

# JAVA 面试题总览

注：此文档为群主将群共享里的所有面试题以及群友面试遇到的题目进行了整理(去掉了重复的和过时的并随时更新)。

更新日期 2017-4-27，欢迎加入群 329019348

## JAVA 基础

1. JAVA 中的几种基本数据类型是什么，各自占用多少字节。
2. String 类能被继承吗，为什么。
3. String, StringBuffer, StringBuilder 的区别。
4. ArrayList 和 LinkedList 有什么区别。
5. 讲讲类的实例化顺序，比如父类静态数据，构造函数，字段，子类静态数据，构造函数，字段，当 new 的时候，他们的执行顺序。
6. 用过哪些 Map 类，都有什么区别，HashMap 是线程安全的吗，并发下使用的 Map 是什么，他们内部原理分别是什么，比如存储方式，hashcode，扩容，默认容量等。
7. JAVA8 的 ConcurrentHashMap 为什么放弃了分段锁，有什么问题吗，如果你来设计，你如何设计。
8. 有没有有顺序的 Map 实现类，如果有，他们是怎么保证有序的。
9. 抽象类和接口的区别，类可以继承多个类么，接口可以继承多个接口么，类可以实现多个接口么。
10. 继承和聚合的区别在哪。
11. IO 模型有哪些，讲讲你理解的 nio，他和 bio 的区别是啥，谈谈 reactor 模型。
12. 反射的原理，反射创建类实例的三种方式是什么。
13. 反射中，Class.forName 和 ClassLoader 区别。
14. 描述动态代理的几种实现方式，分别说出相应的优缺点。
15. 动态代理与 cglib 实现的区别。
16. 为什么 CGlib 方式可以对接口实现代理。
17. final 的用途。
18. 写出三种单例模式实现。
19. 如何在父类中为子类自动完成所有的 hashcode 和 equals 实现？这么做有何优劣。
20. 请结合 OO 设计理念，谈谈访问修饰符 public、private、protected、default 在应用设计中的作用。
21. 深拷贝和浅拷贝区别。
22. 数组和链表数据结构描述，各自的时间复杂度。
23. error 和 exception 的区别，CheckedException, RuntimeException 的区别。
24. 请列出 5 个运行时异常。

25. 在自己的代码中，如果创建一个 `java.lang.String` 类，这个类是否可以被类加载器加载？为什么。
26. 说一说你对 `java.lang.Object` 对象中 `hashCode` 和 `equals` 方法的理解。在什么场景下需要重新实现这两个方法。
27. 在 `jdk1.5` 中，引入了泛型，泛型的存在是用来解决什么问题。
28. 这样的 `a.hashCode()` 有什么用，与 `a.equals(b)` 有什么关系。
29. 有没有可能 2 个不相等的对象有相同的 `hashCode`。
30. `Java` 中的 `HashSet` 内部是如何工作的。
31. 什么是序列化，怎么序列化，为什么序列化，反序列化会遇到什么问题，如何解决。
32. `java8` 的新特性。

## JVM 知识

1. 什么情况下会发生栈内存溢出。
2. `JVM` 的内存结构，`Eden` 和 `Survivor` 比例。
3. `JVM` 内存为什么要分成新生代，老年代，持久代。新生代中为什么要分为 `Eden` 和 `Survivor`。
4. `JVM` 中一次完整的 `GC` 流程是怎样的，对象如何晋升到老年代，说说你知道的几种主要的 `JVM` 参数。
5. 你知道哪几种垃圾收集器，各自的优缺点，重点讲下 `cms` 和 `G1`，包括原理，流程，优缺点。
6. 垃圾回收算法的实现原理。
7. 当出现了内存溢出，你怎么排错。
8. `JVM` 内存模型的相关知识了解多少，比如重排序，内存屏障，`happen-before`，主内存，工作内存等。
9. 简单说说你了解的类加载器，可以打破双亲委派么，怎么打破。
10. 讲讲 `JAVA` 的反射机制。
11. 你们线上应用的 `JVM` 参数有哪些。
12. `g1` 和 `cms` 区别，吞吐量优先和响应优先的垃圾收集器选择。
13. 怎么打出线程栈信息。
14. 请解释如下 `jvm` 参数的含义：

```
-server -Xms512m -Xmx512m -Xss1024K  
-XX:PermSize=256m -XX:MaxPermSize=512m -XX:MaxTenuringThreshold=20  
XX:CMSInitiatingOccupancyFraction=80 -XX:+UseCMSInitiatingOccupancyOnly。
```

