

Anjuke MySQL 规范

`dba.ops.anjuke.inc`

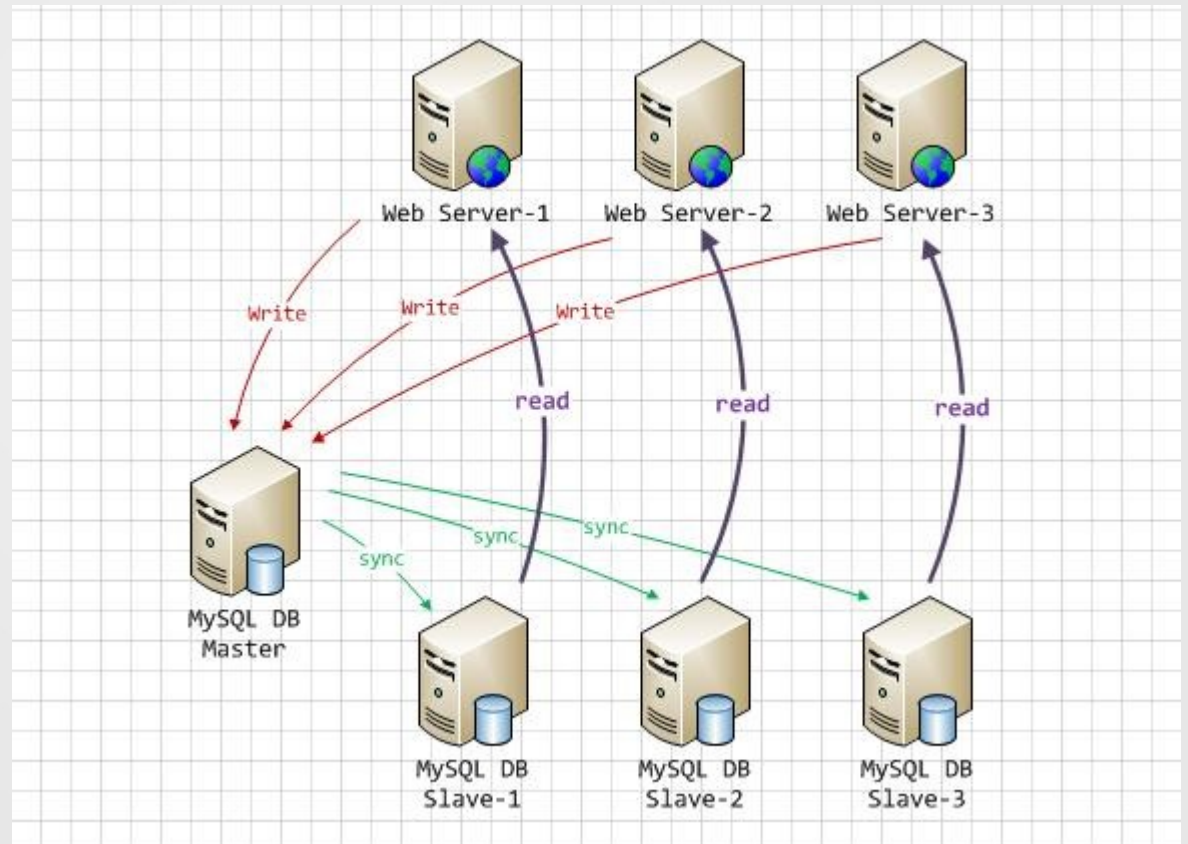
为什么选用 MySQL

- 开源免费
- 易学易用
- 资料完善
- 建站首选



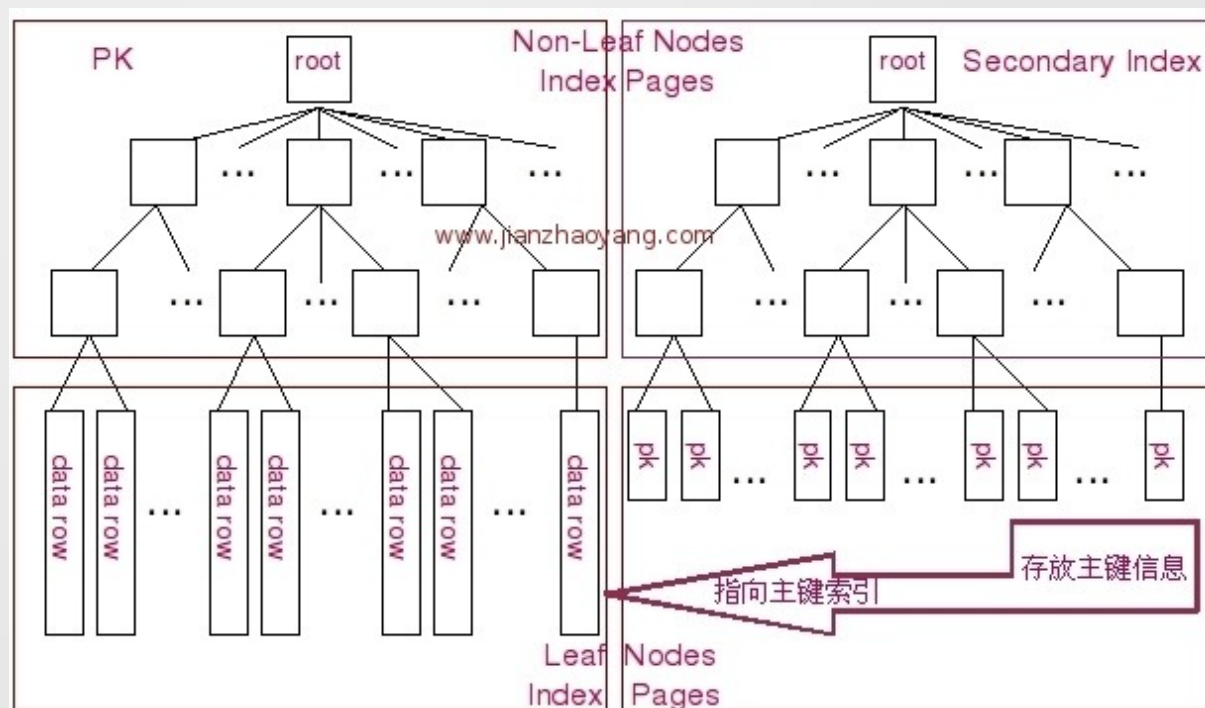
Anjuke.Inc

- 66 台 MySQL
- 13 个 Group
- MySQL 5.1
- Replication
- Proxy



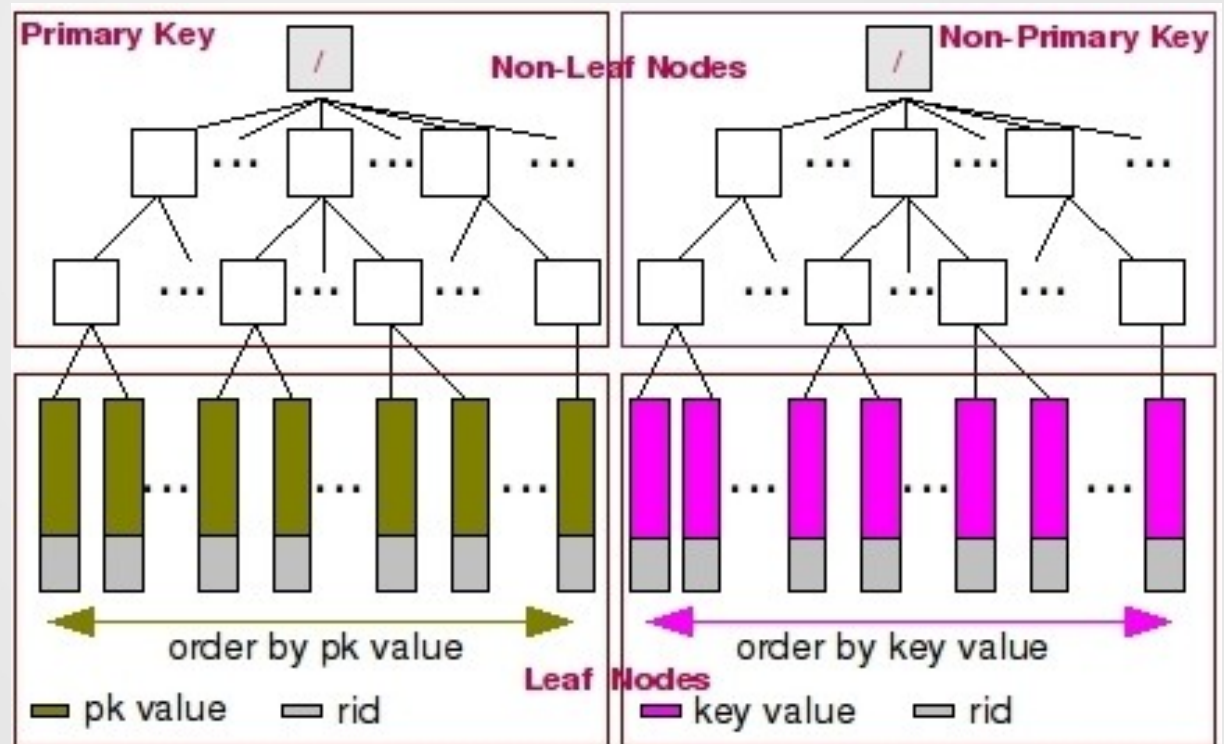
Engine – InnoDB

- .frm、.ibd
- 支持事务
- 必须有主键
- 行级锁
- 良好的读写并发
- 表故障率低
- 缓存 Key 和 Data



Engine – MyISAM

- .frm、.MYD、.MYI
- 不支持事务
- 表级锁
- 表尾追加特性
- 读写并发略低
- 表较脆弱
- 只缓存 Key



建表规范

<1> 如果字段较多，分为基础表和扩展表，最多读写和查询的字段放在基础表内，其他字段放在扩展表里，30 个字段基本是上限了，但此规则不仅限于字段较多的表

<2> 在满足使用的情况下，数据类型尽量选小的，tinyint 和 int，varchar 和 text，timestamp 和 datetime

<3> 必须有主键，可以是 MYSQL 自增主键，也可以程序控制，但必须是 int 型，因为 Master 是 InnoDB 引擎，主键查询效率很高，也有利于排序查询和 limit 查询等

<4> 不要用 NULL 值，必须用 not null，如果实在没值填，可指定 default 0 或 default ''

