1. **什么是SQL？**

结构化查询语言(Structured Query Language)简称SQL，是一种数据库查询语言。  
作用：用于存取数据、查询、更新和管理关系数据库系统。

1. **什么是MySQL?**

MySQL是一个关系型数据库管理系统，由瑞典MySQL AB 公司开发，属于 Oracle 旗下产品。MySQL 是最流行的关系型数据库管理系统之一，在 WEB 应用方面，MySQL是最好的 RDBMS (Relational Database Management System，关系数据库管理系统) 应用软件之一。在Java企业级开发中非常常用，因为 MySQL 是开源免费的，并且方便扩展。

1. **数据库三大范式是什么？**

* 第一范式：每个列都不可以再拆分。
* 第二范式：在第一范式的基础上，非主键列完全依赖于主键，而不能是依赖于主键的一部分。
* 第三范式：在第二范式的基础上，非主键列只依赖于主键，不依赖于其他非主键。
* 在设计数据库结构的时候，要尽量遵守三范式，如果不遵守，必须有足够的理由。比如性能。事实上我们经常会为了性能而妥协数据库的设计。

1. **mysql有关权限的表都有哪几个？**

MySQL服务器通过权限表来控制用户对数据库的访问，权限表存放在mysql数据库里，由mysql\_install\_db脚本初始化。这些权限表分别user，db，table\_priv，columns\_priv和host。下面分别介绍一下这些表的结构和内容：user权限表：记录允许连接到服务器的用户帐号信息，里面的权限是全局级的。db权限表：记录各个帐号在各个数据库上的操作权限。table\_priv权限表：记录数据表级的操作权限。columns\_priv权限表：记录数据列级的操作权限。host权限表：配合db权限表对给定主机上数据库级操作权限作更细致的控制。这个权限表不受GRANT和REVOKE语句的影响。

1. **MySQL的binlog有有几种录入格式？分别有什么区别？**

有三种格式，statement，row和mixed。statement模式下，每一条会修改数据的sql都会记录在binlog中。不需要记录每一行的变化，减少了binlog日志量，节约了IO，提高性能。由于sql的执行是有上下文的，因此在保存的时候需要保存相关的信息，同时还有一些使用了函数之类的语句无法被记录复制。row级别下，不记录sql语句上下文相关信息，仅保存哪条记录被修改。记录单元为每一行的改动，基本是可以全部记下来但是由于很多操作，会导致大量行的改动(比如alter table)，因此这种模式的文件保存的信息太多，日志量太大。mixed，一种折中的方案，普通操作使用statement记录，当无法使用statement的时候使用row。  
此外，新版的MySQL中对row级别也做了一些优化，当表结构发生变化的时候，会记录语句而不是逐行记录。

1. **mysql有哪些数据类型？**

* 1、整数类型，包括TINYINT、SMALLINT、MEDIUMINT、INT、BIGINT，分别表示1字节、2字节、3字节、4字节、8字节整数。任何整数类型都可以加上UNSIGNED属性，表示数据是无符号的，即非负整数。长度：整数类型可以被指定长度，例如：INT(11)表示长度为11的INT类型。长度在大多数场景是没有意义的，它不会限制值的合法范围，只会影响显示字符的个数，而且需要和UNSIGNED ZEROFILL属性配合使用才有意义。例子：假定类型设定为INT(5)，属性为UNSIGNED ZEROFILL，如果用户插入的数据为12的话，那么数据库实际存储数据为00012。
* 2、实数类型，包括FLOAT、DOUBLE、DECIMAL。DECIMAL可以用于存储比BIGINT还大的整型，能存储精确的小数。而FLOAT和DOUBLE是有取值范围的，并支持使用标准的浮点进行近似计算。计算时FLOAT和DOUBLE相比DECIMAL效率更高一些，DECIMAL你可以理解成是用字符串进行处理。
* 3、字符串类型，包括VARCHAR、CHAR、TEXT、BLOBVARCHAR用于存储可变长字符串，它比定长类型更节省空间。VARCHAR使用额外1或2个字节存储字符串长度。列长度小于255字节时，使用1字节表示，否则使用2字节表示。VARCHAR存储的内容超出设置的长度时，内容会被截断。CHAR是定长的，根据定义的字符串长度分配足够的空间。CHAR会根据需要使用空格进行填充方便比较。CHAR适合存储很短的字符串，或者所有值都接近同一个长度。CHAR存储的内容超出设置的长度时，内容同样会被截断。
* 4、枚举类型（ENUM），把不重复的数据存储为一个预定义的集合。有时可以使用ENUM代替常用的字符串类型。ENUM存储非常紧凑，会把列表值压缩到一个或两个字节。ENUM在内部存储时，其实存的是整数。尽量避免使用数字作为ENUM枚举的常量，因为容易混乱。排序是按照内部存储的整数
* 5、日期和时间类型，尽量使用timestamp，空间效率高于datetime，用整数保存时间戳通常不方便处理。如果需要存储微妙，可以使用bigint存储。看到这里，这道真题是不是就比较容易回答了。

