

## 深入理解java虚拟机

### JDK

---

我们可以把Java程序设计语言、Java虚拟机、Java API类库这三部分统称为JDK

可以把Java API类库中的Java SE API子集和Java虚拟机这两部分统称为JRE

## java内存区域与内存溢出异常

---

### 运行时数据区域

#### 程序计数器(线程私有)

程序计数器可以看作是当前线程所执行的字节码的行号指示器。字节码解释器工作时就是通过改变这个计数器的值来选取下一条需要执行的字节码指令，分支、循环、跳转、异常处理、线程恢复等基础功能都需要依赖这个计数器来完成

- 线程私有的内存

java多线程中，为了线程切换后能恢复到正确的执行位置，每条线程都需要有一个独立的程序计数器，各条线程之间计数器互不影响，独立存储，我们称这类内存区域为“线程私有”的内存

- 执行Native方法时，计数器的值

如果线程正在执行的是一个java方法，这个计数器记录的是正在执行的虚拟机字节码指令的地址；如果正在执行的是Native方法，这个计数器值则为空。此内存区域是唯一一个在Java虚拟机规范中没有规定任何OutOfMemoryError情况的区域

#### java虚拟机栈(线程私有)

生命周期与线程相同

- 方法执行时

虚拟机栈描述的是Java方法执行的内存模型：每个方法在执行的同时都会创建一个栈帧(Stack Frame)用于存储局部变量表、操作数栈、动态链接、方法出口等信息。每一个方法从调用直至执行完成的过程，就对应着一个栈帧在虚拟机栈中入栈到出栈的过程

- 局部变量表

局部变量表存放了编译期可知的各种基本数据类型（boolean、byte、char、short、int、float、long、double）、对象引用和 returnAddress类型（指向了一条字节码指令的地址）

- 两种异常状况(重要)

如果线程请求的栈深度大于虚拟机所允许的深度，将抛出StackOverflowError异常；如果虚拟机栈可以动态扩展，如果扩展时无法申请到足够的内存，就会抛出OutOfMemoryError异常

#### 本地方法栈

本地方法栈与虚拟机栈所发挥的作用是非常相似的，它们之间的区别不过是虚拟机栈为虚拟机执行Java方法服务，而本地方法栈则为虚拟机使用到的Native方法服务 与虚拟机栈一样，本地方法栈区域也会抛出StackOverflowError和OutOfMemoryError异常

- 异常  
本地方法栈区域也会抛出StackOverflowError和OutOfMemoryError异常

## java堆(各线程共享)

java堆是被所有线程共享的一块内存区域，在虚拟机启动时创建 java堆区域的唯一目的就是存放对象实例

- "GC堆"  
java堆也被称为 "GC堆"，因为是垃圾收集器管理的主要区域
- 分代收集算法  
新生代和老年代，堆中数据进一步划分的目的是为了更好回收内存，或者更快地分配内存
- java堆的内存空间  
根据Java虚拟机规范的规定，Java堆可以处于物理上不连续的内存空间中，只要逻辑上是连续的即可，就像我们的磁盘空间一样。在实现时，既可以实现成固定大小的，也可以是可扩展的
- 异常  
如果在堆中没有内存完成实例分配，并且堆也无法再扩展时，将会抛出OutOfMemoryError异常

## 方法区(各线程共享)

方法区用于存储已被虚拟机加载的类信息、常量、静态变量、即时编译器编译后的代码等数据

- 异常  
根据Java虚拟机规范的规定，当方法区无法满足内存分配需求时，将抛出 OutOfMemoryError异常

## 运行时常量池(方法区一部分)

常量池用于存放编译期生产的各种字面量和符号引用，这部分内容将在类加载后进入方法区的运行时常量池中存放

- 运行时常量池  
运行时常量池相对于Class文件常量池的另外一个重要特征是具备动态性，Java语言并不要求常量一定只有编译期才能产生，也就是并非预置入Class文件中常量池的内容才能进入方法区运行时常量池，运行期间也可能将新的常量放入池中，这种特性被开发人员利用得比较多的便是String类的intern()方法
- 异常  
当常量池无法再申请到内存时会抛出OutOfMemoryError异常