<5> 如果可以预计到单表数据量会很大，提前做好分表策略，例如 anjuke 的房源主表是按城市拆分的，图片表是根据房源 ID 取模 200 拆分的，还有好多每月一张的表

索引使用规范

<1> 尽量用 primary key

<2> 索引过多会影响数据写入性能

<3> 每次只能用到一个索引

<4> 复合索引，最左匹配，可包含 order by 字段，用以消除排序

如索引 (a,b)，以下 SQL 可使用到该索引

```
select * from TB where a=...
```

```
select * from TB where a=... and b=...
```

```
select * from TB where a=... order by b
```


索引使用规范

<5> 覆盖索引，select 的字段和 where 的字段在同一索引中

select b from TB where a=...

select a from TB where b=... 这样就不行了，最左匹配原则

<6> 不要对索引字段进行计算

select * from TB where mod(a,2)=...

select * from TB where a/10=...

<7> <、<=、>、>=、<>、!= 会打断复合索引的传递，使索引效果大打折扣

索引使用规范

<8> or 条件连接的问题

select * from TB where a=... or b=... 不能完整的使用索引 (a,b)

可以再建一个索引 (b) , 把语句改为

select * from TB where a=...

union

select * from TB where b=...

<9> 字符型字段索引，前缀索引，禁止左 %

