>select * from t1 join t2 on t1.k=t2.k where coalesce(t1.c1, t1.k) > 3 order by t1.c1;<br><br>11)NULL NOT NULL<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.c1=t2.c1 where t1.c1 is NULL;<br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.c1=t2.c1 where t1.c1 is not NULL; |
| 测试结果 | PASS |
| 编号 | 5 \| 配置 \| 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（不支持语句） |
| 测试输入 | select /*+JOIN(hash_join)*/ * from t1 , t2 where t1.k=t2.k and (t1.c1 <> 4 and t2.c1 <= 7) or (t2.c1 = 9) or (t2.c3 = 8);<br>ERROR 7 (0A000): OB-7: Not supported feature or function<br><br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k and t1.c1 > |

| | |
|---|---|
| | 5;<br>ERROR 7 (0A000): OB-7: Not supported feature or function<br><br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c1+t2.c1>10;<br>ERROR 7 (0A000): OB-7: Not supported feature or function<br><br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k ,t3 where t1.c1=t2.c1 and t3.k=t1.k;<br>ERROR 6 (HY000): OB-6: The object is not initialized<br><br>select /*+JOIN(hash_join)*/ * from t1 join t1 on t1.k=t1.k;<br>ERROR 5020 (42S01): table t1 is ambiguous<br><br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.k=t2.k+1;<br>ERROR 7 (0A000): OB-7: Not supported feature or function<br><br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.k+t2.k>10;<br>ERROR 7 (0A000): OB-7: Not supported feature or function<br><br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.k>t2.k;<br>ERROR 7 (0A000): OB-7: Not supported feature or function<br><br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=4;<br>ERROR 7 (0A000): OB-7: Not supported feature or function |
| 测试结果 | |
| 编号 | 6 \| 配置 \| 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（hint 语法错误/ Hint 内数目小于连接数目） |
| 测试输入 | explain select /*+JOIN(HASH_JOIN)*/ * from t1, t2 where t1.k=t2.k;<br>explain select /*+JOIN()*/ * from t1, t2 where t1.k=t2.k;<br>explain select /*+JOIN(HHHH)*/ * from t1, t2 where t1.k=t2.k;<br>PASS<br><br>Hint 内数目小于连接数目<br>explain select /*+JOIN(hash_join, MERGE_JOIN)*/ * from t1, t2, t3, t4 where t2.c3 >0 and t1.k=t2.k and t1.k=t3.k and t3.c1 < 7 and t2.k=t4.k;<br>三个 join 都走 hash join<br>explain select /*+JOIN(MERGE_JOIN, hash_join)*/ * from t1, t2, t3, t4 where t2.c3 >0 and t1.k=t2.k and t1.k=t3.k and t3.c1 < 7 and t2.k=t4.k;<br>三个 join 都走 MERGE_JOIN |

| | | | |
|---|---|---|---|
| | explain select /*+JOIN(hash_join, MERGE_JOIN)*/ * from t1 join t2 on t1.k=t2.k, t3 , t4 where t2.c3 >0 and t1.k=t3.k and t3.c1 < 7 and t2.k=t4.k;<br>MERGE_JOIN，MERGE_JOIN， hash join | | |
| 测试结果 | ERROR | | |
| 编号 | 7 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（数据类型） | | |
| 测试输入 | 1）varchar join varchar<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c2 = t2.c2;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c2 = t2.c2 order by t1.k;<br>select * from t1, t2 where t1.c2 = t2.c2 order by t1.k;<br>PASS<br>2)double join souble<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c3 = t2.c3 ;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c3 = t2.c3 order by t1.k;<br>select * from t1, t2 where t1.c3 = t2.c3 order by t1.k;<br>PASS<br>3)int JOIN varchar<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c1 = t2.c2 ;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c1 = t2.c2 order by t1.k;<br>select * from t1, t2 where t1.c1 = t2.c2 order by t1.k;<br>4) varchar join int<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c2 = t2.c1 ;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c2 = t2.c1 order by t1.k;<br>select * from t1, t2 where t1.c2 = t2.c1 order by t1.k;<br>5) int join double<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c1 = t2.c3 ;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c1 = t2.c3 order by t1.k;<br>select * from t1, t2 where t1.c1 = t2.c3 order by t1.k;<br>6) double join int<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c3 = t2.c1 ;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c3 = t2.c1 order by t1.k;<br>select * from t1, t2 where t1.c3 = t2.c1 order by t1.k;<br>7) double join varchar<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c3 = t2.c2 ;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c3 = t2.c2 order by t1.k;<br>select * from t1, t2 where t1.c3 = t2.c2 order by t1.k; | | |

