

《Java 后端技术面试汇总》

作 者： Java 后端技术 微信公众号

更新日期： 20190301

版本号	修改情况	修改人	修改日期
V2.0			



扫码关注微信公众号：Java 之间

回复：1024，送你 1T 免费视频资源

Java 后端技术

- 技术
- 职场
- 资源
- 资讯
- 面试
- 分享



一、Java 面试通关秘籍第一套

1、基础篇

1.1、Java 基础

- 面向对象的特征：继承、封装和多态
- final, finally, finalize 的区别
- Exception、Error、运行时异常与一般异常有何异同
- 请写出 5 种常见到的 runtime exception
- int 和 Integer 有什么区别，Integer 的值缓存范围
- 包装类，装箱和拆箱
- String、StringBuilder、StringBuffer
- 重载和重写的区别
- 抽象类和接口有什么区别
- 说说反射的用途及实现
- 说说自定义注解的场景及实现
- HTTP 请求的 GET 与 POST 方式的区别
- Session 与 Cookie 区别
- 列出自己常用的 JDK 包
- MVC 设计思想
- equals 与 == 的区别
- hashCode 和 equals 方法的区别与联系

- 什么是 Java 序列化和反序列化，如何实现 Java 序列化？或者请解释 Serializable 接口的作用
- Object 类中常见的方法，为什么 wait notify 会放在 Object 里边？
- Java 的平台无关性如何体现出来的
- JDK 和 JRE 的区别
- Java 8 有哪些新特性

1.2、Java 常见集合

- List 和 Set 区别
- Set 和 hashCode 以及 equals 方法的联系
- List 和 Map 区别
- ArrayList 与 LinkedList 区别
- ArrayList 与 Vector 区别
- HashMap 和 Hashtable 的区别
- HashSet 和 HashMap 区别
- HashMap 和 ConcurrentHashMap 的区别
- HashMap 的工作原理及代码实现，什么时候用到红黑树
- 多线程情况下 HashMap 死循环的问题
- HashMap 出现 Hash DOS 攻击的问题
- ConcurrentHashMap 的工作原理及代码实现，如何统计所有的元素个数
- 手写简单的 HashMap
- 看过那些 Java 集合类的源码

1.3、进程和线程

- 线程和进程的概念、并行和并发的概念
- 创建线程的方式及实现
- 进程间通信的方式
- 说说 CountDownLatch、CyclicBarrier 原理和区别
- 说说 Semaphore 原理
- 说说 Exchanger 原理
- ThreadLocal 原理分析，ThreadLocal 为什么会出现 OOM，出现的深层次原理
- 讲讲线程池的实现原理
- 线程池的几种实现方式
- 线程的生命周期，状态是如何转移的
- 可参考：《[Java 多线程编程核心技术](#)》

1.4、锁机制

- 说说线程安全问题，什么是线程安全，如何保证线程安全
- 重入锁的概念，重入锁为什么可以防止死锁
- 产生死锁的四个条件（互斥、请求与保持、不剥夺、循环等待）
- 如何检查死锁（通过 jConsole 检查死锁）
- volatile 实现原理（禁止指令重排、刷新内存）
- synchronized 实现原理（对象监视器）
- synchronized 与 lock 的区别

- AQS 同步队列
- CAS 无锁的概念、乐观锁和悲观锁
- 常见的原子操作类
- 什么是 ABA 问题，出现 ABA 问题 JDK 是如何解决的
- 乐观锁的业务场景及实现方式
- Java 8 并发包下常见的并发类
- 偏向锁、轻量级锁、重量级锁、自旋锁的概念
- 可参考：《[Java 多线程编程核心技术](#)》

1.5、JVM

- JVM 运行时内存区域划分
- 内存溢出 OOM 和堆栈溢出 SOE 的示例及原因、如何排查与解决
- 如何判断对象是否可以回收或存活
- 常见的 GC 回收算法及其含义
- 常见的 JVM 性能监控和故障处理工具类：jps、jstat、jmap、jinfo、jconsole 等
- JVM 如何设置参数
- JVM 性能调优
- 类加载器、双亲委派模型、一个类的生命周期、类是如何加载到 JVM 中的
- 类加载的过程：加载、验证、准备、解析、初始化
- 强引用、软引用、弱引用、虚引用
- Java 内存模型 JMM

1.6、设计模式

- 常见的设计模式
- 设计模式的六大原则及其含义
- 常见的单例模式以及各种实现方式的优缺点，哪一种最好，手写常见的单利模式
- 设计模式在实际场景中的应用
- Spring 中用到了哪些设计模式
- MyBatis 中用到了哪些设计模式
- 你项目中有使用哪些设计模式
- 说说常用开源框架中设计模式使用分析
- 动态代理很重要！！

1.7、数据结构

- 树（二叉查找树、平衡二叉树、红黑树、B 树、B+树）
- 深度有限算法、广度优先算法
- 克鲁斯卡尔算法、普林母算法、迪克拉斯算法
- 什么是一致性 Hash 及其原理、Hash 环问题
- 常见的排序算法和查找算法：快排、折半查找、堆排序等