## 直接内存

- NIO类  
NIO类，引入了一种基于通道与缓冲区的I/O方式，它可以使用Native函数库直接分配堆外内存，然后通过一个存储在Java堆中的DirectByteBuffer对象作为这块内存的引用进行操作。这样能在一些场景中显著提高性能，因为避免了在Java堆和Native堆中来回复制数据
- 异常  
服务器管理员有时会忽略直接内存，使得各个内存区域总和大于物理内存限制，从而导致动态扩展时出现OutOfMemoryError异常

## 对象

### 对象的创建

类加载 --> 分配内存 -->

- 分配内存
  - 堆中的内存是规整的

*指针碰撞*

假设java堆中的内存是规整的，所有用过的内存都放在一边，空闲的内存放在另一边，中间放着一个指针作为分界点的指示器，那所分配内存就仅仅是把那个指针向空闲空间那边挪动一段与对象大小相等的距离，这种分配方式称为“指针碰撞”
  - 堆中的内存不是规整的

*空闲列表*

如果Java堆中的内存并不是规整的，已使用的内存和空闲的内存相互交错，那就没有办法简单地进行指针碰撞了，虚拟机就必须维护一个列表，记录上哪些内存块是可用的，在分配的时候从列表中找到一块足够大的空间划分给对象实例，并更新列表上的记录，这种分配方式称为“空闲列表”
  - 分配方式的选择

选择哪种分配方式由Java堆是否规整决定，而Java堆是否规整又由所采用的垃圾收集器是否带有压缩整理功能决定。因此，在使用Serial、ParNew等带Compact过程的收集器时，系统采用的分配算法是指针碰撞，而使用CMS这种基于Mark-Sweep算法的收集器时，通常采用空闲列表
  - 分配内存时的线程安全问题(解决方案)
    1. 对分配内存空间的动作进行同步处理
    2. 把内存分配的动作按照线程划分在不同的空间之中进行，即每个线程在Java堆中预先分配一小块内存，称为本地线程分配缓冲。哪个线程要分配内存，就在哪个线程的TLAB上分配，只有TLAB用完并分配新的TLAB时，才需要同步锁定
- 内存空间初始化
  - 对象头

内存空间分配完成后会初始化为0，接下来就是填充对象头

## 对象的内存布局

对象在内存中存储的布局可以分为3块区域：对象头、实例数据和对齐填充

- 对象头

包括两部分信息

  1. 用于存储对象自身的运行时数据

如哈希码、GC分代年龄、锁状态标志、线程持有的锁、偏向线程ID、偏向时间戳等
  2. 类型指针(对象指向它的类元数据的指针，虚拟机通过这个指针来确定这个对象是哪个类的实例)

*对象为一个数组时*

对象头中必须有一块用于记录数组长度的数据，因为虚拟机可以通过普通Java对象的元数据信息确定Java对象的大小，但是从数据的元数据中却无法确定数组的大小
- 实例数据

程序代码中所定义的各种类型的字段内容(包含父类继承下来的和子类中定义的)
- 对齐填充

不是必然需要，主要是占位，保证对象大小是某个字节的整数倍

## 对象的访问定位

Sun HotSpot使用的是直接指针访问

- 句柄访问
  - reference中存储对象的句柄地址  
句柄中包含了对象实例数据与类型数据各自的具体地址信息
- 直接指针访问
  - reference中存储的直接就是对象地址
- 两种对象访问方式各自的优点
  - 句柄访问  
reference中存储的是稳定的句柄地址，在对象被移动时(GC)只会改变句柄中的实例数据指针，而reference本身不需要修改
  - 直接指针访问  
速度更快，它节省了一次指针定位的时间开销，由于对象的访问在Java中非常频繁，因此节省的时间也比较可观

## 垃圾收集器与内存分配策略

---

### 概述

GC已自动化，为什么要去了解

当需要排查各种内存溢出、内存泄漏问题时，当垃圾收集成为系统达到更高并发量的瓶颈时，我们就需要对这些“自动化”的技术实施必要的监控和调节

### 运行时区域

程序计数器、虚拟机栈、本地方法栈 3 个区域随线程生灭(因为是线程私有)，栈中的栈帧随着方法的进入和退出而有条不紊地执行着出栈和入栈操作。而 Java 堆和方法区则不一样，一个接口中的多个实现类需要的内存可能不一样，一个方法中的多个分支需要的内存也可能不一样，我们只有在程序处于运行期才知道那些对象会创建，这部分内存的分配和回收都是动态的，垃圾回收器所关注的就是这部分内存

