目录

[**一、** **多线程** 2](#_Toc471485036)

[**二、** **集合** 10](#_Toc471485037)

[**三、** **Spring** 13](#_Toc471485038)

[**四、** **数据库** 23](#_Toc471485039)

[**五、** **JVM** 26](#_Toc471485040)

[**六、** **NIO** 33](#_Toc471485041)

[**七、** **Redis** 33](#_Toc471485042)

[**八、** **MQ** 33](#_Toc471485043)

[**九、** **其他** 34](#_Toc471485044)

1. **多线程**

#### 1.1 实现方式

##### 1.1.1 Thread

继承Thread类的方法尽管被我列为一种多线程实现方式，但Thread本质上也是实现了Runnable接口的一个实例，它代表一个线程的实例，并且，启动线程的唯一方法就是通过Thread类的start()实例方法。start()方法是一个native方法，它将启动一个新线程，并执行run()方法。这种方式实现多线程很简单，通过自己的类直接extend Thread，并复写run()方法，就可以启动新线程并执行自己定义的run()方法。例如：

public class MyThread extends Thread {

　　public void run() {

　　 System.out.println("MyThread.run()");

　　}

}

在合适的地方启动线程如下：

MyThread myThread1 = new MyThread();

MyThread myThread2 = new MyThread();

myThread1.start();

myThread2.start();

##### 1.1.2 Runnable

如果自己的类已经extends另一个类，就无法直接extends Thread，此时，必须实现一个Runnable接口，如下：

public class MyThread extends OtherClass implements Runnable {

　　public void run() {

　　 System.out.println("MyThread.run()");

　　}

}

为了启动MyThread，需要首先实例化一个Thread，并传入自己的MyThread实例：

MyThread myThread = new MyThread();

Thread thread = new Thread(myThread);

thread.start();

事实上，当传入一个Runnable target参数给Thread后，Thread的run()方法就会调用target.run()，参考JDK源代码：

public void run() {

　　if (target != null) {

　　 target.run();

　　}

}

##### 1.1.3 Future

ExecutorService、Callable、Future这个对象实际上都是属于Executor框架中的功能类。想要详细了解Executor框架的可以访问http://www.javaeye.com/topic/366591 ，这里面对该框架做了很详细的解释。返回结果的线程是在JDK1.5中引入的新特征，确实很实用，有了这种特征我就不需要再为了得到返回值而大费周折了，而且即便实现了也可能漏洞百出。

可返回值的任务必须实现Callable接口，类似的，无返回值的任务必须Runnable接口。执行Callable任务后，可以获取一个Future的对象，在该对象上调用get就可以获取到Callable任务返回的Object了，再结合线程池接口ExecutorService就可以实现传说中有返回结果的多线程了。下面提供了一个完整的有返回结果的多线程测试例子，在JDK1.5下验证过没问题可以直接使用。代码如下：

import java.util.concurrent.\*;

import java.util.Date;

import java.util.List;

import java.util.ArrayList;

/\*\*

\* 有返回值的线程

\*/

@SuppressWarnings("unchecked")

public class Test {

public static void main(String[] args) throws ExecutionException,

InterruptedException {

System.out.println("----程序开始运行----");

Date date1 = new Date();

int taskSize = 5;

// 创建一个线程池

ExecutorService pool = Executors.newFixedThreadPool(taskSize);

// 创建多个有返回值的任务

List<Future> list = new ArrayList<Future>();

for (int i = 0; i < taskSize; i++) {

Callable c = new MyCallable(i + " ");

// 执行任务并获取Future对象

Future f = pool.submit(c);

// System.out.println(">>>" + f.get().toString());

list.add(f);

}

// 关闭线程池

pool.shutdown();

// 获取所有并发任务的运行结果

for (Future f : list) {

// 从Future对象上获取任务的返回值，并输出到控制台

System.out.println(">>>" + f.get().toString());

}

Date date2 = new Date();

System.out.println("----程序结束运行----，程序运行时间【"

+ (date2.getTime() - date1.getTime()) + "毫秒】");

}

}

class MyCallable implements Callable<Object> {

private String taskNum;

MyCallable(String taskNum) {

this.taskNum = taskNum;

}

public Object call() throws Exception {

System.out.println(">>>" + taskNum + "任务启动");

Date dateTmp1 = new Date();

Thread.sleep(1000);

Date dateTmp2 = new Date();

long time = dateTmp2.getTime() - dateTmp1.getTime();

System.out.println(">>>" + taskNum + "任务终止");

return taskNum + "任务返回运行结果,当前任务时间【" + time + "毫秒】";

}

}

代码说明：

上述代码中Executors类，提供了一系列工厂方法用于创先线程池，返回的线程池都实现了ExecutorService接口。

public static ExecutorService newFixedThreadPool(int nThreads)

创建固定数目线程的线程池。

public static ExecutorService newCachedThreadPool()

创建一个可缓存的线程池，调用execute 将重用以前构造的线程（如果线程可用）。如果现有线程没有可用的，则创建一个新线程并添加到池中。终止并从缓存中移除那些已有 60 秒钟未被使用的线程。

public static ExecutorService newSingleThreadExecutor()

创建一个单线程化的Executor。

public static ScheduledExecutorService newScheduledThreadPool(int corePoolSize)

创建一个支持定时及周期性的任务执行的线程池，多数情况下可用来替代Timer类。

ExecutoreService提供了submit()方法，传递一个Callable，或Runnable，返回Future。如果Executor后台线程池还没有完成Callable的计算，这调用返回Future对象的get()方法，会阻塞直到计算完成。