1.8、网络/IO 基础

- BIO、NIO、AIO 的概念
- 什么是长连接和短连接
- Http1.0 和 2.0 相比有什么区别，可参考《[Http 2.0](#)》

- Https 的基本概念
- 三次握手和四次挥手、为什么挥手需要四次
- 从浏览器中输入 URL 到页面加载的发生了什么？可参考《[从输入 URL 到页面加载发生了什么](#)》

2、数据存储和消息队列



2.1、数据库

- MySQL 索引使用的注意事项
- DDL、DML、DCL 分别指什么
- explain 命令
- left join, right join, inner join
- 数据库事物 ACID（原子性、一致性、隔离性、持久性）
- 事物的隔离级别（读未提交、读以提交、可重复读、可序列化读）
- 脏读、幻读、不可重复读
- 数据库的几大范式
- 数据库常见的命令
- 说说分库与分表设计
- 分库与分表带来的分布式困境与应对之策（如何解决分布式下的分库分表，全局表？）
- 说说 SQL 优化之道
- MySQL 遇到的死锁问题、如何排查与解决

- 存储引擎的 InnoDB 与 MyISAM 区别，优缺点，使用场景
- 索引类别（B+树索引、全文索引、哈希索引）、索引的原理
- 什么是自适应哈希索引（AHI）
- 为什么要用 B+tree 作为 MySQL 索引的数据结构
- 聚集索引与非聚集索引的区别
- 遇到过索引失效的情况没，什么时候可能会出现，如何解决
- limit 20000 加载很慢怎么解决
- 如何选择合适的分布式主键方案
- 选择合适的数据存储方案
- 常见的几种分布式 ID 的设计方案
- 常见的数据库优化方案，在你的项目中数据库如何进行优化的

2.2、Redis

- Redis 有哪些数据类型，可参考《[Redis 常见的 5 种不同的数据类型详解](#)》
- Redis 内部结构
- Redis 使用场景
- Redis 持久化机制，可参考《[使用快照和 AOF 将 Redis 数据持久化到硬盘中](#)》
- Redis 集群方案与实现
- Redis 为什么是单线程的？
- 缓存雪崩、缓存穿透、缓存预热、缓存更新、缓存降级
- 使用缓存的合理性问题

- Redis 常见的回收策略

2.3、消息队列

- 消息队列的使用场景
- 消息的重发补偿解决思路
- 消息的幂等性解决思路
- 消息的堆积解决思路
- 自己如何实现消息队列
- 如何保证消息的有序性

3、开源框架和容器

3.1、SSM/Servlet

- Servlet 的生命周期
- 转发与重定向的区别
- BeanFactory 和 ApplicationContext 有什么区别
- Spring Bean 的生命周期
- Spring IOC 如何实现
- Spring 中 Bean 的作用域，默认的是哪一个
- 说说 Spring AOP、Spring AOP 实现原理
- 动态代理（CGLib 与 JDK）、优缺点、性能对比、如何选择
- Spring 事务实现方式、事务的传播机制、默认的事务类别
- Spring 事务底层原理

- Spring 事务失效（事务嵌套），JDK 动态代理给 Spring 事务埋下的坑，可参考《[JDK 动态代理给 Spring 事务埋下的坑！](#)》
- 如何自定义注解实现功能
- Spring MVC 运行流程
- Spring MVC 启动流程
- Spring 的单例实现原理
- Spring 框架中用到了哪些设计模式
- Spring 其他产品 (Spring Boot、Spring Cloud、Spring Security、Spring Data、Spring AMQP 等)
- 有没有用到 Spring Boot，Spring Boot 的认识、原理
- MyBatis 的原理
- 可参考《[为什么会有 Spring](#)》
- 可参考《[为什么会有 Spring AOP](#)》

3.2、Netty

- 为什么选择 Netty
- 说说业务中，Netty 的使用场景
- 原生的 NIO 在 JDK 1.7 版本存在 epoll bug
- 什么是 TCP 粘包/拆包
- TCP 粘包/拆包的解决办法
- Netty 线程模型
- 说说 Netty 的零拷贝

- Netty 内部执行流程

- Netty 重连实现

3.3、Tomcat

- Tomcat 的基础架构 (Server、Service、Connector、Container)
- Tomcat 如何加载 Servlet 的
- Pipeline-Valve 机制
- 可参考：《[四张图带你了解 Tomcat 系统架构！](#)》

4、分布式

4.1、Nginx

- 请解释什么是 C10K 问题或者知道什么是 C10K 问题吗？
- Nginx 简介，可参考《[Nginx 简介](#)》
- 正向代理和反向代理.
- Nginx 几种常见的负载均衡策略
- Nginx 服务器上的 Master 和 Worker 进程分别是什么
- 使用“反向代理服务器”的优点是什么？