| | | | |
|---|---|---|---|
| | 8) varchar join double<br>explain select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c2 = t2.c3;<br>select /*+JOIN(hash_join)*/ * from t1, t2 where t1.c2 = t2.c3 order by t1.k;<br>select * from t1, t2 where t1.c2 = t2.c3 order by t1.k; | | |
| 测试结果 | 不支持不同数据类型连接 | | |
| 编号 | 8 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（连接类型） | | |
| 测试输入 | explain select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k;<br>explain select /*+JOIN(hash_join)*/ * from t1 inner join t2 on t1.k=t2.k;<br>explain select /*+JOIN(hash_join)*/ * from t1 left join t2 on t1.k=t2.k;<br>explain select /*+JOIN(hash_join)*/ * from t1 right join t2 on t1.k=t2.k;<br>explain select /*+JOIN(hash_join)*/ * from t1 full join t2 on t1.k=t2.k;<br><br>select /*+JOIN(hash_join)*/ * from t1 join t2 on t1.k=t2.k;<br>select /*+JOIN(hash_join)*/ * from t1 inner join t2 on t1.k=t2.k;<br>select * from t1 join t2 on t1.k=t2.k;<br><br>select /*+JOIN(hash_join)*/ * from t1 left join t2 on t1.k=t2.k;<br>select * from t1 left join t2 on t1.k=t2.k;<br><br>select /*+JOIN(hash_join)*/ * from t1 right join t2 on t1.k=t2.k;<br>select * from t1 right join t2 on t1.k=t2.k;<br><br>select /*+JOIN(hash_join)*/ * from t1 full join t2 on t1.k=t2.k;<br>select * from t1 full join t2 on t1.k=t2.k;<br><br>explain select /*+JOIN(hash_join, hash_join, hash_join, hash_join, hash_join)*/ * from t1 inner join t2 on t1.k=t2.k left join t3 on t2.k=t3.k right join t4 on t3.k=t4.k full join t5 on t4.k=t5.k join t6 on t5.k=t6.k where t5.c1 between 1 and 5;<br>select /*+JOIN(hash_join, hash_join, hash_join, hash_join, hash_join)*/ * from t1 inner join t2 on t1.k=t2.k left join t3 on t2.k=t3.k right join t4 on t3.k=t4.k full join t5 on t4.k=t5.k join t6 on t5.k=t6.k where t5.c1 between 1 and 5;<br>select * from t1 inner join t2 on t1.k=t2.k left join t3 on t2.k=t3.k right join t4 on t3.k=t4.k full join t5 on t4.k=t5.k join t6 on t5.k=t6.k where t5.c1 between 1 and 5; | | |
| 测试结果 | PASS | | |
| 编号 | 9 | 配置 | 197，单集群，RS+UPS+CS+MS |