#### 1.2 线程池

##### 1.2.1 内部原理

使用部分：

http://www.cnblogs.com/easycloud/p/3726089.html

缺陷：

http://www.ibm.com/developerworks/cn/java/j-jtp0730/

好处：

1.减少在创建和销毁线程上所花的时间以及系统资源的开销

2.如不使用线程池，有可能造成系统创建大量线程而导致消耗完系统内存以及”过度切换”。

**总结:**

一个任务通过execute(Runnable)方法被添加到线程池，任务必须是一个 Runnable类型的对象，任务的执行方法就是调用Runnable类型对象的run()方法。当一个任务通过execute(Runnable)方法欲添加到线程池时，会做一下几步：

1. 如果此时线程池中的数量小于corePoolSize，即使线程池中的线程都处于空闲状态，也要创建新的线程来处理被添加的任务。

2. 如果此时线程池中的数量大于等于corePoolSize，但是缓冲队列 workQueue未满，那么任务被放入缓冲队列。

3. 如果此时线程池中的数量大于corePoolSize，缓冲队列workQueue满，并且线程池中的数量小于maximumPoolSize，建新的线程来处理添加的任务。

4. 如果此时线程池中的数量大于corePoolSize，缓冲队列workQueue满，并且线程池中的数量等于maximumPoolSize，那么通过 handler所指定的策略来处理此任务。也就是处理任务的优先级为：核心线程corePoolSize、任务队列workQueue、最大线程maximumPoolSize，如果三者都满了，使用handler处理被拒绝的任务。

5. 当线程池中的线程数量大于corePoolSize时，如果某线程空闲时间超过keepAliveTime，线程将被终止。这样，线程池可以动态的调整池中的线程数。

##### 1.2.2 常见类型

1、newFixedThreadPool创建一个指定工作线程数量的线程池。每当提交一个任务就创建一个工作线程，如果工作线程数量达到线程池初始的最大数，则将提交的任务存入到池队列中。

2、newCachedThreadPool创建一个可缓存的线程池。这种类型的线程池特点是：

1).工作线程的创建数量几乎没有限制(其实也有限制的,数目为Interger. MAX\_VALUE), 这样可灵活的往线程池中添加线程。

2).如果长时间没有往线程池中提交任务，即如果工作线程空闲了指定的时间(默认为1分钟)，则该工作线程将自动终止。终止后，如果你又提交了新的任务，则线程池重新创建一个工作线程。

3、newSingleThreadExecutor创建一个单线程化的Executor，即只创建唯一的工作者线程来执行任务，如果这个线程异常结束，会有另一个取代它，保证顺序执行(我觉得这点是它的特色)。单工作线程最大的特点是可保证顺序地执行各个任务，并且在任意给定的时间不会有多个线程是活动的 。

4、newScheduleThreadPool创建一个定长的线程池，而且支持定时的以及周期性的任务执行，类似于Timer。

1. **集合**

#### 2.1 HashMap

##### 2.1.1 Hash算法

典型的哈希算法包括 MD2、MD4、MD5 和 SHA-1。它是一种单向密码体制,即它是一个从明文到密文的不可逆的映射,只有加密过程,没有解密过程。同时,哈希函数可以将任意长度的输入经过变化以后得到固定长度的输出。哈希函数的这种单向特征和输出数据长度固定的特征使得它可以生成消息或者数据。

##### 2.1.2 内部原理

http://www.cnblogs.com/chenssy/p/3521565.html

初始容量，加载因子。这两个参数是影响HashMap性能的重要参数，其中容量表示哈希表中桶的数量，初始容量是创建哈希表时的容量，加载因子是哈希表在其容量自动增加之前可以达到多满的一种尺度，它衡量的是一个散列表的空间的使用程度，负载因子越大表示散列表的装填程度越高，反之愈小。对于使用链表法的散列表来说，查找一个元素的平均时间是O(1+a)，因此如果负载因子越大，对空间的利用更充分，然而后果是查找效率的降低；如果负载因子太小，那么散列表的数据将过于稀疏，对空间造成严重浪费。系统默认负载因子为0.75，一般情况下我们是无需修改的。

当我们想一个HashMap中添加一对key-value时，系统首先会计算key的hash值，然后根据hash值确认在table中存储的位置。若该位置没有元素，则直接插入。否则迭代该处元素链表并依此比较其key的hash值。如果两个hash值相等且key值相等(e.hash == hash && ((k = e.key) == key || key.equals(k))),则用新的Entry的value覆盖原来节点的value。如果两个hash值相等但key值不等 ，则将该节点插入该链表的链头。

面试部分：

http://www.admin10000.com/document/3322.html

##### 2.1.3 线程不安全的原因

addEntry（操作），resize（扩容：双倍）

#### 2.2 ConcurrentHashMap

##### 2.2.1 内部原理

查找segment中的位置，分段实现线程安全，HashEntry内部为final ，所以删除时，复制删除节点前所有元素在和后面的重新组成新segment；添加时在表头添加。

#### 2.3 ArrayList

##### 2.3.1 内部原理

(01) ArrayList 实际上是通过一个数组去保存数据的。当我们构造ArrayList时；若使用默认构造函数，则ArrayList的默认容量大小是10。

(02) 当ArrayList容量不足以容纳全部元素时，ArrayList会重新设置容量：新的容量=“(原始容量x3)/2 + 1”。

(03) ArrayList的克隆函数，即是将全部元素克隆到一个数组中。

(04) ArrayList实现java.io.Serializable的方式。当写入到输出流时，先写入“容量”，再依次写入“每一个元素”；当读出输入流时，先读取“容量”，再依次读取“每一个元素”。