### 判断对象是否死亡(重点)

垃圾收集器在对堆进行回收前，第一件事情就是要确定这些对象之中哪些还“存活”着，哪些已经“死去”（即不可能再被任何途径使用的对象）

### 引用技术算法

- 算法思想  
给对象中添加一个引用计数器，每当有一个地方引用它时，计数器值就加1；当引用失效时，计数器值就减1；任何时刻计数器为0的对象就是不可能再被使用的
- 好处  
实现简单，判定效率也很高
- 缺点
  1. 很难解决对象之间的相互循环引用的问题
  2. 当两个对象互相指向对方，而且这两个对象再无任何其他引用时，实际上这两个对象已经不可能在被访问，但因为互相引用，导致他们的引用计数都不为0，所以引用计数算法无法通知GC收集器回收它们

## 可达性分析算法

- 算法思想

通过一系列称为 "GC Roots" 的对象作为起始点，从这些节点开始向下搜索，搜索所走过的路径称为引用链，当一个对象到GC Roots没有任何引用链相连(就是说从GC Roots到这个对象不可达)时，则证明此对象是不可用的

- 可作为GC Roots的对象

- 虚拟机栈(栈帧中的本地变量表)中引用的对象
- 方法区中类静态属性引用的对象
- 方法区中常量引用的对象
- 本地方法栈中JNI(即一般说的Native方法)引用的对象

- 再谈引用

我们希望能描述这样一类对象：当内存空间还足够时，则能保留在内存之中；如果内存空间在进行垃圾收集后还是非常紧张，则可以抛弃这些对象

引用被分为下面几类：

- 强引用
  - 指在程序代码中普遍存在的
  - 类似 "Object obj = new Object()" 这类引用，只要强引用还存在，垃圾收集器永远不会回收掉被引用的对象
- 软引用
  - 用来描述一些还有用但并非必需的对象
  - 对于软引用关联着的对象，在系统将要发生内存溢出异常之前，将会把这些对象列进回收范围之中进行第二次回收。如果这次回收还没有足够的内存，才会抛出内存溢出异常
- 弱引用
  - 也是用来描述非必需对象，但比软引用更弱一些
  - 被弱引用关联的对象只能生存到下一次垃圾收集发生之前，当垃圾收集器工作时，无论当前内存是否足够，都会回收掉只被弱引用关联的对象
- 虚引用
  - 也被称为幽灵引用或者幻影引用，它是最弱的一种引用关系
  - 一个对象是否有虚引用的存在，完全不会对其生存时间构成影响，也无法通过虚引用来取得一个对象实例
  - 为一个对象设置虚引用关联的唯一目的就是能在这个对象被收集器回收时收到一个系统通知

- 生存还是死亡

- 即使在可达性分析算法中不可达的对象，也并非是非死不可的，这时候它们暂时处于“缓刑”阶段，要真正宣告一个对象死亡，至少要经历两次标记过程
- 如果对象在进行可达性分析后发现没有与GC Roots相连接的引用链，那它将会被第一次标记并且进行一次筛选，筛选的条件是此对象是否有必要执行finalize()方法。当对象没有覆盖finalize()方法，或者finalize()方法已经被虚拟机调用过，虚拟机将这两种情况都视为“没有必要执行”
- 如果这个对象被判定为有必要执行 finalize() 方法，那么这个对象竟会放置在一个叫做 F-Queue 的队列中，并在稍后由一个由虚拟机自动建立的、低优先级的 Finalizer 线程去执行它。这里所谓的“执行”是指虚拟机会触发这个方法，并不承诺或等待他运行结束。finalize() 方法是对象逃脱死亡命运的最后一次机会，稍后 GC 将对 F-Queue 中的对象进行第二次小规模标记，如果对象要在 finalize() 中成功拯救自己——只要重新与引用链上的任何一个对象建立关联即可
- finalize() 方法只会被系统自动调用一次

