# 3年JAVA技能总结-数据库篇

1. 数据库三级范式

1NF:所有属性都是不可分的基本数据项。要求字段具有原子性，不可再分解；(只要是关系型数据库都满足1NF)

2NF:满足1NF，并且每个非主属性完全依赖于码。解决：消除复合主键就可避免出现部分以来，可增加单列关键字。

3NF:满足2NF，不存在非主属性对码的传递函数依赖。比如某个字段a依赖于主键，而一些字段依赖字段a，这就是传递依赖。解决：将一个实体信息的数据放在一个表内实现。

sql语句分类：

DDL：数据定义语言（create drop）

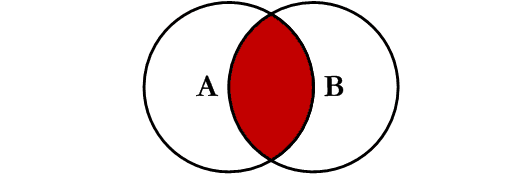
DML：数据操作语句（insert update delete）

DQL：数据查询语句（select ）

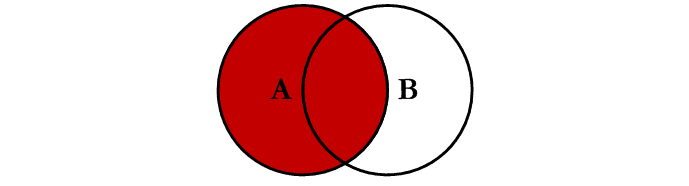
DCL：数据控制语句，进行授权和权限回收（grant revoke）

TPL：数据事务语句（commit collback savapoint）

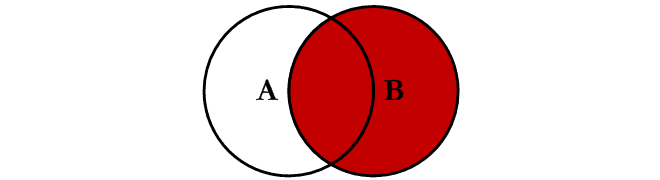
1. MySql各种join

INNER JOIN(内连接) 

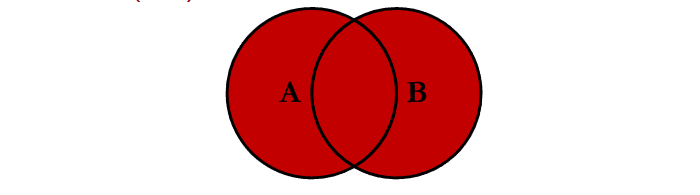
SELECT \* FORM TABLE\_A  A  INNER JOIN TABLE\_B B ON A.KEY = B.KEY

LEFT JOIN (左连接)

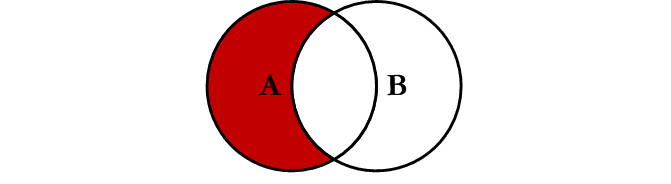
SELECT \* FORM TABLE\_A  A LEFT JOIN TABLE\_B B ON A.KEY = B.KEY

RIGHT JOIN (右连接)

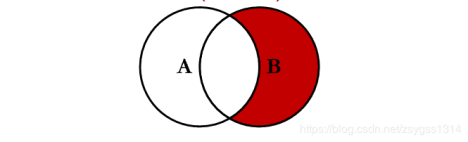
SELECT \* FROM TABLE\_A A RIGHT TABLE\_B B ON  A.KEY = B.KEY

OUTER JOIN (外连接)

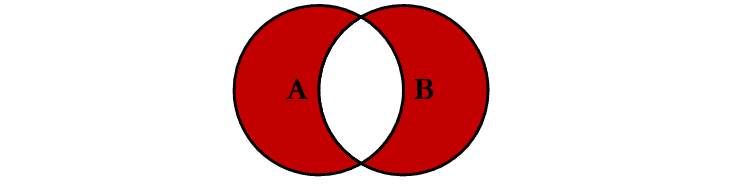
SELECT \* FROM TABLE\_A A FULL OUTER JOIN TABLE\_B B ON A.KEY = B.KEY

LEFT JOIN EXCLUDING INNER JOIN (左连接 - 内连接)

SELECT \* FROM TABLE\_A A LEFT JOIN TABLE\_B ON A.KEY = B.KEY WHERE B.KEY IS NULL

RIGHT JOIN EXCLUDING INNER JOIN (右连接 - 内连接)

SELECT \* FROM TABLE\_A A RIGHT JOIN TABLE\_B B ON A.KEY = B.KEY WHERE A.KEY IS NULL

OUTER JOIN EXCLUDING INNER JOIN (外连接 - 内连接)