<10> 注意数据类型

如上，有索引 (b) , 且字段 b 是字符型

select * from TB where b=5; 是用不了索引 (b) 的

SQL 语句规范

- <1> SQL 中不要使用 库名 . 表名 的写法
- <2> 不要在 SQL 中有函数计算
- <3> 不要有过多的 where 条件匹配
- <4> 不要有过多的 order by 字段
- <5> 禁止子查询，禁止复杂的多表 join 查询，可以拆分成多条简单 SQL
- <6> 只 select 需要的字段，避免 select *
- <7> select 必须要使用到索引，可以用 explain 查看执行计划

SQL 语句规范

<8> 禁止使用 order by rand() 实现乱序效果

<9> 禁止大偏移量的 limit M,N 查询，使用主键游标，where PK > ... limit N

<10> 减少 count() 查询，用计算器取代

彩蛋：count(*)、count(1)、count(字段) 区别

<11> update 和 delete 不能使用 limit N，要用主键，where PK= 或 where PK in (...)

<12> update 时注意对数据进行安全处理

<13> 跑 JOB 修正数据时要掌握好“节奏”

DBRT 提交规范

DBRT 中必须要有以下信息

<1> 库名

<2> 完成时间

<3> 新建表必须回答的 5 个问题

- 单天数据写入量

- 高峰期查询频率

- 是否在关键业务页面，PV 前三名的页面

- 有否 JOB 读写该表

- 主要的 insert、update、delete、select 语句

Thank you !

Talk Time

Anjuke Inc.