1. **MyISAM索引与InnoDB索引的区别？**

* InnoDB索引是聚簇索引，MyISAM索引是非聚簇索引。
* InnoDB的主键索引的叶子节点存储着行数据，因此主键索引非常高效。
* MyISAM索引的叶子节点存储的是行数据地址，需要再寻址一次才能得到数据。
* InnoDB非主键索引的叶子节点存储的是主键和其他带索引的列数据，因此查询时做到覆盖索引会非常高效。

1. **InnoDB引擎的4大特性**

* 插入缓冲（insert buffer)
* 二次写(double write)
* 自适应哈希索引(ahi)
* 预读(read ahead)

1. **什么是索引？**

* 索引是一种特殊的文件(InnoDB数据表上的索引是表空间的一个组成部分)，它们包含着对数据表里所有记录的引用指针。
* 索引是一种数据结构。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数及其变种B+数。
* 更通俗的说，索引就相当于目录。为了方便查找书中的内容，通过对内容建立索引形成目录。索引是一个文件，它是要占据物理空间的。

1. **索引有哪些优缺点？**

* 索引的优点：可以大大加快数据的检索速度，这也是创建索引的最主要的原因。通过使用索引，可以在查询的过程中，使用优化隐藏器，提高系统的性能。
* 索引的缺点：时间方面：创建索引和维护索引要耗费时间，具体地，当对表中的数据进行增加、删除和修改的时候，索引也要动态的维护，会降低增/改/删的执行效率；空间方面：索引需要占物理空间。

1. **索引有哪几种类型？**

* 主键索引: 数据列不允许重复，不允许为NULL，一个表只能有一个主键。
* 唯一索引: 数据列不允许重复，允许为NULL值，一个表允许多个列创建唯一索引。可以通过 ALTER TABLE table\_name ADD UNIQUE (column); 创建唯一索引可以通过 ALTER TABLE table\_name ADD UNIQUE (column1,column2); 创建唯一组合索引
* 普通索引: 基本的索引类型，没有唯一性的限制，允许为NULL值。可以通过ALTER TABLE table\_name ADD INDEX index\_name (column);创建普通索引可以通过ALTER TABLE table\_name ADD INDEX index\_name(column1, column2, column3);创建组合索引。
* 全文索引： 是目前搜索引擎使用的一种关键技术。可以通过ALTER TABLE table\_name ADD FULLTEXT (column);创建全文索引

1. **索引的数据结构（b树，hash）**

索引的数据结构和具体存储引擎的实现有关，在MySQL中使用较多的索引有Hash索引，B+树索引等，而我们经常使用的InnoDB存储引擎的默认索引实现为：B+树索引。对于哈希索引来说，底层的数据结构就是哈希表，因此在绝大多数需求为单条记录查询的时候，可以选择哈希索引，查询性能最快；其余大部分场景，建议选择BTree索引。

* **1. B树索引**

mysql通过存储引擎取数据，基本上90%的人用的就是InnoDB了，按照实现方式分，InnoDB的索引类型目前只有两种：BTREE（B树）索引和HASH索引。B树索引是Mysql数据库中使用最频繁的索引类型，基本所有存储引擎都支持BTree索引。通常我们说的索引不出意外指的就是（B树）索引（实际是用B+树实现的，因为在查看表索引时，mysql一律打印BTREE，所以简称为B树索引）