SELECT \* FROM TABLE\_A FULL OUTER JOIN TABLE\_B B ON A.KEY = B.KEY WHERE A.KEY IS NULL OR B.KEY IS NULL

事务的十大特性ACID

原子性：不可分割的操作单元，事务中所有操作，要么全部成功；要么撤回到执行事务之前的状态

一致性：如果在执行事务之前数据库是一致的，那么在执行事务之后数据库也还是一致的；

隔离性：事务操作之间彼此独立和透明互不影响。事务独立运行。这通常使用锁来实现。一个事务处理后的结果，影响了其他事务，那么其他事务会撤回。事务的100%隔离，需要牺牲速度。

持久性：事务一旦提交，其结果就是永久的。即便发生系统故障，也能恢复。

事务隔离级别

未提交读(Read Uncommitted)：允许脏读，其他事务只要修改了数据，即使未提交，本事务也能看到修改后的数据值。也就是可能读取到其他会话中未提交事务修改的数据

提交读(Read Committed)：只能读取到已经提交的数据。Oracle等多数数据库默认都是该级别 (不重复读)。

可重复读(Repeated Read)：可重复读。无论其他事务是否修改并提交了数据，在这个事务中看到的数据值始终不受其他事务影响。

串行读(Serializable)：完全串行化的读，每次读都需要获得表级共享锁，读写相互都会阻塞

MySQL数据库(InnoDB引擎)默认使用可重复读（ Repeatable read)

索引相关

数据库索引，是数据库管理系统中一个排序的数据结构，以协助快速查询、更新数据库表中数据。索引的实现通常使用 B\_TREE。B\_TREE 索引加速了数据访问，因为存储引擎不会再去扫描整张表得到需要的数据；相反，它从根节点开始，根节点保存了子节点的指针，存储引擎会根据指针快速寻找数据。

MyISAM引擎使用B+Tree作为索引结构，叶节点的data域存放的是数据记录的地址，即：MyISAM索引文件和数据文件是分离的，MyISAM的索引文件仅仅保存数据记录的地址。MyISAM中索引检索的算法为首先按照B+Tree搜索算法搜索索引，如果指定的Key存在，则取出其data域的值，然后以data域的值为地址，读取相应数据记录。MyISAM的索引方式也叫做“非聚集”的。

InnoDB引擎也使用B+Tree作为索引结构，但是InnoDB的数据文件本身就是索引文件，叶节点data域保存了完整的数据记录。这个索引的key是数据表的主键，因此InnoDB表数据文件本身就是主索引。这种索引叫做“聚焦索引”。InnoDB的辅助索引的data域存储相应记录主键的值而不是地址。换句话说，InnoDB的所有辅助索引都引用主键作为data域。聚集索引这种实现方式使得按主键的搜索十分高效，但是辅助索引搜索需要检索两遍索引：首先检索辅助索引获得主键，然后用主键到主索引中检索获得记录。InnoDB的索引实现后，不建议使用过长的字段作为主键，因为所有辅助索引都引用主索引，过长的主索引会令辅助索引变得过大。在Innodb中也不建议使用非单调的字段作为主键，因为InnoDB数据文件本身是一颗B+Tree，非单调的主键会造成在插入新记录时数据文件为了维持B+Tree的特性而频繁的分裂调整，十分低效，建议使用自增字段作为主键。

MySQL数据库的四类索引:

　　index ---- 普通索引,数据可以重复，没有任何限制。

　　unique ---- 唯一索引,要求索引列的值必须唯一，但允许有空值；如果是组合索引，那么列值的组合必须唯一。

　　primary key ---- 主键索引,是一种特殊的唯一索引，一个表只能有一个主键，不允许有空值，一般是在创建表的同时创建主键索引。

　　组合索引 ---- 在多个字段上创建的索引，只有在查询条件中使用了创建索引时的第一个字段，索引才会被使用。

　　fulltext ---- 全文索引,是对于大表的文本域：char，varchar，text列才能创建全文索引，主要用于查找文本中的关键字，并不是直接与索引中的值进行比较。fulltext更像是一个搜索引擎，配合match against操作使用，而不是一般的where语句加like。

　　注:全文索引目前只有MyISAM存储引擎支持全文索引，InnoDB引擎5.6以下版本还不支持全文索引

　　所有存储引擎对每个表至少支持16个索引，总索引长度至少为256字节，索引有两种存储类型，包括B型树索引和哈希索引。

　　索引可以提高查询的速度，但是创建和维护索引需要耗费时间，同时也会影响插入的速度，如果需要插入大量的数据时，最好是先删除索引，插入数据后再建立索引。

索引生效条件

　　假设index（a,b,c）

最左前缀匹配：模糊查询时，使用%匹配时：’a%‘会使用索引，’%a‘不会使用索引

条件中有or，索引不会生效

a and c，a生效，c不生效

b and c，都不生效

a and b > 5 and c,a和b生效，c不生效。

检测索引的效果：

show status like '%handler\_read%' 越大越好

## **Mysql中的锁类型**

* MyISAM支持表锁，InnoDB支持表锁和行锁，默认为行锁
* 表级锁：开销小，加锁快，不会出现死锁。锁定粒度大，发生锁冲突的概率最高，并发量最低
* 行级锁：开销大，加锁慢，会出现死锁。锁力度小，发生锁冲突的概率小，并发度最高