## 开源框架知识

1. 简单讲讲 `tomcat` 结构，以及其类加载器流程。
2. `tomcat` 如何调优，涉及哪些参数。
3. 讲讲 `Spring` 加载流程。

4. 讲讲 Spring 事务的传播属性。
5. Spring 如何管理事务的。
6. Spring 怎么配置事务（具体说出一些关键的 xml 元素）。
7. 说说你对 Spring 的理解，非单例注入的原理？它的生命周期？循环注入的原理，aop 的实现原理，说说 aop 中的几个术语，它们是怎么相互工作的。
8. Springmvc 中 DispatcherServlet 初始化过程。
9. netty 的线程模型 netty 的线程模型 netty 如何基于 reactor 模型上实现的。
10. netty 的 fashwheeltimer 的用法，实现原理，是否出现过调用不够准时，怎么解决。
11. netty 的心跳处理在弱网下怎么办。
12. netty 的通讯协议是什么样的。

## 操作系统

1. Linux 系统下你关注过哪些内核参数，说说你知道的。
2. Linux 下 IO 模型有几种，各自的含义是什么。
3. epoll 和 poll 有什么区别。
4. 平时用到哪些 Linux 命令。
5. 用一行命令查看文件的最后五行。
6. 用一行命令输出正在运行的 java 进程。
7. 介绍下你理解的操作系统中线程切换过程。
8. 进程和线程的区别。
9. top 命令之后有哪些内容，有什么作用。

## 多线程

1. 多线程的几种实现方式，什么是线程安全。
2. volatile 的原理，作用，能代替锁么。
3. 画一个线程的生命周期状态图。
4. sleep 和 wait 的区别。
5. Lock 与 Synchronized 的区别。
6. synchronized 的原理是什么，一般用在什么地方(比如加在静态方法和非静态方法的区别，静态方法和非静态方法同时执行的时候会有影响吗)解释以下名词：重排序，自旋锁，偏向锁，轻量级锁，可重入锁，公平锁，非公平锁，乐观锁，悲观锁。
7. 用过哪些原子类，他们的原理是什么。
8. JUC 下研究过哪些并发工具，讲讲原理。
9. 用过线程池吗，如果用过，请说明原理，并说说 newCache 和 newFixed 有什么区别，构造函数的各个参数的含义是什么，比如 coreSize, maxsize 等。
10. 线程池的关闭方式有几种，各自的区别是什么。

11. 假如有一个第三方接口，有很多个线程去调用获取数据，现在规定每秒钟最多有 10 个线程同时调用它，如何做到。
12. spring 的 controller 是单例还是多例，怎么保证并发的安全。
13. 用三个线程按顺序循环打印 abc 三个字母，比如 abcabcabc。
14. ThreadLocal 用过么，用途是什么，原理是什么，用的时候要注意什么。
15. 如果让你实现一个并发安全的链表，你会怎么做。
16. 有哪些无锁数据结构，他们实现的原理是什么。
17. 讲讲 java 同步机制的 wait 和 notify。
18. 多线程如果线程挂住了怎么办。
19. countdowlatch 和 cyclicbarrier 的内部原理和用法，以及相互之间的差别(比如 countdownlatch 的 await 方法和是怎么实现的)。
20. 对 AbstractQueuedSynchronizer 了解多少，讲讲加锁和解锁的流程，独占锁和公平锁有什么不同。
21. 使用 synchronized 修饰静态方法和非静态方法有什么区别。
22. 简述 ConcurrentLinkedQueue 和 LinkedBlockingQueue 的用处和不同之处。
23. 导致线程死锁的原因？怎么解除线程死锁。
24. 非常多个线程（可能是不同机器），相互之间需要等待协调，才能完成某种工作，问怎么设计这种协调方案。
25. 用过读写锁吗，原理是什么，一般在什么场景下用。
26. 开启多个线程，如果保证顺序执行，有哪几种实现方式，或者如何保证多个线程都执行完再拿到结果。
27. 延迟队列的实现方式，delayQueue 和时间轮算法的异同。