* **2. B+tree性质**

n棵子tree的节点包含n个关键字，不用来保存数据而是保存数据的索引。  
所有的叶子结点中包含了全部关键字的信息，即指向含这些关键字记录的指针，且叶子结点本身依关键字的大小自小而大顺序链接。  
所有的非终端结点可以看成是索引部分，结点中仅含其子数中的最大（或最小）关键字。  
B+ 树中，数据对象的插入和删除仅在叶节点上进行。  
B+树有2个头指针，一个是树的根节点，一个是最小关键码的叶节点。

* **3. 哈希索引**

简要说下，类似于数据结构中简单实现的HASH表（散列表）一样，当我们在mysql中用哈希索引时，主要就是通过Hash算法（常见的Hash算法有直接定址法、平方取中法、折叠法、除数取余法、随机数法），将数据库字段数据转换成定长的Hash值，与这条数据的行指针一并存入Hash表的对应位置；如果发生Hash碰撞（两个不同关键字的Hash值相同），则在对应Hash键下以链表形式存储。当然这只是简略模拟图。

**索引的基本原理**

索引用来快速地寻找那些具有特定值的记录。如果没有索引，一般来说执行查询时遍历整张表。  
索引的原理很简单，就是把无序的数据变成有序的查询把创建了索引的列的内容进行排序对排序结果生成倒排表在倒排表内容上拼上数据地址链在查询的时候，先拿到倒排表内容，再取出数据地址链，从而拿到具体数据

**索引算法有哪些？**

索引算法有 BTree算法和Hash算法

* **1. BTree算法**

BTree是最常用的mysql数据库索引算法，也是mysql默认的算法。因为它不仅可以被用在=,>,>=,<,<=和between这些比较操作符上，而且还可以用于like操作符，只要它的查询条件是一个不以通配符开头的常量。

* **2. Hash算法**

Hash Hash索引只能用于对等比较，例如=,<=>（相当于=）操作符。由于是一次定位数据，不像BTree索引需要从根节点到枝节点，最后才能访问到页节点这样多次IO访问，所以检索效率远高于BTree索引。

1. **索引设计的原则？**

* 适合索引的列是出现在where子句中的列，或者连接子句中指定的列。
* 基数较小的类，索引效果较差，没有必要在此列建立索引
* 使用短索引，如果对长字符串列进行索引，应该指定一个前缀长度，这样能够节省大量索引空间
* 不要过度索引。索引需要额外的磁盘空间，并降低写操作的性能。在修改表内容的时候，索引会进行更新甚至重构，索引列越多，这个时间就会越长。所以只保持需要的索引有利于查询即可。

1. **创建索引的原则**

索引虽好，但也不是无限制的使用，最好符合一下几个原则最左前缀匹配原则，组合索引非常重要的原则，mysql会一直向右匹配直到遇到范围查询(>、<、between、like)就停止匹配，比如a = 1 and b = 2 and c > 3 and d = 4 如果建立(a,b,c,d)顺序的索引，d是用不到索引的，如果建立(a,b,d,c)的索引则都可以用到，a,b,d的顺序可以任意调整。较频繁作为查询条件的字段才去创建索引更新频繁字段不适合创建索引若是不能有效区分数据的列不适合做索引列(如性别，男女未知，最多也就三种，区分度实在太低)尽量的扩展索引，不要新建索引。比如表中已经有a的索引，现在要加(a,b)的索引，那么只需要修改原来的索引即可。定义有外键的数据列一定要建立索引。对于那些查询中很少涉及的列，重复值比较多的列不要建立索引。对于定义为text、image和bit的数据类型的列不要建立索引。

1. **创建索引时需要注意什么？**

* 非空字段：应该指定列为NOT NULL，除非你想存储NULL。在mysql中，含有空值的列很难进行查询优化，因为它们使得索引、索引的统计信息以及比较运算更加复杂。你应该用0、一个特殊的值或者一个空串代替空值；
* 取值离散大的字段：（变量各个取值之间的差异程度）的列放到联合索引的前面，可以通过count()函数查看字段的差异值，返回值越大说明字段的唯一值越多字段的离散程度高；
* 索引字段越小越好：数据库的数据存储以页为单位一页存储的数据越多一次IO操作获取的数据越大效率越高。