#### 2.4 BlockingQueue

##### 2.4.1 内部原理

阻塞队列（BlockingQueue）是一个支持两个附加操作的队列。这两个附加的操作是：在队列为空时，获取元素的线程会等待队列变为非空。当队列满时，存储元素的线程会等待队列可用。阻塞队列常用于生产者和消费者的场景，生产者是往队列里添加元素的线程，消费者是从队列里拿元素的线程。阻塞队列就是生产者存放元素的容器，而消费者也只从容器里拿元素。

阻塞队列提供了四种处理方法:

1.抛出异常：是指当阻塞队列满时候，再往队列里插入元素，会抛出IllegalStateException("Queue full")异常。当队列为空时，从队列里获取元素时会抛出NoSuchElementException异常 。

2.返回特殊值：插入方法会返回是否成功，成功则返回true。移除方法，则是从队列里拿出一个元素，如果没有则返回null

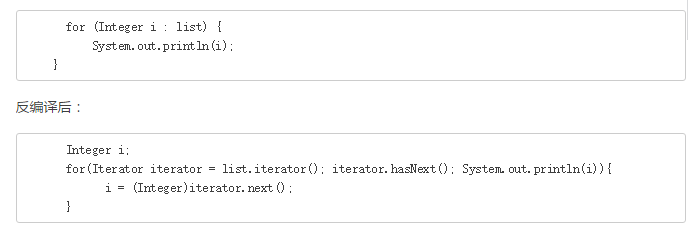
3.一直阻塞：当阻塞队列满时，如果生产者线程往队列里put元素，队列会一直阻塞生产者线程，直到拿到数据，或者响应中断退出。当队列空时，消费者线程试图从队列里take元素，队列也会阻塞消费者线程，直到队列可用。

4.超时退出：当阻塞队列满时，队列会阻塞生产者线程一段时间，如果超过一定的时间，生产者线程就会退出。

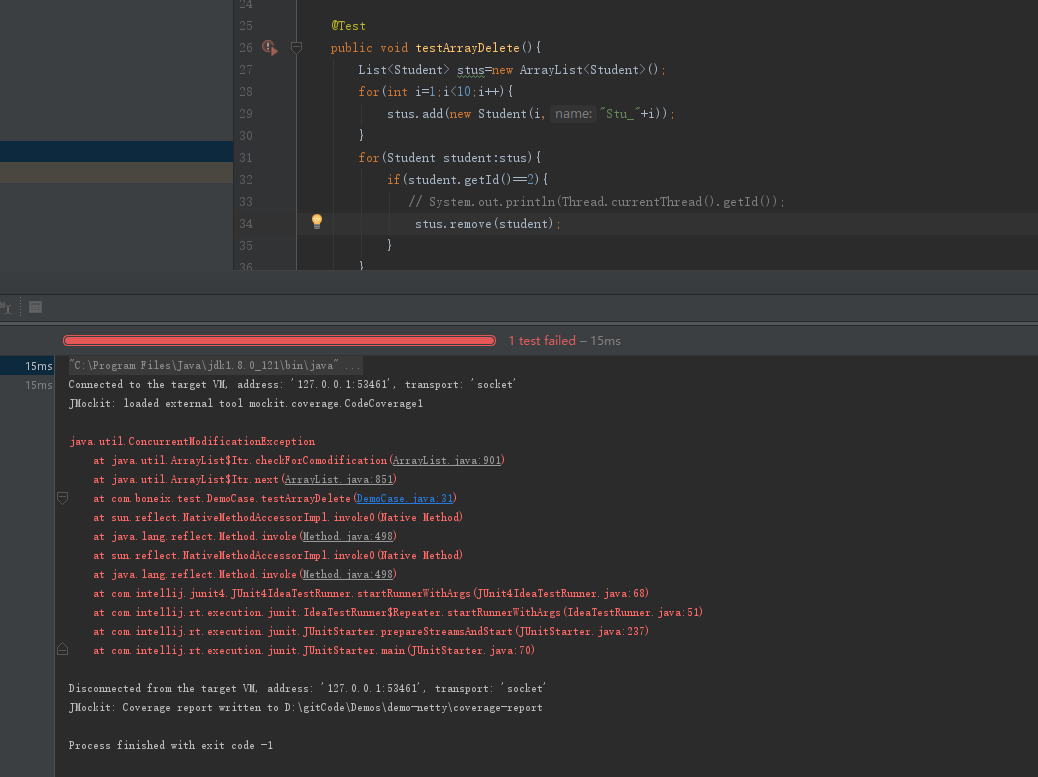
#### 2.5 面试相关

##### 2.5.1 增强型for循环

增强型for循环底层通过Iterator实现，

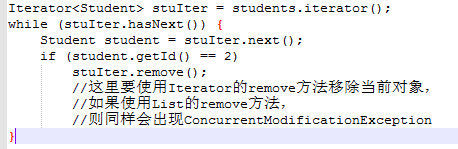


Iterator是工作在一个独立的线程中，并且拥有一个 mutex 锁。 Iterator被创建之后会建立一个指向原来对象的单链索引表，当原来的对象数量发生变化时，这个索引表的内容不会同步改变，所以当索引指针往后移动的时候就找不到要迭代的对象，所以按照 fail-fast 原则 Iterator 会马上抛出java.util.ConcurrentModificationException异常。