4.2、分布式其他

- 谈谈业务中使用分布式的场景
- Session 分布式方案
- Session 分布式处理
- 分布式锁的应用场景、分布式锁的产生原因、基本概念

- 分布式锁的常见解决方案
- 分布式事务的常见解决方案
- 集群与负载均衡的算法与实现
- 说说分库与分表设计，可参考《[数据库分库分表策略的具体实施方案](#)》
- 分库与分表带来的分布式困境与应对之策

4.3、Dubbo

- 什么是 Dubbo，可参考《[Dubbo 入门](#)》
- 什么是 RPC、如何实现 RPC、RPC 的实现原理，可参考《[基于 HTTP 的 RPC 实现](#)》
- Dubbo 中的 SPI 是什么概念
- Dubbo 的基本原理、执行流程

5、微服务



5.1、微服务

- 前后端分离是如何做的？
- 微服务哪些框架
- Spring Cloud 的常见组件有哪些？可参考《[Spring Cloud 概述](#)》
- 领域驱动有了解吗？什么是领域驱动模型？充血模型、贫血模型
- JWT 有了解吗，什么是 JWT，可参考《[前后端分离利器之 JWT](#)》
- 你怎么理解 RESTful
- 说说如何设计一个良好的 API

- 如何理解 RESTful API 的幂等性
- 如何保证接口的幂等性
- 说说 CAP 定理、BASE 理论
- 怎么考虑数据一致性问题
- 说说最终一致性的实现方案
- 微服务的优缺点，可参考《[微服务批判](#)》
- 微服务与 SOA 的区别
- 如何拆分服务、水平分割、垂直分割
- 如何应对微服务的链式调用异常
- 如何快速追踪与定位问题
- 如何保证微服务的安全、认证

5.2、安全问题

- 如何防范常见的 Web 攻击、如何方式 SQL 注入
- 服务端通信安全攻防
- HTTPS 原理剖析、降级攻击、HTTP 与 HTTPS 的对比

5.3、性能优化

- 性能指标有哪些
- 如何发现性能瓶颈
- 性能调优的常见手段
- 说说你在项目中如何进行性能调优

6、其他

6.1、设计能力

- 说说你在项目中使用过的 UML 图
- 你如何考虑组件化、服务化、系统拆分
- 秒杀场景如何设计
- 可参考：《[秒杀系统的技术挑战、应对策略以及架构设计总结一二！](#)》

6.2、业务工程

- 说说你的开发流程、如何进行自动化部署的
- 你和团队是如何沟通的
- 你如何进行代码评审
- 说说你对技术与业务的理解
- 说说你在项目中遇到感觉最难 Bug，是如何解决的
- 介绍一下工作中的一个你认为最有价值的项目，以及在这个过程中的角色、解决的问题、你觉得你们项目还有哪些不足的地方

6.3、软实力

- 说说你的优缺点、亮点
- 说说你最近在看什么书、什么博客、在研究什么新技术、再看那些开源项目的源代码
- 说说你觉得最有意义的技术书籍
- 工作之余做什么事情、平时是如何学习的，怎样提升自己的能力
- 说说个人发展方向方面的思考
- 说说你认为的服务端开发工程师应该具备哪些能力

- 说说你认为是架构师是什么样的，架构师主要做什么
- 如何看待加班的问题



二、Java 面试通关秘籍第二套

1、Java 相关

- ArrayList 与 LinkedList 默认空间是多少；
- ArrayList 与 LinkedList 区别与各自的优势 List 和 Map 区别；
- 谈谈 HashMap，哈希表解决 hash 冲突的方法；
- 为什么要重写 hashCode()和 equals()以及他们之间的区别与关系；
- Object 的 hashCode()是怎么计算的？
- 若 hashCode 方法永远返回 1 或者一个常量会产生什么结果？
- Java Collections 和 Arrays 的 sort 方法默认的排序方法是什么；
- 引用计数法与 GC Root 可达性分析法区别；
- 浅拷贝和深拷贝的区别；
- String s = "abc"和 String s = new String("abc")区别；
- HashSet 方法里面的 hashCode 存在哪,如果重写 equals 不重写 hashCode 会怎么样？
- 反射的作用与实现原理；
- Java 中的回调机制；
- 模板方法模式；
- 开闭原则说一下；
- 发布/订阅使用场景；
- KMP 算法（一种改进的字符串匹配算法）；

- JMM 里边的原子性、可见性、有序性是如何体现出来的，JMM 中内存屏障是什么意思，

2、多线程

- AtomicInteger 底层实现原理；
- synchronized 与 ReentrantLock 哪个是公平锁；
- CAS 机制会出现什么问题；
- 用过并发包下边的哪些类；
- 一个线程连着调用 start 两次会出现什么情况？
- wait 方法能不能被重写，wait 能不能被中断；
- 线程池的实现？四种线程池？重要参数及原理？任务拒接策略有哪几种？
- 线程状态以及 API 怎么操作会发生这种转换；
- 常用的避免死锁方法；

3、JVM

- Minor GC 与 Full GC 分别在什么时候发生？什么时候触发 Full GC；
- GC 收集器有哪些？CMS 收集器与 G1 收集器的特点。
- Java 在什么时候会出现内存泄漏；
- Java 中的大对象如何进行存储；
- rt.jar 被什么类加载器加载，什么时间加载；
- 自己写的类被什么加载，什么时间加载；
- 自己写的两个不同的类是被同一个类加载器加载的吗？为什么？