1. **使用索引查询一定能提高查询的性能吗？**

通常，通过索引查询数据比全表扫描要快。但是我们也必须注意到它的代价。  
索引需要空间来存储，也需要定期维护， 每当有记录在表中增减或索引列被修改时，索引本身也会被修改。 这意味着每条记录的INSERT，DELETE，UPDATE将为此多付出4，5 次的磁盘I/O。 因为索引需要额外的存储空间和处理，那些不必要的索引反而会使查询反应时间变慢。使用索引查询不一定能提高查询性能，索引范围查询(INDEX RANGE SCAN)适用于两种情况:

* 基于一个范围的检索，一般查询返回结果集小于表中记录数的30%
* 基于非唯一性索引的检索

1. **百万级别或以上的数据如何删除？**

关于索引：由于索引需要额外的维护成本，因为索引文件是单独存在的文件,所以当我们对数据的增加,修改,删除,都会产生额外的对索引文件的操作,这些操作需要消耗额外的IO,会降低增/改/删的执行效率。所以，在我们删除数据库百万级别数据的时候，查询MySQL官方手册得知删除数据的速度和创建的索引数量是成正比的。所以我们想要删除百万数据的时候可以先删除索引（此时大概耗时三分多钟）然后删除其中无用数据（此过程需要不到两分钟）删除完成后重新创建索引(此时数据较少了)创建索引也非常快，约十分钟左右。与之前的直接删除绝对是要快速很多，更别说万一删除中断,一切删除会回滚。那更是坑了。

1. **什么是最左前缀原则？什么是最左匹配原则？**

顾名思义，就是最左优先，在创建多列索引时，要根据业务需求，where子句中使用最频繁的一列放在最左边。  
最左前缀匹配原则，非常重要的原则，mysql会一直向右匹配直到遇到范围查询(>、<、between、like)就停止匹配，比如a = 1 and b = 2 and c > 3 and d = 4 如果建立(a,b,c,d)顺序的索引，d是用不到索引的，如果建立(a,b,d,c)的索引则都可以用到，a,b,d的顺序可以任意调整。  
=和in可以乱序，比如a = 1 and b = 2 and c = 3 建立(a,b,c)索引可以任意顺序，mysql的查询优化器会帮你优化成索引可以识别的形式

1. **B树和B+树的区别**

在B树中，你可以将键和值存放在内部节点和叶子节点；但在B+树中，内部节点都是键，没有值，叶子节点同时存放键和值。  
B+树的叶子节点有一条链相连，而B树的叶子节点各自独立。

**使用B树的好处**

B树可以在内部节点同时存储键和值，因此，把频繁访问的数据放在靠近根节点的地方将会大大提高热点数据的查询效率。这种特性使得B树在特定数据重复多次查询的场景中更加高效。

**使用B+树的好处**

由于B+树的内部节点只存放键，不存放值，因此，一次读取，可以在内存页中获取更多的键，有利于更快地缩小查找范围。 B+树的叶节点由一条链相连，因此，当需要进行一次全数据遍历的时候，B+树只需要使用O(logN)时间找到最小的一个节点，然后通过链进行O(N)的顺序遍历即可。而B树则需要对树的每一层进行遍历，这会需要更多的内存置换次数，因此也就需要花费更多的时间

1. **什么是聚簇索引？何时使用聚簇索引与非聚簇索引？**

聚簇索引：将数据存储与索引放到了一块，找到索引也就找到了数据  
非聚簇索引：将数据存储于索引分开结构，索引结构的叶子节点指向了数据的对应行，myisam通过key\_buffer把索引先缓存到内存中，当需要访问数据时（通过索引访问数据），在内存中直接搜索索引，然后通过索引找到磁盘相应数据，这也就是为什么索引不在key buffer命中时，速度慢的原因。

1. **非聚簇索引一定会回表查询吗？**

不一定，这涉及到查询语句所要求的字段是否全部命中了索引，如果全部命中了索引，那么就不必再进行回表查询。  
举个简单的例子，假设我们在员工表的年龄上建立了索引，那么当进行select age from employee where age < 20的查询时，在索引的叶子节点上，已经包含了age信息，不会再次进行回表查询。