- 回收方法区 在堆中，尤其是在新生代中，一次垃圾回收一般可以回收 70% ~ 95% 的空间，而永久代的垃圾收集效率远低于此

- 永久代垃圾回收
  - 永久代垃圾回收主要两部分内容：废弃的常量和无用的类
    - 判断废弃常量：一般是判断没有该常量的引用
    - 判断无用的类：要以下三个条件都满足
      1. 该类所有的实例都已经回收，也就是 Java 堆中不存在该类的任何实例
      2. 加载该类的 `ClassLoader` 已经被回收
      3. 该类对应的 `java.lang.Class` 对象没有任何地方被引用，无法在任何地方通过反射访问该类的方法
- 垃圾收集算法
  - 标记-清除算法
    - 分为两个阶段：
      1. 标记：标记出所有需要回收的对象
      2. 清除：标记完成后统一回收所有被标记的对象
    - 不足
      1. 效率问题：标记和清除效率都不高
      2. 空间问题：标记清除之后会产生大量不连续的内存碎片
  - 复制算法

## 虚拟机类加载机制

---

### 类加载的时机

#### 类的生命周期

1. 加载
2. 验证
3. 准备
4. 解析
5. 初始化
6. 使用
7. 卸载

#### 需立即对类进行"初始化"的5种情况

1. 遇到到`new`、`getstatic`、`putstatic`或`invokestatic`这4条字节码指令时，如果类没有进行过初始化，则需要先触发其初始化。
  2. 使用`java.lang.reflect`包的方法对类进行反射调用的时候，如果类没有进行过初始化，则需要先触发其初始化
  3. 当初始化一个类的时候，如果发现其父亲还没有进行过初始化，则需要先触发其父类的初始化
  4. 当虚拟机启动时，用户需要指定一个要执行的主类(包含`main()`方法的那个类)，虚拟机会先初始化这个主类
  5. 当使用JDK1.7的动态语言支持时，如果一个`java.lang.invoke.MethodHandle`实例最后的解析结果果`REF_getStatic`、`REF_putStatic`、`REF_invokeStatic`的方法句柄，并且这个方法句柄所对应的类没有进行过初始化，则需要先触发其初始化 这5种情况的行为数据对一个类进行主动引用
- 对于静态字段，只有直接定义这个字段的类才会被初始化，因此通过其子类来引用父类中定义的静态字段，只会触发父类的初始化而不会触发子类的初始化

- 被动引用例子

当调用SubClass.value时只会输出"SuperClass init!"

对于静态字段，只有直接定义这个字段的类才会被初始化，因此通过其子类来引用父类中定义的静态字段，只会触发父类的初始化而不会触发子类的初始化

```
public class SuperClass {
    static {
        System.out.println("SuperClass init!");
    }
    public static int value = 1127;
}
```

```
public class SubClass extends SuperClass {
    static {
        System.out.println("SubClass init!");
    }
}
```

//当调用ConstClass.HELLOWORLD时并没有输出"ConstClass init!"

//因为虽然引用了ConstClass类中的常量HELLOWORLD，但其实在编译阶段通过常量传播优化，已经将此常量的值存到了main方法所在类的常量池中，以后main方法所在类对常量ConstClass.HELLOWORLD的引用实际都被转化为main方法所在类对自身常量池的引用了

```
public class ConstClass {
    static {
        System.out.println("ConstClass init!");
    }
    public static final String HELLOWORLD = "hello world!"
}
```

```
public class NotInitialization {
    public static void main(String[] args) {
        System.out.println(SubClass.value);
        /**
         * output : SuperClass init!
         *
         * 通过子类引用父类的静态对象不会导致子类的初始化
         * 只有直接定义这个字段的类才会被初始化
         */

        SuperClass[] sca = new SuperClass[10];
        /**
         * output :
         *
         * 通过数组定义来引用类不会触发此类的初始化
         * 虚拟机在运行时动态创建了一个数组类
         */

        System.out.println(ConstClass.HELLOWORLD);
        /**
         * output :
         *
         * 常量在编译阶段会存入调用类的常量池中，本质上并没有直接引用到定义类常量的类，
         * 因此不会触发定义常量的类的初始化。
         * “hello world” 在编译期常量传播优化时已经存储到 NotInitialization 常量池中
         */
    }
}
```

```
了。  
        */  
    }  
}
```

## 类加载的过程

### 加载

虚拟机需要完成下面三件事

1. 通过一个类的全限定名来获取定义此类的二进制字节流
2. 将这个字节流所代表的静态存储结构转化为方法区的运行时数据结构
3. 在内存中生成一个代表这个类的`java.lang.Class`对象，作为方法区这个类的各种数据的访问入口