- 为什么新生代内存需要有两个 Survivor 区？
- 几种常用的内存调试工具：jmap、jstack、jconsole；
- 类加载的五个过程：加载、验证、准备、解析、初始化；
- G1 停顿吗，CMS 回收步骤，CMS 为什么会停顿，停顿时间；
- 栈主要存的数据是什么，堆呢？
- 堆分为哪几块，比如说新生代老生代，那么新生代又分为什么？
- 软引用和弱引用的使用场景（软引用可以实现缓存，弱引用可以用来在回调函数中防止内存泄露）；

4、数据库

- 数据库索引，什么是全文索引，全文索引中的倒排索引是什么原理
- 数据库最佳左前缀原则是什么？
- 数据库的三大范式；
- 悲观锁和乐观锁的原理和应用场景；
- 左连接、右连接、内连接、外连接、交叉连接、笛卡儿积等；
- 一般情况下数据库宕机了如何进行恢复（什么是 Write Ahead Log 机制，什么是 Double Write 机制，什么是 Check Point）；
- 什么是 redo 日志、什么是 undo 日志；
- 数据库中的隔离性是怎样实现的；原子性、一致性、持久性又是如何实现的；
- 什么是组合索引，组合索引什么时候会失效；
- 关系型数据库和非关系型数据库区别；

- 数据库死锁如何解决；
- MySQL 并发情况下怎么解决（通过事务、隔离级别、锁）；
- MySQL 中的 MVCC 机制是什么意思，根据具体场景，MVCC 是否有问题；
- MySQL 数据库的隔离级别，以及如何解决幻读；

5、缓存服务器

- Redis 中 zSet 跳跃表问题；
- Redis 的 set 的应用场合？
- Redis 高级特性了解吗？
- Redis 的 pipeline 有什么用处？
- Redis 集群宕机如何处理，怎么样进行数据的迁移；
- Redis 的集群方案；
- Redis 原子操作怎么用比较好；
- Redis 过期策略是怎么实现的呢？

6、SSM 相关

- Spring 中 @Autowired 和 @Resource 注解的区别？
- Spring 声明一个 bean 如何对其进行个性化定制；
- MyBatis 有什么优势；
- MyBatis 如何做事务管理；

7、操作系统

- Linux 静态链接和动态链接；
- 什么是 IO 多路复用模型 (select、poll、epoll) ；
- Linux 中的 grep 管道用处？Linux 的常用命令？
- 操作系统中虚拟地址、逻辑地址、线性地址、物理地址的概念及区别；
- 内存的页面置换算法；
- 内存的页面置换算法；
- 进程调度算法，操作系统是如何调度进程的；
- 父子进程、孤儿进程、僵死进程等概念；
- fork 进程时的操作；
- kill 用法，某个进程杀不掉的原因 (僵死进程；进入内核态，忽略 kill 信号)；
- 系统管理命令（如查看内存使用、网络情况）；
- find 命令、awk 使用；
- Linux 下排查某个死循环的线程；

8、网络相关

- 数据链路层是做什么的？
- 数据链路层的流量控制？
- 网络模型的分层、IP 和 Mac 地址在那个层、TCP 和 HTTP 分别在那个层；
- TCP 滑动窗口；
- TCP 为什么可靠；
- TCP 的同传，拆包与组装包是什么意思；

- Https 和 Http 有什么区别；
- Http 为什么是无状态的；
- TCP 三次握手，为什么不是三次，为什么不是四次；
- TCP 的拥塞控制、流量控制详细说明？
- Http1.0 和 Http2.0 的区别；
- 两个不同 ip 地址的计算机之间如何通信；
- 地址解析协议 ARP；
- OSI 七层模型分别对应着五层模型的哪一部分；
- TCP 三次握手数据丢失了怎么办？那如果后面又找到了呢？

9、分布式相关

- 消息队列使用的场景介绍和作用（应用耦合、异步消息、流量削锋等）；
- 如何解决消息队列丢失消息和重复消费问题；
- Kafka 使用过吗，什么是幂等性？怎么保证一致性，持久化怎么做，分区 partition 的理解，LEO 是什么意思，如何保证多个 partition 之间数据一致性的（ISR 机制），为什么 Kafka 可以这么快（基于磁盘的顺序读写）；
- 异步队列怎么实现；
- 你项目的并发是多少？怎么解决高并发问题？单机情况下 Tomcat 的并发大概是多少，MySQL 的并发大致是多少？
- 什么是 C10K 问题；
- 高并发情况下怎么办；