1. **联合索引是什么？为什么需要注意联合索引中的顺序？**

MySQL可以使用多个字段同时建立一个索引，叫做联合索引。在联合索引中，如果想要命中索引，需要按照建立索引时的字段顺序挨个使用，否则无法命中索引。  
MySQL使用索引时需要索引有序，假设现在建立了"name，age，school"的联合索引，那么索引的排序为: 先按照name排序，如果name相同，则按照age排序，如果age的值也相等，则按照school进行排序。  
当进行查询时，此时索引仅仅按照name严格有序，因此必须首先使用name字段进行等值查询，之后对于匹配到的列而言，其按照age字段严格有序，此时可以使用age字段用做索引查找，以此类推。因此在建立联合索引的时候应该注意索引列的顺序，一般情况下，将查询需求频繁或者字段选择性高的列放在前面。此外可以根据特例的查询或者表结构进行单独的调整。

1. **什么是数据库事务？**

事务是一个不可分割的数据库操作序列，也是数据库并发控制的基本单位，其执行的结果必须使数据库从一种一致性状态变到另一种一致性状态。事务是逻辑上的一组操作，要么都执行，要么都不执行。

**事物的四大特性(ACID)介绍一下?**

* 原子性： 事务是最小的执行单位，不允许分割。事务的原子性确保动作要么全部完成，要么完全不起作用；
* 一致性： 执行事务前后，数据保持一致，多个事务对同一个数据读取的结果是相同的；
* 隔离性： 并发访问数据库时，一个用户的事务不被其他事务所干扰，各并发事务之间数据库是独立的；
* 持久性： 一个事务被提交之后。它对数据库中数据的改变是持久的，即使数据库发生故障也不应该对其有任何影响。

1. **什么是脏读？幻读？不可重复读？**

脏读(Drity Read)：某个事务已更新一份数据，另一个事务在此时读取了同一份数据，由于某些原因，前一个RollBack了操作，则后一个事务所读取的数据就会是不正确的。  
不可重复读(Non-repeatable read):在一个事务的两次查询之中数据不一致，这可能是两次查询过程中间插入了一个事务更新的原有的数据。  
幻读(Phantom Read):在一个事务的两次查询中数据笔数不一致，例如有一个事务查询了几列(Row)数据，而另一个事务却在此时插入了新的几列数据，先前的事务在接下来的查询中，就会发现有几列数据是它先前所没有的。

1. **什么是事务的隔离级别？MySQL的默认隔离级别是什么？**

为了达到事务的四大特性，数据库定义了4种不同的事务隔离级别，由低到高依次为Read uncommitted、Read committed、Repeatable read、Serializable，这四个级别可以逐个解决脏读、不可重复读、幻读这几类问题。  
SQL 标准定义了四个隔离级别：READ-UNCOMMITTED(读取未提交)： 最低的隔离级别，允许读取尚未提交的数据变更，可能会导致脏读、幻读或不可重复读。READ-COMMITTED(读取已提交)： 允许读取并发事务已经提交的数据，可以阻止脏读，但是幻读或不可重复读仍有可能发生。REPEATABLE-READ(可重复读)： 对同一字段的多次读取结果都是一致的，除非数据是被本身事务自己所修改，可以阻止脏读和不可重复读，但幻读仍有可能发生。SERIALIZABLE(可串行化)： 最高的隔离级别，完全服从ACID的隔离级别。所有的事务依次逐个执行，这样事务之间就完全不可能产生干扰，也就是说，该级别可以防止脏读、不可重复读以及幻读。  
Mysql 默认采用的 REPEATABLE\_READ隔离级别 Oracle 默认采用的 READ\_COMMITTED隔离级别

**隔离级别与锁的关系**

* 在Read Uncommitted级别下，读取数据不需要加共享锁，这样就不会跟被修改的数据上的排他锁冲突
* 在Read Committed级别下，读操作需要加共享锁，但是在语句执行完以后释放共享锁；
* 在Repeatable Read级别下，读操作需要加共享锁，但是在事务提交之前并不释放共享锁，也就是必须等待事务执行完毕以后才释放共享锁。
* SERIALIZABLE 是限制性最强的隔离级别，因为该级别锁定整个范围的键，并一直持有锁，直到事务完成。