所以 Iterator 在工作的时候是不允许被迭代的对象被改变的。但你可以使用 Iterator 本身的方法 remove() 来删除对象，Iterator.remove() 方法会在删除当前迭代对象的同时维护索引的一致性。

正确的在遍历的同时删除元素的姿势：



1. **Spring**

#### 3.1 注入方式

##### 3.1.1 构造器注入

constructor-arg

##### 3.1.2 属性注入

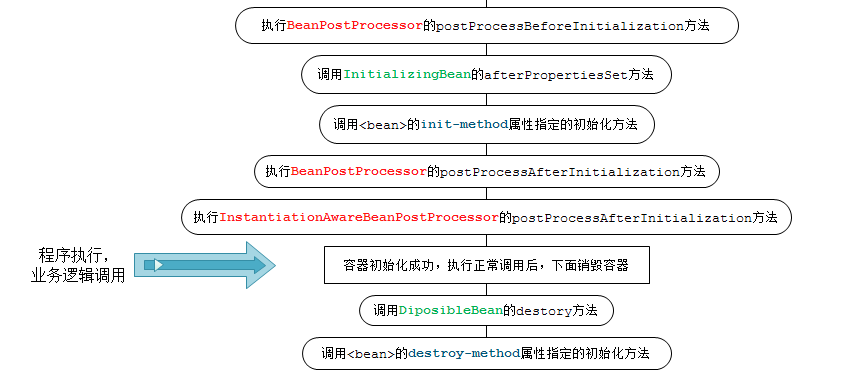
property

##### 3.1.3 静态工厂注入

factory-method

#### 3.2 Bean生命周期





1.Spring容器读取XML文件中bean的定义并实例化bean。

2.Spring根据bean的定义设置属性值。

3.如果该Bean实现了BeanNameAware接口，Spring将bean的id传递给setBeanName()方法。

4.如果该Bean实现了BeanFactoryAware接口，Spring将beanfactory传递给setBeanFactory()方法。

5.如果任何bean BeanPostProcessors 和该bean相关，Spring调用postProcessBeforeInitialization()方法。

6.如果该Bean实现了InitializingBean接口，调用Bean中的afterPropertiesSet方法。如果bean有初始化函数声明，调用相应的初始化方法。

7.如果任何bean BeanPostProcessors 和该bean相关，调用postProcessAfterInitialization()方法。

8.如果该bean实现了DisposableBean，调用destroy()方法。

#### 3.3 自定义标签

##### 3.3.1 如何解析

(诸如aop等头部标签是在xml头部文件中指定对应的handler)定义handler类实现NamespaceHandlerSupport接口，在init()方法中通过registerBeanDefinitionParser注册解析类（实现BeanDefinitionParser接口，getBeanClass()方法用来返回对应的模板类，doParse()用来具体解析xml配置）