- 分布式理论，什么是 CAP 理论，什么是 Base 理论，什么是 Paxos 理论；
- 分布式协议的选举算法；
- 说一下你对微服务的理解，与 SOA 的区别；
- Dubbo 的基本原理，RPC，支持哪些通信方式，服务的调用过程；
- Dubbo 如果有一个服务挂掉了怎么办；
- 分布式事务，操作两个表不在一个库，如何保证一致性。
- 分布式系统中，每台机器如何产生一个唯一的随机值；
- 系统的量级、pv、uv 等；
- 什么是 Hash 一致性算法？分布式缓存的一致性，服务器如何扩容（哈希环）；
- 正向代理、反向代理；
- 什么是客户端负载均衡策略、什么是服务器端负载均衡策略；
- 如何优化 Tomcat，常见的优化方式有哪些；
- Nginx 的 Master 和 Worker，Nginx 是如何处理请求的；

10、系统设计相关

- 如何防止表单重复提交（Token 令牌环等方式）；
- 有一个 url 白名单，需要使用正则表达式进行过滤，但是 url 量级很大，大概亿级，那么如何优化正则表达式？如何优化亿级的 url 匹配呢？
- 常见的 Nginx 负载均衡策略；已有两台 Nginx 服务器了，倘若这时候再增加一台服务器，采用什么负载均衡算法比较好？

- 扫描二维码登录的过程解析；
- 如何设计一个生成唯一 UUID 的算法？
- 实现一个负载均衡的算法，服务器资源分配为 70%、20%、10%；
- 有三个线程 T1 T2 T3，如何保证他们按顺序执行；
- 三个线程循环输出 ABCABCABC....

11、安全相关

- 什么是 XSS 攻击，XSS 攻击的一般表现形式有哪些？如何防止 XSS 攻击；

三、Java 面试通关秘籍第三套

1、基础题

- 怎么解决 Hash 冲突；（开放地址法、链地址法、再哈希法、建立公共溢出区等）
- 写出一个必然会产生死锁的伪代码；
- Spring IoC 涉及到的设计模式；（工厂模式、单利模式。。）
- toString()方法什么情况下需要重写；
- 判断对象相等时，什么情况下只需要重写 equals()，什么情况下需要重写 equals(),hashCode()？
- Set 内存放的元素为什么不可以重复，内部是如何保证和实现的？
- 如何保证分布式缓存的一致性(分布式缓存一致性 hash 算法?)？分布式 session 实现？
- Java 8 流式迭代的好处？
- 项目中用到的 JDK 的哪些特性？
- 说一下 TreeMap 的实现原理？红黑树的性质？红黑树遍历方式有哪些？如果 key 冲突如何解决？setColor()方法在什么时候用？什么时候会进行旋转和颜色转换？
- Spring 的 bean 的创建时机？依赖注入的时机？
- ArrayList 和 LinkedList 的删除一个元素的时间复杂度；(ArrayList 是 $O(N)$ ，LinkedList 是 $O(1)$) ；

- CopyOnWriteArrayList 是什么；
- 序列化和反序列化底层如何实现的（ObjectOutputStream、ObjectInputStream、readObject writeObject）；
- 如何调试多线程的程序；
- 一个线程连着调用 start 两次会出现什么情况？（由于状态只有就绪、阻塞、执行，状态是无法由执行转化为执行的，所以会报不合法的状态！）
- HashMap 在什么时候时间复杂度是 $O(1)$ ，什么时候是 $O(n)$ ，什么时候又是 $O(\log n)$ ；
- wait 方法能不能被重写？（wait 是 final 类型的，不可以被重写，不仅如此，notify 和 notifyall 都是 final 类型的），wait 能不能被中断；
- 一个 Controller 调用两个 Service，这两 Service 又都分别调用两个 Dao，问其中用到了几个数据库连接池的连接？

2、网络基础

- HTTP、TCP、UDP 的区别和联系；
- TCP 和 UDP 各自的优势，知道哪些使用 UDP 协议的成功案例；
- TCP 和 UDP 各用了底层什么协议；
- 单个 UDP 报文最大容量；
- 单个 TCP 报文最大容量；
- TCP 报头格式、UDP 报头格式；
- Server 遭遇 SYN Flood 应当怎么处理；

- Web 开发中如何防范 XSS?
- 拆包和粘包的问题，如何解决，如果我们的包没有固定长度的话，我们的应用程序应该如何解决；

3、操作系统

- 为什么要内存对齐；
- 为什么会有大端小端，htol 这一类函数的作用；
- top 显示出来的系统信息都是什么含义；（重要！）
- Linux 地址空间，怎么样进行寻址的；
- Linux 如何查找目录或者文件的；

4、分布式其他

- 分库与分表带来的分布式困境与应对之策；
- Solr 如何实现全天 24 小时索引更新；

5、Redis

- Redis 插槽的分配（key 的有效部分使用 CRC16 算法计算出哈希值，再将哈希值对 16384 取余，得到插槽值）；
- Redis 主从是怎么选取的（一种是主动切换，另一种是使用 sentinel 自动方式）；
- Redis 复制的过程；
- Redis 队列应用场景；