存储过程

我们常用的操作数据库语言SQL语句在执行的时候需要要先编译，然后执行，而存储过程（Stored Procedure）是一组为了完成特定功能的SQL语句集，经编译后存储在数据库中，用户通过指定存储过程的名字并给定参数（如果该存储过程带有参数）来调用执行它。

一个存储过程是一个可编程的函数，它在数据库中创建并保存。它可以有SQL语句和一些特殊的控制结构组成。当希望在不同的应用程序或平台上执行相同的函数，或者封装特定功能时，存储过程是非常有用的。数据库中的存储过程可以看做是对编程中面向对象方法的模拟。它允许控制数据的访问方式。

优点：

(1).存储过程增强了SQL语言的功能和灵活性。存储过程可以用流控制语句编写，有很强的灵活性，可以完成复杂的判断和较复杂的运算。

(2).存储过程允许标准组件是编程。存储过程被创建后，可以在程序中被多次调用，而不必重新编写该存储过程的SQL语句。而且数据库专业人员可以随时对存储过程进行修改，对应用程序源代码毫无影响。

(3).存储过程能实现较快的执行速度。如果某一操作包含大量的Transaction-SQL代码或分别被多次执行，那么存储过程要比批处理的执行速度快很多。因为存储过程是预编译的。在首次运行一个存储过程时查询，优化器对其进行分析优化，并且给出最终被存储在系统表中的执行计划。而批处理的Transaction-SQL语句在每次运行时都要进行编译和优化，速度相对要慢一些。

(4).存储过程能过减少网络流量。针对同一个数据库对象的操作（如查询、修改），如果这一操作所涉及的Transaction-SQL语句被组织程存储过程，那么当在客户计算机上调用该存储过程时，网络中传送的只是该调用语句，从而大大增加了网络流量并降低了网络负载。

1. .存储过程可被作为一种安全机制来充分利用。系统管理员通过执行某一存储过程的权限进行限制，能够实现对相应的数据的访问权限的限制，避免了非授权用户对数据的访问，保证了数据的安全。

delete、drop、truncate区别

truncate 和 delete只删除数据，不删除表结构 ,drop删除表结构，并且释放所占的空间。

删除数据的速度，drop> truncate > delete

delete属于DML语言，需要事务管理，commit之后才能生效。drop和truncate属于DDL语言，操作立刻生效，不可回滚。

使用场合：

当你不再需要该表时， 用 drop;

当你仍要保留该表，但要删除所有记录时， 用 truncate;

当你要删除部分记录时（always with a where clause), 用 delete.

CHAR和VARCHAR的区别：

CHAR和VARCHAR类型在存储和检索方面有所不同

CHAR列长度固定为创建表时声明的长度，长度值范围是1到255

当CHAR值被存储时，它们被用空格填充到特定长度，检索CHAR值时需删除尾随空格。

## **存储引擎 MyISAM 和 InnoDB区别：**

1. InnoDB支持事务，MyISAM不支持。
2. MyISAM适合查询以及插入为主的应用，InnoDB适合频繁修改以及涉及到安全性较高的应用。
3. InnoDB支持外键，MyISAM不支持。
4. 从MySQL5.5.5以后，InnoDB是默认引擎。
5. MyISAM支持全文类型索引，而InnoDB不支持全文索引。
6. InnoDB中不保存表的总行数，select count(\*) from table时，InnoDB需要扫描整个表计算有多少行，但MyISAM只需简单读出保存好的总行数即可。注：当count(\*)语句包含where条件时MyISAM也需扫描整个表。
7. 对于自增长的字段，InnoDB中必须包含只有该字段的索引，但是在MyISAM表中可以和其他字段一起建立联合索引。
8. 清空整个表时，InnoDB是一行一行的删除，效率非常慢。MyISAM则会重建表。MyisAM使用delete语句删除后并不会立刻清理磁盘空间，需要定时清理，命令：OPTIMIZE table dept;
9. InnoDB支持行锁（某些情况下还是锁整表，如 update table set a=1 where user like ‘%lee%’）
10. Myisam创建表生成三个文件：.frm 数据表结构 、 .myd 数据文件 、 .myi 索引文件，Innodb只生成一个 .frm文件，数据存放在ibdata1.log
11. 现在一般都选用InnoDB，主要是MyISAM的全表锁，读写串行问题，并发效率锁表，效率低，MyISAM对于读写密集型应用一般是不会去选用的。
12. 应用场景：
    * MyISAM不支持事务处理等高级功能，但它提供高速存储和检索，以及全文搜索能力。如果应用中需要执行大量的SELECT查询，那么MyISAM是更好的选择。
    * InnoDB用于需要事务处理的应用程序，包括ACID事务支持。如果应用中需要执行大量的INSERT或UPDATE操作，则应该使用InnoDB，这样可以提高多用户并发操作的性能。

1. nosql-redis
2. nosql-Mongodb
3. nosql-HBase