## TCP 与 HTTP

1. http1.0 和 http1.1 有什么区别。
2. TCP 三次握手和四次挥手的流程，为什么断开连接要 4 次，如果握手只有两次，会出现什么。
3. TIME\_WAIT 和 CLOSE\_WAIT 的区别。
4. 说说你知道的几种 HTTP 响应码，比如 200, 302, 404。
5. 当你用浏览器打开一个链接的时候，计算机做了哪些工作步骤。
6. TCP/IP 如何保证可靠性，说说 TCP 头的结构。
7. 如何避免浏览器缓存。
8. 简述 Http 请求 get 和 post 的区别以及数据包格式。
9. 简述 HTTP 请求的报文格式。
10. HTTPS 的加密方式是什么，讲讲整个加密解密流程。

# 架构设计与分布式

1. 常见的缓存策略有哪些，如何做到缓存与 DB 里的数据一致性，你们项目中用到了什么缓存系统，如何设计的。
2. 用 java 自己实现一个 LRU。
3. 分布式集群下如何做到唯一序列号。
4. 设计一个秒杀系统，30 分钟没付款就自动关闭交易。
5. 如何使用 redis 和 zookeeper 实现分布式锁？有什么区别优缺点，会有什么问题，分别适用什么场景。（延伸：如果知道 redlock，讲讲他的算法实现，争议在哪里）
6. 如果有人恶意创建非法连接，怎么解决。
7. 分布式事务的原理，优缺点，如何使用分布式事务，2pc 3pc 的区别，解决了哪些问题，还有哪些问题没解决，如何解决，你自己项目里涉及到分布式事务是怎么处理的。
8. 什么是一致性 hash。
9. 什么是 restful，讲讲你理解的 restful。
10. 如何设计建立和保持 100w 的长连接。
11. 如何防止缓存击穿和雪崩。
12. 解释什么是 MESI 协议(缓存一致性)。
13. 说说你知道的几种 HASH 算法，简单的也可以。
14. 什么是 paxos 算法，什么是 zab 协议。
15. 一个在线文档系统，文档可以被编辑，如何防止多人同时对同一份文档进行编辑更新。
16. 线上系统突然变得异常缓慢，你如何查找问题。
17. 说说你平时用到的设计模式。
18. Dubbo 的原理，有看过源码么，数据怎么流转的，怎么实现集群，负载均衡，服务注册和发现，重试转发，快速失败策略是怎样的。
19. 一次 RPC 请求的流程是什么。
20. 自己实现过 rpc 么，原理可以简单讲讲。Rpc 要解决什么问题。
21. 异步模式的用途和意义。
22. 缓存数据过期后的更新如何设计。
23. 编程中自己都怎么考虑一些设计原则的，比如开闭原则，以及在工作中的应用。
24. 设计一个社交网站中的“私信”功能，要求高并发、可扩展等等。画一下架构图。
25. MVC 模式，即常见的 MVC 框架。
26. 聊下曾经参与设计的服务器架构并画图，谈谈遇到的问题，怎么解决的。
27. 应用服务器怎么监控性能，各种方式的区别。
28. 如何设计一套高并发支付方案，架构如何设计。
29. 如何实现负载均衡，有哪些算法可以实现。
30. Zookeeper 的用途，选举的原理是什么。
31. Zookeeper watch 机制原理。
32. Mybatis 的底层实现原理。

33. 请思考一个方案，设计一个可以控制缓存总体大小的自动适应的本地缓存。
34. 请思考一个方案，实现分布式环境下的 `countDownLatch`。
35. 后台系统怎么防止请求重复提交。
36. 如何看待缓存的使用（本地缓存，集中式缓存），简述本地缓存和集中式缓存和优缺点。  
本地缓存在并发使用时的注意事项。
37. 描述一个服务从发布到被消费的详细过程。
38. 讲讲你理解的服务治理。
39. 如何做到接口的幂等性。
40. 如何做限流策略，令牌桶和漏斗算法的使用场景。
41. 什么叫数据一致性，你怎么理解数据一致性。
42. 分布式服务调用方，不依赖服务提供方的话，怎么处理服务方挂掉后，大量无效资源请求的浪费，如果只是服务提供方吞吐不高的时候该怎么做，如果服务挂了，那么一会重启，该怎么做到最小的资源浪费，流量半开的实现机制是什么。
43. `dubbo` 的泛化调用怎么实现的，如果是你，你会怎么做。