- Redis 主节点宕机了怎么办，还有没有同步的数据怎么办；

6、系统设计开放性题目

- 秒杀系统设计，超卖怎么搞；
- 你们的图片时怎么存储的，对应数据库时如何保存图片的信息的？
- 假如成都没有一座消防站，现在问你要建立几座消防站，每个消防站要配多少名消防官兵，多少辆消防车，请你拿出一个方案；
- 基于数组实现一个循环阻塞队列；
- 常见的 ipv4 地址的展现形式如 “168.0.0.1”，请实现 ip 地址和 int 类型的相互转换。（使用位移的方式）
- 现网某个服务部署在多台 Linux 服务器上，其中一台突然出现 CPU 100% 的情况，而其他服务器正常，请列举可能导致这种情况发生的原因？如果您遇到这样的情况，应如何定位？内存？CPU？发布？debug？请求量？

7、大数据量问题（后边会有专题单独讨论）

- 给定 a、b 两个文件，各存放 50 亿个 url，每个 url 各占 64 字节，内存限制是 4G，让你找出 a、b 文件共同的 url？
- 海量日志数据，提取出某日访问百度次数最多的那个 IP；
- 一个文本文件，大约有一万行，每行一个词，要求统计出其中最频繁出现的前 10 个词，请给出思想，给出时间复杂度分析。

此话题后边会有专门的文章探讨，如果有等不及的小伙伴，可以移步参考：

1、https://blog.csdn.net/v_july_v/article/details/6279498

2、https://blog.csdn.net/v_july_v/article/details/7382693

8、逻辑思维题

- 有两根粗细均匀的香（烧香拜佛的香），每一根烧完都花一个小时，怎么样能够得到 15min？
- 假定你有 8 个撞球，其中有 1 个球比其他的球稍重,如果只能利用天平来断定哪一个球重,要找到较重的球,要称几次？（2 次）；
- 实验室里有 1000 个一模一样的瓶子，但是其中的一瓶有毒。可以用实验室的小白鼠来测试哪一瓶是毒药。如果小白鼠喝掉毒药的话，会在一个星期的时候死去，其他瓶子里的药水没有任何副作用。请问最少用多少只小白鼠可以在一个星期以内查出哪瓶是毒药；（答案是 10 只）
- 假设有一个池塘，里面有无穷多的水。现有 2 个空水壶，容积分别为 5 升和 6 升。问题是如何只用这 2 个水壶从池塘里取得 3 升的水；



四、Java 面试通关秘籍第四套

1、Java 基础

- 为什么 JVM 调优经常会将-Xms 和 -Xmx 参数设置成一样；
- Java 线程池的核心属性以及处理流程；
- Java 内存模型，方法区存什么；
- CMS 垃圾回收过程；
- Full GC 次数太多了，如何优化；
- 直接内存如何管理的；
- Java 线程池的几个参数的意义和实现机制；
- Java 线程池使用无界任务队列和有界任务队列的优劣对比；
- CountDownLatch 和 CyclicBarrier 的区别；
- Java 中有哪些同步方案（重量级锁、显式锁、并发容器、并发同步器、CAS、volatile、AQS 等）
- 如果你的项目出现了内存泄露，怎么监控这个问题呢；
- 标记清除和标记整理的区别和优缺点，为何标记整理会发生 stop the world；
- 线程池，如何根据 CPU 的核数来设计线程大小，如果是计算机密集型的呢，如果是 IO 密集型的呢？
- 让你设计一个 cache 如何设计；
- String 中 hashCode 是怎么实现的；
- JDK 中哪些实现了单例模式？

- 多个线程同时读写，读线程的数量远远大于写线程，你认为应该如何解决并发的问题？你会选择加什么样的锁？
- 线程池内的线程如果全部忙，提交一个新的任务，会发生什么？队列全部塞满了之后，还是忙，再提交会发生什么？
- synchronized关键字锁住的是什么东西？在字节码中是怎么表示的？在内存中的对象上表现为什么？
- wait/notify/notifyAll方法需不需要被包含在 synchronized 块中？这是为什么？
- ExecutorService 你一般是怎么用的？是每个 Service 放一个还是个项目放一个？有什么好处？

2、数据库

- InnoDB 的插入缓冲和两次写的概率和意义；
- 如果建了一个单列索引，查询的时候查出 2 列，会用到这个单列索引吗？（会用到）
- 如果建了一个包含多个列的索引，查询的时候只用了第一列，能不能用上这个索引？查三列呢？
- 接上题，如果 where 条件后面带有一个 $i + 5 < 100$ 会使用到这个索引吗？
- like %aaa%会使用索引吗？like aaa%呢？
- drop、truncate、delete 的区别？

- 平时你们是怎么监控数据库的？慢 SQL 是怎么排查的？（慢查询日志）
- 你们数据库是否支持 emoji 表情，如果不支持，如何操作？选择什么编码方式？如果支持一个表情占几个字节？(utf8mb4)；
- 如果查询很慢，你会想到的第一个方式是什么？（数据库索引）