| | | | |
|---|---|---|---|
| 测试目的 | 测试 JOIN 功能（表别名） | | |
| 测试输入 | explain select /*+JOIN(hash_join)*/ * from t1 tt1 join t2 tt2 on tt1.k = tt2.c1;<br>select /*+JOIN(hash_join)*/ * from t1 tt1 join t2 tt2 on tt1.k = tt2.c1;<br>select /*+JOIN(hash_join)*/ * from t1 tt1 , t2 tt2 where tt1.k = tt2.c1;<br><br>select * from t1 tt1 join t2 tt2 on tt1.k = tt2.c1; | | |
| 测试结果 | PASS | | |
| 编号 | 10 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（单表） | | |
| 测试输入 | Explain select /*+JOIN(hash_join)*/ * from t1 where t1.k=t1.c1;<br>select /*+JOIN(hash_join)*/ * from t1 where t1.k=t1.c1;<br>select * from t1 where t1.k=t1.c1;<br>没有走 hash_join<br>select /*+JOIN(hash_join)*/ * from t1 where t1.k=t1.c2;<br>select * from t1 where t1.k=t1.c2;<br><br>select /*+JOIN(hash_join)*/ * from t1 where t1.k=t1.c3;<br>select * from t1 where t1.k=t1.c3;<br><br>Explain select /*+JOIN(hash_join)*/ * from t1;<br>没有走 hash_join | | |
| 测试结果 | 正常执行不走 join | | |
| 编号 | 11 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（二级索引） | | |
| 测试输入 | T3t4 表有索引<br>explain select /*+JOIN(hash_join)*/ * from t1, t3 where t1.k=t3.c2 and t3.c1<100;<br>select /*+JOIN(hash_join)*/ * from t2, t3 where t2.c1 = t3.c1 and t3.c1<100 order by t2.k;<br>select * from t2, t3 where t2.c1 = t3.c1 and t3.c1<100 order by t2.k;<br><br>explain select /*+JOIN(hash_join)*/ * from t4, t3 where t4.c1 = t3.c1 and t3.c1<100;<br>select /*+JOIN(hash_join)*/ * from t4, t3 where t4.c1 = t3.c1 and t3.c1<100 order by t3.k;<br>select * from t4, t3 where t4.c1 = t3.c1 and t3.c1<100 order by t3.k; | | |
| 测试结果 | PASS | | |
| 编号 | 12 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试每日合并期间 JOIN 功能 | | |
| 测试输入 | 每日合并<br>explain select /*+JOIN(hash_join)*/ * from t1, t3 where t1.k=t3.c2 and | | |

| | | | |
|---|---|---|---|
| | t3.c1<100;<br>select /*+JOIN(hash_join)*/ * from t2, t3 where t2.c1 = t3.c1 and t3.c1<100 order by t2.k;<br>select * from t2, t3 where t2.c1 = t3.c1 and t3.c1<100 order by t2.k; | | |
| 测试结果 | PASS | | |
| 编号 | 13 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试基本 JOIN 功能（子查询） | | |
| 测试输入 | explain select /*+JOIN(hash_join)*/ * from (select * from t1 where t1.c3 <= 0.8) tt1 join t2 on tt1.k=t2.k where t2.c1 between 1 and 7;<br>explain select /*+JOIN(hash_join)*/ * from (select * from t1 where t1.c3 <= 0.8) tt1 join (select * from t2 where t2.c1 between 1 and 7) tt2 on tt1.k=tt2.k;<br>select * from (select * from t1 where t1.c3 >= 0.8) tt1 join t2 on tt1.k=t2.k where t2.c1 between 1 and 7;<br>select * from (select * from t1 where t1.c3 >= 0.8) tt1 join (select * from t2 where t2.c1 between 1 and 7) tt2 on tt1.k=tt2.k;<br>select /*+JOIN(hash_join)*/ * from (select * from t1 where t1.c3 >= 0.8) tt1 join t2 on tt1.k=t2.k where t2.c1 between 1 and 7;<br>select /*+JOIN(hash_join)*/ * from (select * from t1 where t1.c3 >= 0.8) tt1 join (select * from t2 where t2.c1 between 1 and 7) tt2 on tt1.k=tt2.k;<br><br>explain select /*+JOIN(MERGE_JOIN, hash_join, MERGE_JOIN, hash_join, hash_join)*/ * from (select t3.k from (select t1.k from t1, t2 where t1.k = t2.k and t1.c1 > 4) tmp1 join t3 on tmp1.k=t3.k) tmp2 left join t4 on tmp2.k=t4.k, t5, t6 where t4.c1 < 9 and t4.k=t5.k and t5.c2>=5 and t5.c2=t6.c1;<br>select /*+JOIN(MERGE_JOIN, hash_join, MERGE_JOIN, hash_join, hash_join)*/ * from (select t3.k from (select t1.k from t1, t2 where t1.k = t2.k and t1.c1 > 4) tmp1 join t3 on tmp1.k=t3.k) tmp2 left join t4 on tmp2.k=t4.k, t5, t6 where t4.c1 < 9 and t4.k=t5.k and t5.c2>=5 and t5.c2=t6.c1;<br>select * from (select t3.k from (select t1.k from t1, t2 where t1.k = t2.k and t1.c1 > 4) tmp1 join t3 on tmp1.k=t3.k) tmp2 left join t4 on tmp2.k=t4.k, t5, t6 where t4.c1 < 9 and t4.k=t5.k and t5.c2>=5 and t5.c2=t6.c1; | | |
| 测试结果 | PASS | | |
| 编号 | 14 | 配置 | 197，单集群，RS+UPS+CS+MS |
| 测试目的 | 测试 Hash Join 并行执行 | | |
| 测试输入 | 多线程对相同表相同属性进行 Hash Join<br>explain select /*+JOIN(hash_join)*/ * from t1, t3 where t1.k=t3.c2 and t3.c1<100; | | |