1. **按照锁的粒度分数据库锁有哪些？**

* 行级锁:行级锁是Mysql中锁定粒度最细的一种锁，表示只针对当前操作的行进行加锁。行级锁能大大减少数据库操作的冲突。其加锁粒度最小，但加锁的开销也最大。行级锁分为共享锁 和 排他锁。特点：开销大，加锁慢；会出现死锁；锁定粒度最小，发生锁冲突的概率最低，并发度也最高。
* 表级锁: 表级锁是MySQL中锁定粒度最大的一种锁，表示对当前操作的整张表加锁，它实现简单，资源消耗较少，被大部分MySQL引擎支持。最常使用的MYISAM与INNODB都支持表级锁定。表级锁定分为表共享读锁（共享锁）与表独占写锁（排他锁）。特点：开销小，加锁快；不会出现死锁；锁定粒度大，发出锁冲突的概率最高，并发度最低。
* 页级锁:页级锁是MySQL中锁定粒度介于行级锁和表级锁中间的一种锁。表级锁速度快，但冲突多，行级冲突少，但速度慢。所以取了折衷的页级，一次锁定相邻的一组记录。

1. **从锁的类别上分MySQL都有哪些锁呢？**

从锁的类别上来讲，有共享锁和排他锁。共享锁: 又叫做读锁。 当用户要进行数据的读取时，对数据加上共享锁。共享锁可以同时加上多个。排他锁: 又叫做写锁。 当用户要进行数据的写入时，对数据加上排他锁。排他锁只可以加一个，他和其他的排他锁，共享锁都相斥。

1. **InnoDB存储引擎的锁的算法有哪三种？**

* Record lock：单个行记录上的锁
* Gap lock：间隙锁，锁定一个范围，不包括记录本身
* Next-key lock：record+gap 锁定一个范围，包含记录本身

1. **什么是死锁？怎么解决？**

死锁是指两个或多个事务在同一资源上相互占用，并请求锁定对方的资源，从而导致恶性循环的现象。  
常见的解决死锁的方法如果不同程序会并发存取多个表，尽量约定以相同的顺序访问表，可以大大降低死锁机会。在同一个事务中，尽可能做到一次锁定所需要的所有资源，减少死锁产生概率；对于非常容易产生死锁的业务部分，可以尝试使用升级锁定颗粒度，通过表级锁定来减少死锁产生的概率；  
如果业务处理不好可以用分布式事务锁或者使用乐观锁

1. **数据库的乐观锁和悲观锁是什么？怎么实现的？**

数据库管理系统（DBMS）中的并发控制的任务是确保在多个事务同时存取数据库中同一数据时不破坏事务的隔离性和统一性以及数据库的统一性。乐观并发控制（乐观锁）和悲观并发控制（悲观锁）是并发控制主要采用的技术手段。  
悲观说：假定会发生并发冲突，屏蔽一切可能违反数据完整性的操作。在查询完数据的时候就把事务锁起来，直到提交事务。实现方式：使用数据库中的锁机制  
乐观锁：假设不会发生并发冲突，只在提交操作时检查是否违反数据完整性。在修改数据的时候把事务锁起来，通过version的方式来进行锁定。实现方式：乐一般会使用版本号机制或CAS算法实现。

1. **大表数据查询，怎么优化？**

* 优化shema、sql语句+索引；
* 第二加缓存，memcached, redis；
* 主从复制，读写分离；
* 垂直拆分，根据你模块的耦合度，将一个大的系统分为多个小的系统，也就是分布式系统
* 水平切分，针对数据量大的表，这一步最麻烦，最能考验技术水平，要选择一个合理的sharding key, 为了有好的查询效率，表结构也要改动，做一定的冗余，应用也要改，sql中尽量带sharding key，将数据定位到限定的表上去查，而不是扫描全部的表

1. **超大分页怎么处理？**

超大的分页一般从两个方向上来解决:数据库层面,这也是我们主要集中关注的(虽然收效没那么大),类似于select \* from table where age > 20 limit 1000000,10这种查询其实也是有可以优化的余地的. 这条语句需要load1000000数据然后基本上全部丢弃,只取10条当然比较慢. 当时我们可以修改为select \* from table where id in (select id from table where age > 20 limit 1000000,10).这样虽然也load了一百万的数据,但是由于索引覆盖,要查询的所有字段都在索引中,所以速度会很快. 同时如果ID连续的好,我们还可以select \* from table where id > 1000000 limit 10,效率也是不错的,优化的可能性有许多种,但是核心思想都一样,就是减少load的数据从需求的角度减少这种请求…主要是不做类似的需求(直接跳转到几百万页之后的具体某一页.只允许逐页查看或者按照给定的路线走,这样可预测,可缓存)以及防止ID泄漏且连续被人恶意攻击