3、Linux 基础



- Linux 下可以在 /proc 目录下可以查看 CPU 的核心数等；cat /proc/下边会有很多系统内核信息可供显示；
- 说一下栈的内存是怎么分配的；
- Linux 各个目录有了解过吗？/etc、/bin、/dev、/lib、/sbin 这些常见的目录主要作用是什么？
- 说一下栈帧的内存是怎么分配的；
- Linux 下排查某个死循环的线程；
- 动态链接和静态链接的区别；
- 进程的内存分布；
- 如何查找一个进程打开所有的文件；
- 说一下常使用的协议及其对应的端口；
- 为什么会有内核态，保护模式你知道吗？
- 文件是怎么在磁盘上存储的？
- 有了进程为何还要线程呢，不同进程和线程他们之间有什么不同。（进程是资源管理的最小单位，线程是程序执行的最小单位。在操作系统设计上，从

进程演化出线程，最主要的目的就是更好的支持 SMP 以及减小（进程/线程）上下文切换开销。）

- InnoDB 聚集索引 B+ 树叶子节点和磁盘什么顺序相同；
- 文件系统，进程管理和调度，内存管理机制、虚地址保护模式；

4、网络基础

- HTTP1.0 和 HTTP1.1 的区别；
- DHCP 如何实现分配 IP 的；发现阶段（DHCP 客户端在网络中广播发送 DHCP DISCOVER 请求报文，发现 DHCP 服务器，请求 IP 地址租约）、提供阶段（DHCP 服务器通过 DHCP OFFER 报文向 DHCP 客户端提供 IP 地址预分配）、选择阶段（DHCP 客户端通过 DHCP REQUEST 报文确认选择第一个 DHCP 服务器为它提供 IP 地址自动分配服务）和确认阶段（被选择的 DHCP 服务器通过 DHCP ACK 报文把在 DHCP OFFER 报文中准备的 IP 地址租约给对应 DHCP 客户端）。
- OSI 七层模型，每层都说下自己的理解和知道的，说的越多越好；

5、框架相关

- Servlet 如何保证单例模式,可不可以编程多例的哪？
- Dubbo 请求流程以及原理；
- Spring 框架如何实现事务的；
- 如果一个接口有 2 个不同的实现, 那么怎么来 Autowire 一个指定的实现？
(可以使用 Qualifier 注解限定要注入的 Bean，也可以使用 Qualifier 和

Autowire 注解指定要获取的 bean，也可以使用 Resource 注解的 name 属性指定要获取的 Bean)

- Spring 框架中需要引用哪些 jar 包，以及这些 jar 包的用途；
- Spring Boot 没有放到 web 容器里为什么能跑 HTTP 服务？
- Spring 中循环注入是什么意思，可不可以解决，如何解决；
- Spring 的声明式事务 @Transaction 注解一般写在什么位置？抛出了异常会自动回滚吗？有没有办法控制不触发回滚？
- MyBatis 怎么防止 SQL 注入；
- Tomcat 本身的参数你一般会怎么调整？
- 了解哪几种序列化协议？如何选择合适的序列化协议；
- Redis 渐进式 rehash 过程？
- 比如我有个电商平台，做每日订单的异常检测，服务端代码应该写；



五、Java 面试通关秘籍第五套

1、Java 相关

- 乐观悲观锁的设计，如何保证原子性，解决的问题；
- char 和 double 的字节，以及在内存的分布是怎样；
- 对象内存布局，然后讲下对象的死亡过程？
- 对象头，详细讲下；
- sync 原理详细，sync 内抛异常会怎样，死锁吗？还是释放掉？怎么排查死锁？死锁会怎样？有没有什么更好的替代方案？
- 详细讲一下集合，HashSet 源码，HashMap 源码，如果要线程安全需要怎么做？
- 多线程是解决什么问题的？线程池解决什么问题？
- 线程池，如何设计的，里面的参数有多少种，里面的工作队列和线程队列是怎样的结构，如果给你，怎样设计线程池？
- AQS 原理，ReentrantLock 源码，设计原理，整体过程。
- 继续聊多线程源码，sync 原理，然后一个场景设计题；
- float f = 1.4f;double d = 1.4d; 与 float f = 1.5f;double d = 1.5d; 是否为 true，内存是怎样的；
- split 的源码，split("a|b|c");得出多少个数组；
- 把所有认识熟用的 JUC(java.util.concurrent(简称 JUC)包)下的类写出来，讲下使用，然后讲下原生的线程操作；

- 开闭原则，解析工厂方法模式，建造者模式，区别。手撸出来。
- 讲下 JVM 的大页模式，JVM 内存模型；
- 什么是敏捷开发，防御性编程，并行编程。Team Leader 的思考；
- 逃逸分析是什么，作用是什么，用途是什么；
- 怎么认为一个类是线程安全？线程安全的定义是什么？Java 有多少个关键字进行同步？为什么这样设计？（聊了一大堆，一堆为什么）；
- 两个线程设计题。记得一个是：t1,t2,t3，让 t1, t2 执行完才执行 t3，原生实现。
- 写个后缀表达式，为什么要设计后缀表达式，有什么好处？然后写下中缀。
- 我看你做过性能优化，比如你怎么分析项目里面的 OOM 的，内存泄露呢？详细说思路；
- 说下多线程，我们什么时候需要分析线程数，怎么分析，分析什么因素；
- 抽象方法和类方法的区别，static 的抽象方法可以吗？
- 说下 Java 的克隆体系；
- 涉及 OOM、JVM 优化、源码问题、数据库优化、多线程等问题；
- CPU 高？什么情况 CPU 高？解决什么问题？
- 你有遇到过临界区问题吗？有遇到过吗？你在项目遇到这个问题是怎样解决的？
- volatile 关键字作用；
- Java 的多态怎么实现；
- 解释一下自旋；