#### 3.4 Aop

##### 3.4.1 配置



#### 3.5 事务

##### 3.5.1 配置

根据代理机制的不同，总结了五种Spring事务的配置方式，配置文件如下：

第一种方式：每个Bean都有一个代理

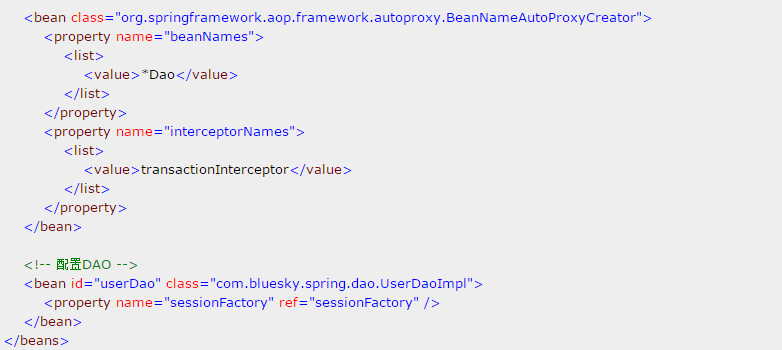


第二种方式：所有Bean共享一个代理基类



第三种方式：使用拦截器



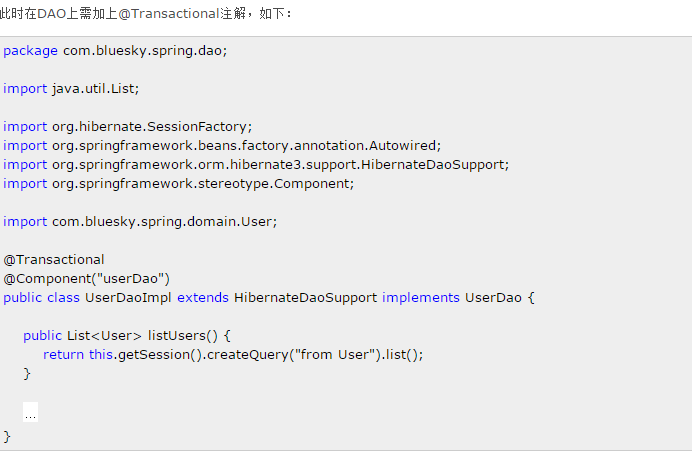


第四种方式：使用tx标签配置的拦截器

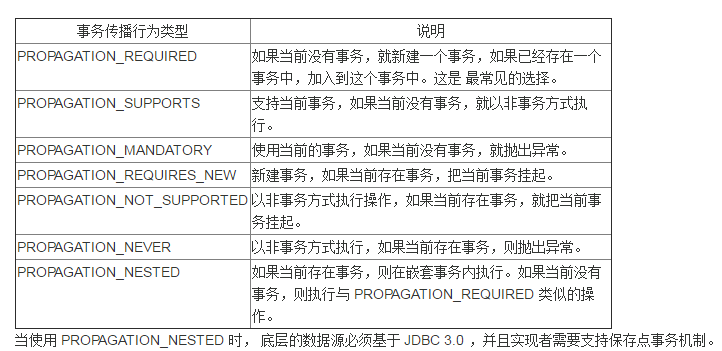


第五种方式：全注解





##### 3.5.2 传播性



##### 3.5.3 隔离性

数据库并发操作存在的异常情况：

1. 更新丢失（Lost update）： 两个事务都同时更新一行数据但是第二个事务却中途失败退出导致对数据两个修改都失效了这是系统没有执 行任何锁操作因此并发事务并没有被隔离开来。

2. 脏读取（Dirty Reads）： 一个事务开始读取 了某行数据但是另外一个事务已经更新了此数据但没有能够及时提交。这是相当危险很可能所有操作都被回滚。

3. 不可重复读取（Non-repeatable Reads）： 一 个事务对同一行数据重复读取两次但是却得到了不同结果。例如在两次读取中途有另外一个事务对该行数据进行了修改并提交。

4. 两次更新问题（Second lost updates problem）： 无法重复读取特例，有两个并发事务同时读取同一行数据然后其中一个对它进行修改提交而另一个也进行了修改提交这就会造成 第一次写操作失效。

5. 幻读（Phantom Reads）： 也称为幻像（幻 影）。事务在操作过程中进行两次查询，第二次查询结果包含了第一次查询中未出现的数据（这里并不要求两次查询SQL语句相同）这是因为在两次查询过程中有 另外一个事务插入数据造成的。

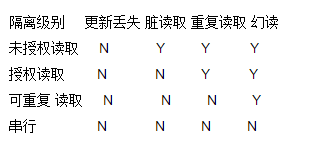
为了避免上面出现几种情况在标准SQL规范中定义了4个事务隔离级别，不同隔离级别对事务处理不同 。

1.未授权读取（Read Uncommitted）： 也称 未提交读。允许脏读取但不允许更新丢失，如果一个事务已经开始写数据则另外一个数据则不允许同时进行写操作但允许其他事务读此行数据。该隔离级别可以通过 “排他写锁”实现。事务隔离的最低级别，仅可保证不读取物理损坏的数据。与READ COMMITTED 隔离级相反，它允许读取已经被其它用户修改但尚未提交确定的数据。

2. 授权读取（Read Committed）： 也称提交 读。允许不可重复读取但不允许脏读取。这可以通过“瞬间共享读锁”和“排他写锁”实现，读取数据的事务允许其他事务继续访问该行数据，但是未提交写事务将 会禁止其他事务访问该行。SQL Server 默认的级别。在此隔离级下，SELECT 命令不会返回尚未提交（Committed） 的数据，也不能返回脏数据。

3. 可重复读取（Repeatable Read）： 禁止 不可重复读取和脏读取。但是有时可能出现幻影数据，这可以通过“共享读锁”和“排他写锁”实现，读取数据事务将会禁止写事务（但允许读事务），写事务则禁 止任何其他事务。在此隔离级下，用SELECT 命令读取的数据在整个命令执行过程中不会被更改。此选项会影响系统的效能，非必要情况最好不用此隔离级。

4. 串行（Serializable）： 也称可串行读。提 供严格的事务隔离，它要求事务序列化执行，事务只能一个接着一个地执行，但不能并发执行。如果仅仅通过“行级锁”是无法实现事务序列化的，必须通过其他机 制保证新插入的数据不会被刚执行查询操作事务访问到。事务隔离的最高级别，事务之间完全隔离。如果事务在可串行读隔离级别上运行，则可以保证任何并发重叠 事务均是串行的。



##### 3.5.4 不生效的情况

1、首先使用如下代码 确认你的bean 是代理对象吗？

必须是spring定义（通过XML或注解定义都可以）的Bean才接受事务。

直接new出来的对象添加事务是不起作用的。

可以通过以下方式判断是否是代理对象：

AopUtils.isAopProxy(Object object)

AopUtils.isCglibProxy(Object object) //cglib

AopUtils.isJdkDynamicProxy(Object object) //jdk动态代理

2、如果你使用了springmvc，可能是context:component-scan重复扫描引起的。

3、如使用MySQL且引擎是MyISAM造成的（因为不支持事务），改成InnoDB即可。

#### 3.6 面试相关

##### 3.6.1 BeanFactory和FactoryBean的区别

BeanFactory： 以Factory结尾，表示它是一个工厂类，是用于管理Bean的一个工厂;

FactoryBean：以Bean结尾，表示它是一个Bean，不同于普通Bean的是：它是实现了FactoryBean<T>接口的Bean，根据该Bean的Id从BeanFactory中获取的bean实际上是FactoryBean的getObject()返回的对象，而不是FactoryBean本身， 如果要获取FactoryBean对象，可以在id前面加一个&符号来获取。