1. **为什么要尽量设定一个主键？**

主键是数据库确保数据行在整张表唯一性的保障，即使业务上本张表没有主键，也建议添加一个自增长的ID列作为主键。设定了主键之后，在后续的删改查的时候可能更加快速以及确保操作数据范围安全。

1. **主键使用自增ID还是UUID？**

推荐使用自增ID，不要使用UUID。  
因为在InnoDB存储引擎中，主键索引是作为聚簇索引存在的，也就是说，主键索引的B+树叶子节点上存储了主键索引以及全部的数据(按照顺序)，如果主键索引是自增ID，那么只需要不断向后排列即可，如果是UUID，由于到来的ID与原来的大小不确定，会造成非常多的数据插入，数据移动，然后导致产生很多的内存碎片，进而造成插入性能的下降。  
总之，在数据量大一些的情况下，用自增主键性能会好一些。  
关于主键是聚簇索引，如果没有主键，InnoDB会选择一个唯一键来作为聚簇索引，如果没有唯一键，会生成一个隐式的主键。

1. **字段为什么要求定义为not null？**

null值会占用更多的字节，且会在程序中造成很多与预期不符的情况。  
如果要存储用户的密码散列，应该使用什么字段进行存储？  
密码散列，盐，用户身份证号等固定长度的字符串应该使用char而不是varchar来存储，这样可以节省空间且提高检索效率。

1. **数据库结构优化？**

一个好的数据库设计方案对于数据库的性能往往会起到事半功倍的效果。  
需要考虑数据冗余、查询和更新的速度、字段的数据类型是否合理等多方面的内容。  
将字段很多的表分解成多个表：对于字段较多的表，如果有些字段的使用频率很低，可以将这些字段分离出来形成新表。因为当一个表的数据量很大时，会由于使用频率低的字段的存在而变慢。  
增加中间表：对于需要经常联合查询的表，可以建立中间表以提高查询效率。通过建立中间表，将需要通过联合查询的数据插入到中间表中，然后将原来的联合查询改为对中间表的查询。  
增加冗余字段：设计数据表时应尽量遵循范式理论的规约，尽可能的减少冗余字段，让数据库设计看起来精致、优雅。但是，合理的加入冗余字段可以提高查询速度。表的规范化程度越高，表和表之间的关系越多，需要连接查询的情况也就越多，性能也就越差。

1. **MySQL数据库cpu飙升到500%的话他怎么处理？**

当 cpu 飙升到 500%时，先用操作系统命令 top 命令观察是不是 mysqld 占用导致的，如果不是，找出占用高的进程，并进行相关处理。  
如果是 mysqld 造成的， show processlist，看看里面跑的 session 情况，是不是有消耗资源的 sql 在运行。找出消耗高的 sql，看看执行计划是否准确， index 是否缺失，或者实在是数据量太大造成。  
一般来说，肯定要 kill 掉这些线程(同时观察 cpu 使用率是否下降)，等进行相应的调整(比如说加索引、改 sql、改内存参数)之后，再重新跑这些 SQL。  
也有可能是每个 sql 消耗资源并不多，但是突然之间，有大量的 session 连进来导致 cpu 飙升，这种情况就需要跟应用一起来分析为何连接数会激增，再做出相应的调整，比如说限制连接数等。

1. **主从复制的作用？**

* 主数据库出现问题，可以切换到从数据库。
* 可以进行数据库层面的读写分离。
* 可以在从数据库上进行日常备份。

1. **MySQL主从复制解决的问题？**

* 数据分布：随意开始或停止复制，并在不同地理位置分布数据备份
* 负载均衡：降低单个服务器的压力
* 高可用和故障切换：帮助应用程序避免单点失败
* 升级测试：可以用更高版本的MySQL作为从库

1. **MySQL主从复制工作原理？**

* 在主库上把数据更高记录到二进制日志
* 从库将主库的日志复制到自己的中继日志
* 从库读取中继日志的事件，将其重放到从库数据中。