- 解释一下信号量;
- 什么情况下会触发类加载;
- Java 内存抖动严重, 优化的思路;

2、数据库相关

- SQL 优化思路, 联合索引与底层树结构的映像关系, 索引结构 (B+、B-), 为什么用这样的结构;
- 讲下 MySQL 的集群? 集群遇到过什么问题? sql 的优化?
- 你目前为止遇到的最大数据量是多少? 知道 100 万时候怎么设计吗? 1000 万呢? 过几十亿呢?
- MySQL 有多少个参数可调, 除了最大连接数。全部列出来, 一个个分析。
- 聊下优化过的索引, 怎么优化;
- 红黑树和平衡树的区别, 为什么数据库不用红黑树;
- mysql 有哪些锁, 意向锁有什么用;
- 数据库高并发下的优化思路;
- 数据库什么情况下索引会失效;

3、数据结构和操作系统相关

- 数据结构学过吧, 聊一下? 学过什么结构? 讲下树和队列? B 树呢?
- 操作系统学过吧, 聊一下? 讲一下系统内存是怎样的? 分段分页虚拟内存?
- 页面置换算法呢? 多少种? 有最优的置换算法吗?
- 你学过什么课程? 然后聊下操作系统, 内核、用户之类。

- 反转链表手撸;
- 快排, 给一串数组, 把具体每次 partition 写下, 最终结果也写 45, 32, 41, 35, 38, 20, 50;
- 一个整数 status, 判断第 K 个比特位是否为比特 1;
- 把递归实现的快排改成非递归, 你知道非递归有什么好处吗;
- 举例使用分治思想的算法;

4、网络相关

- 讲下请求头细节?
- Http 和 Https? Http1.0,1.1,2.0, 讲下长连接和短连接? Https 是怎样的?
如果我篡改了公钥呢? 怎么防止?
- Get 和 Post, 讲下区别, 要我模拟出抓包来。
- 详细讲下 Cookie 和 Session, Token, OAuth2.0 协议;
- 拥塞算法知道吗? 哪些, 分别怎样?
- 学过计算机网络是吧? socket 熟悉吗? 对它的读写缓冲区有理解吗? 怎么的? 那滑动窗口是怎样的? 为什么这样设计?
- 再聊下 Http 的 Http basic authentication;
- Https 的过程;

5、框架相关

- 聊下 Spring 源码, 知道多少, 都聊一下;
- 聊下 Spring 注解, @Autowire, @Resource, 以及他们的解析过程;

- 聊一下架构，接入层架构，服务层架构。聊下技术栈，Spring Boot, Spring Cloud、Docker;
- Spring ioc 的具体优势，和直接 New 一个对象有什么区别;
- Servlet 生命周期，是否单例，为什么是单例;
- Spring Mvc 初始化过程;

6、分布式相关

- 多少种 RPC 框架?
- 一致性哈希是干嘛的?
- 搭建高并发高可用系统需要怎样设计? 考虑哪些东西，有多少说多少。
- 你对缓存有什么理解? 缓存是解决什么问题? 后端缓存有哪些，分别解决什么问题?
- 聊一下分布式锁;
- 你是怎么设计系统缓存的，为什么，什么场景;
- 也来说下，削峰的多种实现，Redis? MQ?
- 为什么用 mq 就能削峰? 解决什么问题?

7、设计题

- 有几台机器存储着几亿淘宝搜索日志，你只有一台 2g 的电脑，怎么选出搜索热度最高的十个搜索关键词;
- 如何设计算法压缩一段 URL;

- 有一个页面能同时展示两个广告，现在有五个广告，设计算法使五个广告展示概率为 1:2:3:4:5；
- 有 25 匹马，五个赛道，用最少的比赛次数将 25 匹马排序；

8、其他相关

- Tomcat 缓存，聊下缓存的整体理解，知道多少种缓存；
- 解释下 Mucene 原理，倒排索引，怎样进行中文分词，基于什么进行分词；
- TopN 的大数据量题；
- 你对接入层要思考什么东西？遇到过哪些问题？搭建系统要考量哪些因素？
- 然后项目问题，优化问题；
- 熟悉 maven 是吧？我们来聊下 Maven 的源码原理，Maven 冲突的时候，怎么选依赖包，我们怎么查，我们遇到两个不一样的版本，我们应该如何去选择，为什么？
- 项目如何分组，性能优化小组应该做哪些；
- 我们来说下接入层的搭建，认知分析；
- 问下项目的系统构建，思考，为什么这样构建？
- 如何判断一段代码的好坏；