1. **数据库**

#### 4.1 MySQL

##### 4.1.1 常见类型

**MyISAM**

MyISAM表无法处理事务，这就意味着有事务处理需求的表，不能使用MyISAM存储引擎。MyISAM存储引擎特别适合在以下几种情况下使用：

1.选择密集型的表。MyISAM存储引擎在筛选大量数据时非常迅速，这是它最突出的优点。

2.插入密集型的表。MyISAM的并发插入特性允许同时选择和插入数据。例如：MyISAM存储引擎很适合管理邮件或Web服务器日志数据。

**InnoDB**

InnoDB是一个健壮的事务型存储引擎，这种存储引擎已经被很多互联网公司使用，为用户操作非常大的数据存储提供了一个强大的解决方案。我的电脑上安装的MySQL 5.6.13版，InnoDB就是作为默认的存储引擎。InnoDB还引入了行级锁定和外键约束，在以下场合下，使用InnoDB是最理想的选择：

1.更新密集的表。InnoDB存储引擎特别适合处理多重并发的更新请求。

2.事务。InnoDB存储引擎是支持事务的标准MySQL存储引擎。

3.自动灾难恢复。与其它存储引擎不同，InnoDB表能够自动从灾难中恢复。

4.外键约束。MySQL支持外键的存储引擎只有InnoDB。

5.支持自动增加列AUTO\_INCREMENT属性。

一般来说，如果需要事务支持，并且有较高的并发读取频率，InnoDB是不错的选择。

**MEMORY**

使用MySQL Memory存储引擎的出发点是速度。为得到最快的响应时间，采用的逻辑存储介质是系统内存。虽然在内存中存储表数据确实会提供很高的性能，但当mysqld守护进程崩溃时，所有的Memory数据都会丢失。获得速度的同时也带来了一些缺陷。它要求存储在Memory数据表里的数据使用的是长度不变的格式，这意味着不能使用BLOB和TEXT这样的长度可变的数据类型，VARCHAR是一种长度可变的类型，但因为它在MySQL内部当做长度固定不变的CHAR类型，所以可以使用。

一般在以下几种情况下使用Memory存储引擎：

1.目标数据较小，而且被非常频繁地访问。在内存中存放数据，所以会造成内存的使用，可以通过参数max\_heap\_table\_size控制Memory表的大小，设置此参数，就可以限制Memory表的最大大小。

2.如果数据是临时的，而且要求必须立即可用，那么就可以存放在内存表中。

3.存储在Memory表中的数据如果突然丢失，不会对应用服务产生实质的负面影响。

**MERGE**

MERGE存储引擎是一组MyISAM表的组合，这些MyISAM表结构必须完全相同，尽管其使用不如其它引擎突出，但是在某些情况下非常有用。说白了，Merge表就是几个相同MyISAM表的聚合器；Merge表中并没有数据，对Merge类型的表可以进行查询、更新、删除操作，这些操作实际上是对内部的MyISAM表进行操作。Merge存储引擎的使用场景。

对于服务器日志这种信息，一般常用的存储策略是将数据分成很多表，每个名称与特定的时间端相关。例如：可以用12个相同的表来存储服务器日志数据，每个表用对应各个月份的名字来命名。当有必要基于所有12个日志表的数据来生成报表，这意味着需要编写并更新多表查询，以反映这些表中的信息。与其编写这些可能出现错误的查询，不如将这些表合并起来使用一条查询，之后再删除Merge表，而不影响原来的数据，删除Merge表只是删除Merge表的定义，对内部的表没有任何影响。

**ARCHIVE**

Archive是归档的意思，在归档之后很多的高级功能就不再支持了，仅仅支持最基本的插入和查询两种功能。在MySQL 5.5版以前，Archive是不支持索引，但是在MySQL 5.5以后的版本中就开始支持索引了。Archive拥有很好的压缩机制，它使用zlib压缩库，在记录被请求时会实时压缩，所以它经常被用来当做仓库使用。

##### 4.1.2 优化方式

1. EXPLAIN 你的 SELECT 查询;

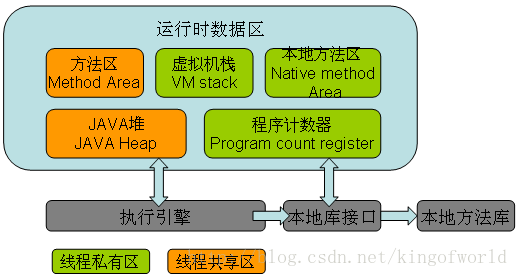
2. 当只要一行数据时使用 LIMIT 1;

3. 为搜索字段建索引;

1. **JVM**

#### 5.1 JVM 内存模型

http://www.cnblogs.com/AloneSword/p/4262255.html

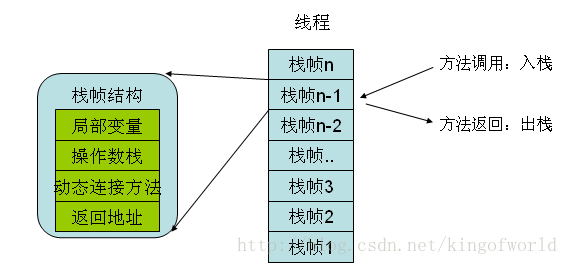


**程序计数器**

多线程时，当线程数超过CPU数量或CPU内核数量，线程之间就要根据时间片轮询抢夺CPU时间资源。因此每个线程有要有一个独立的程序计数器，记录下一条要运行的指令。线程私有的内存区域。如果执行的是JAVA方法，计数器记录正在执行的java字节码地址，如果执行的是native方法，则计数器为空。

**虚拟机栈**

线程私有的，与线程在同一时间创建。管理JAVA方法执行的内存模型。每个方法执行时都会创建一个桢栈来存储方法的的变量表、操作数栈、动态链接方法、返回值、返回地址等信息。栈的大小决定了方法调用的可达深度（递归多少层次，或嵌套调用多少层其他方法，-Xss参数可以设置虚拟机栈大小）。栈的大小可以是固定的，或者是动态扩展的。如果请求的栈深度大于最大可用深度，则抛出stackOverflowError；如果栈是可动态扩展的，但没有内存空间支持扩展，则抛出OutofMemoryError。使用jclasslib工具可以查看class类文件的结构。下图为栈帧结构图：