| | |
|---|---|
| | select /*+JOIN(hash_join)*/ * from t2, t3 where t2.c1 = t3.c1 and t3.c1<100 order by t2.k;<br>select * from t2, t3 where t2.c1 = t3.c1 and t3.c1<100 order by t2.k; |
| 测试结果 | PASS |

## 性能测试

| 编号 | 1 | 配置 | 三集群+6ms |
|---|---|---|---|
| 测试目的 | 测试 join 的深度对性能的影响 | | |
| 测试输入 | 1、2 张 500w 的表 inner join、3 张 500w 的表 inner join、4 张 500w 的表 inner join、5 张 500w 的表 inner join、6 张 500w 的表 inner join。<br>2、2 张 1000w 的表 inner join、3 张 1000w 的表 inner join、4 张 1000w 的表 inner join、5 张 1000w 的表 inner join、6 张 1000w 的表 inner join。 | | |
| 测试结果 | 随着 join 深度增加，Merge Join 和 Bloomfilter Join 所费时间大幅度上升，而 Hash Join 所耗时间缓慢上升，性能最好。 | | |

| 编号 | 2 | 配置 | 三集群+6ms |
|---|---|---|---|
| 测试目的 | 测试 Hash Join 在选择率为 10%的情况下，非重复值个数对性能的影响（连接属性为 int 类型） | | |
| 测试输入 | 左右表都为 2000 万条数据，左表的非重复值个数为（500w,800w,1000w,1600w,2000w）。 | | |
| 测试结果 | Join 表数目为 2000w 时，随着左表非重复数的增加，Hash Join 所耗时间远低于 merge join 和 bloom filter join，hash join 性能远好于 Bloom filter Join 和 Merge Join。 | | |

| 编号 | 3 | 配置 | 三集群+6ms |
|---|---|---|---|
| 测试目的 | 测试 Hash Join 在非重复值个数为 500 万下，选择率对性能的影响（连接属性为 int 类型）。 | | |
| 测试输入 | 左表 1000 万条数据（非重复数 500 万，选择率 0.1%，1%，10%，30%，60%，90%） | | |
| 测试结果 | 在非重复值个数为 500w 时，Join 左右表数目为 1000 万时，Hash Join 所耗时间低于 merge join 和 bloom filter join，hash join 性能远好于 Bloom filter Join 和 Merge Join。 | | |

| 编号 | 4 | 配置 | 三集群+6ms |
|---|---|---|---|
| 测试目的 | 测试 Hash Join 在选择率为 10%的情况下，非重复值个数对性能的影响（连接属性为 varchar 类型） | | |
| 测试输入 | 左右表 2000 万条数据（非重复数 500w,800w,1000w,1600w,2000w，连接属性为 varchar 类型） | | |

| | |
|---|---|
| 测试结果 | 在将左表的非重复值个数增加到（500w,800w,1000w,1600w,2000w）时，Join 表数目为 2000w 时，Hash Join 性能远好于 Bloom filter Join 和 Merge Join。当连接属性为 varchar 类型时，相较于 Bloom filter Join 和 Merge Join，Hash Join 性能更优。 |

| | | | |
|---|---|---|---|
| 编号 | 5 | 配置 | 三集群+6ms |
| 测试目的 | 测试 Hash Join 在非重复值个数为 500 万下，选择率对性能的影响（连接属性为 varchar 类型） | | |
| 测试输入 | 左右表 1000 万条数据（非重复数 500 万，选择率 0.1%，1%，10%，30%，60%，90%） | | |
| 测试结果 | Hash Join 性能远好于 Bloom filter Join 和 Merge Join。<br>连接属性为 varchar 类型时，相较于 Bloom filter Join 和 Merge Join，Hash Join 性能更优。 | | |