### 验证

验证时连接阶段的第一步，这一阶段的目的是为了确保Class文件的字节流中包含的信息符合当前虚拟机的要求，并且不会危害虚拟机自身的安全

- 文件格式验证  
验证字节流是否符合Class文件格式的规范，并且能被当前版本的虚拟机处理
- 元数据验证  
对字节码描述的信息进行语义分析，以保证其描述的信息符合Java语言规范的要求
- 字节码验证  
通过数据流和控制流分析，确定程序语义是合法的、符合逻辑的
- 符号引用验证  
对类自身以外(常量池中的各种符号引用)的信息进行匹配性校验

### 准备

准备阶段是正式为类变量(被`static`修饰的变量)分配内存并设置类变量初始值的阶段，这些变量所使用的内存都将在方法区中进行分配

### 解析

解析阶段是虚拟机将常量池内的符号引用替换为直接引用的过程

### 初始化

在准备阶段，变量已经赋过一次系统要求的初始值，而在初始化阶段，则根据程序员通过程序制定的主观计划去初始化类变量和其他资源

## 类加载器

当写完一个java文件的时候，编译器会把它编译成一个由字节码组成的class文件，当程序运行时，JVM会首先寻找包含有`main()`方法的类，把这个class文件中的字节码数据读入进来，转化为JVM中运行是对应的Class对象。执行这个动作的，就叫类加载器

### 启动类加载器



这个类加载器负责将一些核心的，被JVM识别的类加载进来，用C++实现，与JVM是一体的

- 扩展类加载器  
这个类加载器用来加载Java的扩展库
- 应用程序类加载器  
用于加载我们自己定义编写的类
- 用户自己实现的加载器  
当实际需要自己掌控类加载过程时才会用到，一般没有用到

## 双亲委派模型

- 层次关系  
双亲委派模型要求除了顶层的启动类加载器外，其余的类加载器都应当有自己的父类加载器
- 工作过程  
如果一个类加载器收到了类加载的请求，它首先不会自己去尝试加载这个类，而是把这个请求委派给父类加载器去完成，每一个层次的类加载器都是如此，因此所有的加载请求最终都应该传送到顶层的启动类加载器中，只有当父加载器反馈自己无法完成这个加载请求(它的搜索范围中没有找到所需的类)时，子加载器才会尝试自己去加载
- 特点 Java类随着它的类加载一起具备了一种带有优先级的层次关系  
例如类`java.lang.Object`，它存放在`rt.jar`之中，无论哪一个类加载器要加载这个类，最终都是委派给处于模型最顶端的启动类加载器进行加载，因此`Object`类在程序的各种类加载器环境中都是同一个类
- 实现  
先检查是否已经被加载过，若没有加载则调用父加载器的`loadClass()`方法，若父加载器为空则默认使用启动类加载器作为父加载器。如果父类加载失败，抛出`ClassNotFoundException`异常后，再调用自己的`findClass()`方法进行加载

# 虚拟机字节码执行引擎

---

## 运行时栈帧结构

栈帧是用于支持虚拟机进行方法调用和方法执行的数据结构，它是虚拟机运行时数据区中的虚拟机栈的栈元素

栈帧存储了方法的局部变量表、操作数栈、动态链接和方法返回地址等信息

每一个方法从调用开始至执行完成的过程，都对应着一个栈帧在虚拟机栈里面从入栈到出栈的过程

栈帧包括了局部变量表、操作数栈、动态链接、方法返回地址和一些额外的附加信息

一个栈帧需要分配多少内存，不会受到程序运行期变量数据的影响，而仅仅取决于具体的虚拟机实现

### 局部变量表

是一组变量值存储空间，用于存放方法参数和方法内部定义的局部变量

局部变量表的容量以变量槽为最小单位

一个变量槽可以存放一个32位以内的数据类型，Java中占用32位以内的数据类型有`boolean`、`byte`、`char`、`short`、`int`、`float`、`reference`和`returnAddress`(少见) 8种类型

- **reference**类型

表示对一个对象实例的引用，虚拟机规范即没有说明它的长度，也没有明确指出这种引用应有怎样的结构

1. 从此引用中直接或间接地查找到对象在**Java**堆中的数据存放的起始地址索引
2. 此引用中直接或间接地查找到对象所属数据类型在方法区中的存储的类型信息，否则无法实现**Java**语言规范中定义的语法约束约束