**本地方法区**

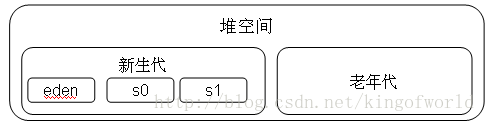
和虚拟机栈功能相似，但管理的不是JAVA方法，是本地方法，本地方法是用C实现的。

**JAVA堆**

线程共享的，存放所有对象实例和数组。垃圾回收的主要区域。可以分为新生代和老年代(tenured)。

新生代用于存放刚创建的对象以及年轻的对象，如果对象一直没有被回收，生存得足够长，老年对象就会被移入老年代。

新生代又可进一步细分为eden、survivorSpace0(s0,from space)、survivorSpace1(s1,to space)。刚创建的对象都放入eden,s0和s1都至少经过一次GC并幸存。如果幸存对象经过一定时间仍存在，则进入老年代(tenured)。



**方法区**

线程共享的，用于存放被虚拟机加载的类的元数据信息：如常量、静态变量、即时编译器编译后的代码。也成为永久代。如果hotspot虚拟机确定一个类的定义信息不会被使用，也会将其回收。回收的基本条件至少有：所有该类的实例被回收，而且装载该类的ClassLoader被回收

#### 5.2 OOM的类型及原因

##### 5.2.1 Heap

**java.lang.OutOfMemoryError: Java heap space**

**分析:**

此OOM是由于JVM中heap的最大值不满足需要，将设置heap的最大值调高即可，参数样例为：-Xmx2G

JVM堆的设置是指Java程序运行过程中JVM可以调配使用的内存空间的设置.JVM在启动的时候会自动设置Heap size的值，其初始空间(即-Xms)是物理内存的1/64，最大空间(-Xmx)是物理内存的1/4。可以利用JVM提供的-Xmn -Xms -Xmx等选项可进行设置。Heap size 的大小是Young Generation 和Tenured Generaion 之和。

提示：在JVM中如果98％的时间是用于GC且可用的Heap size 不足2％的时候将抛出此异常信息。

提示：Heap Size 最大不要超过可用物理内存的80％，一般的要将-Xms和-Xmx选项设置为相同，而-Xmn为1/4的-Xmx值。

**解决参考:**

调高heap的最大值，即-Xmx的值调大。

##### 5.2.2 StackOverflowError

**Exception in thread "main" java.lang.StackOverflowError**

**分析:**

如果线程请求的栈深度大于虚拟机所允许的最大深度，将抛出StackOverflowError异常。

如果虚拟机在扩展栈时无法申请到足够的内存空间，则抛出OutOfMemoryError异常。

一般在单线程程序情况下无法产生OutOfMemoryError异常，使用多线程方式也会出现OutOfMemeoryError，因为栈是线程私有的，线程多也会方法区溢出。

**解决参考:**

检查程序是否有深度递归。

##### 5.2.3 Perm

**java.lang.OutOfMemoryError: PermGen space**

**分析**:

PermGen space的全称是Permanent Generation space,是指内存的永久保存区域,这块内存主要是被JVM存放Class和Meta信息的,Class在被Loader时就会被放到PermGen space中,它和存放类实例(Instance)的Heap区域不同,GC(Garbage Collection)不会在主程序运行期对PermGen space进行清理，所以如果你的应用中有很多CLASS的话,就很可能出现PermGen space错误,这种错误常见在web服务器对JSP进行pre compile的时候。如果你的WEB APP下都用了大量的第三方jar, 其大小超过了jvm默认的大小那么就会产生此错误信息了。由于JVM在默认的情况下，Perm默认为64M，而很多程序需要大量的Perm区内 存，尤其使用到像spring等框架的时候，由于需要使用到动态生成类，而这些类不能被GC自动释放，所以导致OutOfMemoryError: PermGen space异常。解决方法很简单，增大JVM的 -XX:MaxPermSize 启动参数，就可以解决这个问题，如过使用的是默认变量通常是64M，改成128M就可以了，-XX:MaxPermSize=128m。如果已经是128m，就改成 256m。我一般在服务器上为安全起见，改成256m。

**解决参考:**

调高Perm的最大值，即-XX:MaxPermSize的值调大。

另外，注意一点，Perm一般是在JVM启动时加载类进来，如果是JVM运行较长一段时间而不是刚启动后溢出的话，很有可能是由于运行时有类被动态加载进来，此时建议用CMS策略中的类卸载配置。

如：-XX:+UseConcMarkSweepGC -XX:+CMSClassUnloadingEnabled

##### 5.2.4 GC

**java.lang.OutOfMemoryError: GC overhead limit exceeded**

**分析**:

此OOM是由于JVM在GC时，对象过多，导致内存溢出，建议调整GC的策略，在一定比例下开始GC而不要使用默认的策略，或者将新代和老代设置合适的大小，需要进行微调存活率。

**解决参考**:

改变GC策略，在老代80%时就是开始GC，并且将-XX:SurvivorRatio（-XX:SurvivorRatio=8）和-XX:NewRatio（-XX:NewRatio=4）设置的更合理。