## 算法

1. 10 亿个数字里里面找最小的 10 个。
2. 有 1 亿个数字，其中有 2 个是重复的，快速找到它，时间和空间要最优。
3. 2 亿个随机生成的无序整数，找出中间大小的值。
4. 给一个不知道长度的（可能很大）输入字符串，设计一种方案，将重复的字符排重。
5. 遍历二叉树。
6. 有  $3n+1$  个数字，其中  $3n$  个中是重复的，只有 1 个是不重复的，怎么找出来。
7. 写一个字符串反转函数。
8. 常用的排序算法，快排，归并、冒泡。快排的最优时间复杂度，最差复杂度。冒泡排序的优化方案。
9. 二分查找的时间复杂度，优势。
10. 一个已经构建好的 `TreeSet`，怎么完成倒排序。
11. 什么是 B+ 树，B-树，列出实际的使用场景。
12. 一个单向链表，删除倒数第 N 个数据。
13. 200 个有序的数组，每个数组里面 100 个元素，找出 top20 的元素。
14. 单向链表，查找中间的那个元素。

## 数据库知识

1. 数据库隔离级别有哪些，各自的含义是什么，MySQL 默认的隔离级别是是什么。
2. MySQL 有哪些存储引擎，各自优缺点。
3. 高并发下，如何做到安全的修改同一行数据。

4. 乐观锁和悲观锁是什么，**INNODB** 的标准行级锁有哪 2 种，解释其含义。
5. **SQL** 优化的一般步骤是什么，怎么看执行计划，如何理解其中各个字段的含义。
6. 数据库会死锁吗，举一个死锁的例子，**mysql** 怎么解决死锁。
7. **MySQL** 的索引原理，索引的类型有哪些，如何创建合理的索引，索引如何优化。
8. 聚集索引和非聚集索引的区别。
9. **select for update** 是什么含义，会锁表还是锁行或是其他。
10. **Btree** 怎么分裂的，什么时候分裂，为什么是平衡的。
11. 数据库的 **ACID** 是什么。
12. 某个表有近千万数据，**CRUD** 比较慢，如何优化。
13. **MySQL** 怎么优化 **table scan** 的。
14. 如何写 **sql** 能够有效的使用到复合索引。
15. **mysql** 中 **in** 和 **exists** 区别。
16. 数据库自增主键可能的问题。
17. **MVCC** 的含义。

## 消息队列

1. 用过哪些 **MQ**，和其他 **mq** 比较有什么优缺点，**MQ** 的连接是线程安全的吗，你们公司的 **MQ** 服务架构怎样的。
2. **MQ** 系统的数据如何保证不丢失。
3. **rabbitmq** 如何实现集群高可用。
4. **kafka** 吞吐量高的原因。
5. **kafka** 和其他消息队列的区别，**kafka** 主从同步怎么实现。
6. 利用 **mq** 怎么实现最终一致性。

## Redis, Memcached

1. **redis** 的 **list** 结构相关的操作。
2. **Redis** 的数据结构都有哪些。
3. **Redis** 的使用要注意什么，讲讲持久化方式，内存设置，集群的应用和优劣势，淘汰策略等。
4. **redis2** 和 **redis3** 的区别，**redis3** 内部通讯机制。
5. 当前 **redis** 集群有哪些玩法，各自优缺点，场景。
6. **Memcache** 的原理，哪些数据适合放在缓存中。
7. **redis** 和 **memcached** 的内存管理的区别。
8. **Redis** 的并发竞争问题如何解决，了解 **Redis** 事务的 **CAS** 操作吗。
9. **Redis** 的选举算法和流程是怎样的。
10. **redis** 的持久化的机制，**aof** 和 **rdb** 的区别。

11. redis 的集群怎么同步的数据的。
12. 知道哪些 redis 的优化操作。

## 搜索

1. elasticsearch 了解多少，说说你们公司 es 的集群架构，索引数据大小，分片有多少，以及一些调优手段。elasticsearch 的倒排索引是什么。
2. elasticsearch 索引数据多了怎么办，如何调优，部署。
3. lucence 内部结构是什么。