##### 5.2.5 Native Thread Created

**java.lang.OutOfMemoryError: unable to create new native thread**

**分析**:

这个异常问题本质原因是我们创建了太多的线程，而能创建的线程数是有限制的，导致了异常的发生。

**解决参考:**

如果JVM内存调的过大或者可利用率小于20%，可以建议将heap及perm的最大值下调，并将线程栈调小，即-Xss调小，如：-Xss128k。

在JVM内存不能调小的前提下，将-Xss设置较小，如：-Xss:128k。

##### 5.2.6 Allocate Huge Array

**Exception in thread "main": java.lang.OutOfMemoryError: Requested array size exceeds VM limit**

**分析**:

此类信息表明应用程序试图分配一个大于堆大小的数组。例如，如果应用程序new一个数组对象，大小为512M，但是最大堆大小为256M，因此OutOfMemoryError会抛出，因为数组的大小超过虚拟机的限制。

**解决参考:**

（1）、首先检查heap的-Xmx是不是设置的过小

（2）、如果heap的-Xmx已经足够大，那么请检查应用程序是不是存在bug，例如：应用程序可能在计算数组的大小时，存在算法错误，导致数组的size很大，从而导致巨大的数组被分配。

##### 5.2.7 Small Swap

Exception in thread "main": java.lang.OutOfMemoryError: request <size> bytes for <reason>. Out of swap space?

**分析**:

抛出这类错误，是由于从native堆中分配内存失败，并且堆内存可能接近耗尽。这类错误可能跟应用程序没有关系，例如下面两种原因也会导致错误的发生：

（1）操作系统配置了较小的交换区

（2）系统的另外一个进程正在消耗所有的内存

**解决参考:**

（1）、检查os的swap是不是没有设置或者设置的过小

（2）、检查是否有其他进程在消耗大量的内存，从而导致当前的JVM内存不够分配。

注意：虽然有时<reason>部分显示导致OOM的原因，但大多数情况下，<reason>显示的是提示分配失败的源模块的名称，所以有必要查看日志文件，如crash时的hs文件。

##### 5.2.8 Exhausted Native Memory

**java.lang.OutOfMemoryErr java.io.FileInputStream.readBytes(Native Method)**

**分析**:

从错误日志来看，在OOM后面没有提示引起OOM的原因，进一步查看stack trace发现，导致OOM的原因是由Native Method的调用引起的，另外检查[Java](http://lib.csdn.net/base/java)heap，发现heap的使用正常，因而需要考虑问题的发生是由于Native memory被耗尽导致的。

**解决参考:**

从根本上来说，解决此问题的方法应该通过检测发生问题时的环境下，native memory为什么被占用或者说为什么native memory越来越小，从而去解决引起Native memory减小的问题。但是如果此问题不容易分析时，可以通过以下方法或者结合起来处理。  
（1）、cpu和os保证是64位的，并且jdk也换为64位的。  
（2）、将java heap的-Xmx尽量调小，但是保证在不影响应用使用的前提下。  
（3）、限制对native memory的消耗，比如：将thread的-Xss调小，并且限制产生大量的线程；限制文件的io操作次数和数量；限制网络的使用等等。

#### 5.3 GC算法

1、**标记-清除算法(Mark-Sweep)**

从根节点开始标记所有可达对象，其余没标记的即为垃圾对象，执行清除。但回收后的空间是不连续的。

2、**复制算法(copying)**

将内存分成两块，每次只使用其中一块，垃圾回收时，将标记的对象拷贝到另外一块中，然后完全清除原来使用的那块内存。复制后的空间是连续的。复制算法适用于新生代，因为垃圾对象多于存活对象，复制算法更高效。在新生代串行垃圾回收算法中，将eden中标记存活的对象拷贝未使用的s1中，s0中的年轻对象也进入s1，如果s1空间已满，则进入老年代；这样交替使用s0和s1。这种改进的复制算法，既保证了空间的连续性，有避免了大量的内存空间浪费。

3、**标记-压缩算法(Mark-compact)**

适合用于老年代的算法（存活对象多于垃圾对象）。

标记后不复制，而是将存活对象压缩到内存的一端，然后清理边界外的所有对象。

#### 5.4 优化参数

-XX:+PrintGCDetails 打印垃圾回收信息

-Xms 为Heap区域的初始值，线上环境需要与-Xmx设置为一致，否则capacity的值会来回飘动

-Xmx 为Heap区域的最大值

-Xss（或-ss） 线程栈大小（指一个线程的native空间）1.5以后是1M的默认大小

-XX:PermSize与-XX:MaxPermSize 方法区（永久代）的初始大小和最大值（但不是本地方法区）

-XX:NewRatio 老年代与新生代比率

-XX:SurvivorRatio Eden与Survivor的占用比例。例如8表示，一个survivor区占用 1/8 的Eden内存，即1/10的新生代内存，为什么不是1/9？因为我们的新生代有2个survivor，即S1和S22。所以survivor总共是占用新生代内存的 2/10，Eden与新生代的占比则为 8/10。

-XX:MaxHeapFreeRatio GC后，如果发现空闲堆内存占到整个预估的比例小于这个值，则减小堆空间。

-XX:MinHeapFreeRatio GC后，如果发现空闲堆内存占到整个预估的比例大于这个值，则增大堆空间。

-XX:NewSize 新生代大小

1. **NIO**

#### 6.1 基本概念

1.面向缓冲

2.非阻塞IO

1. **Redis**

#### 7.2 分布式锁实现

1.setnx

2.setex

1. **MQ**

#### 7.3 概念及作用

实现系统解耦

1. **其他**