


第6讲 | 动态代理是基于什么原理？

2018-05-17 杨晓峰





第6讲 | 动态代理是基于什么原理？

杨晓峰

- 00:00 / 11:24

编程语言通常有各种不同的分类角度，动态类型和静态类型就是其中一种分类角度，简单区分就是语言类型信息是在运行时检查，还是编译期检查。

与其近似的还有一个对比，就是所谓强类型和弱类型，就是不同类型变量赋值时，是否需要显式地（强制）进行类型转换。

那么，如何分类Java语言呢？通常认为，Java是静态的强类型语言，但是因为提供了类似反射等机制，也具备了部分动态类型语言的能力。

言归正传，今天我要问你的问题是，谈谈Java反射机制，动态代理是基于什么原理？

典型回答 **拼课微信：699250**

反射机制是Java语言提供的一种基础功能，赋予程序在运行时自省（Introspect，官方用语）的能力。通过反射我们可以直接操作类或者对象，比如获取某个对象的类定义，获取类声明的属性和方法，调用方法或者构造对象，甚至可以运行时修改类定义。

动态代理是一种方便运行时动态构建代理、动态处理代理方法调用的机制，很多场景都是利用类似机制做到的，比如用来包装RPC调用、面向切面的编程（AOP）。

实现动态代理的方式很多，比如JDK自身提供的动态代理，就是主要利用了上面提到的反射机制。还有其他的实现方式，比如利用传说中更高性能的字节码操作机制，类似ASM、cglib（基于ASM）、Javassist等。

考点分析

这个题目给我的第一印象是稍微有点诱导的嫌疑，可能会下意识地以为动态代理就是利用反射机制实现的，这么说也不算错但稍微有些不全。功能才是目的，实现的方法有很多。总的来说，这道题目考察的是Java语言的另外一种基础机制：反射，它就像是一种魔法，引入运行时自省能力，赋予了Java语言令人意外的活力，通过运行时操作元数据或对象，Java可以灵活地操作运行时才能确定的信息。而动态代理，则是延伸出来的一种广泛应用于产品开发中的技术，很多繁琐的重复编程，都可以被动态代理机制优雅地解决。

从考察知识点的角度，这道题涉及的知识点比较庞杂，所以面试官能够扩展或者深挖的内容非常多，比如：

- 考察你对反射机制的了解和掌握程度。
- 动态代理解决了什么问题，在你业务系统中的应用场景是什么？
- JDK动态代理在设计 and 实现上与cglib等方式有什么不同，进而如何取舍？

这些考点似乎不是短短一篇文章能够囊括的，我会在知识扩展部分尽量梳理一下。

知识扩展

1. 反射机制及其演进

对于Java语言的反射机制本身，如果你去看一下java.lang或java.lang.reflect包下的相关抽象，就会有一个很直观的印象了。Class、Field、Method、Constructor等，这些完全就是我们去操作类和对象的元数据对应。反射各种典型用例的编程，相信有太多文章或书籍进行过详细的介绍，我不再赘述了，至少你需要掌握基本场景编程，这里是官方提供的参考文档：<https://docs.oracle.com/javase/tutorial/reflect/index.html>。

关于反射，有一点我需要特意提一下，就是反射提供的AccessibleObject.setAccessible(boolean flag)。它的子类也大都重写了这个方法，这里的所谓accessible可以理解成修饰成员的public、protected、private，这意味着我们可以在运行时修改成员访问限制！

setAccessible的应用场景非常普遍，遍布我们的日常开发、测试、依赖注入等各种框架中。比如，在O/R Mapping框架中，我们为一个Java实体对象，运行时自动生成setter、getter的逻辑，这是加载或者持久化数据非常必要的，框架通常可以利用反射做这个事情，而不需要开发者手动写类似的重复代码。

另一个典型场景就是绕过API访问控制。我们日常开发时可能被迫要调用内部API去做些事情，比如，自定义的高性能NIO框架需要显式地释放DirectBuffer，使用反射绕开限制是一种常见办法。

但是，在Java 9以后，这个方法的使用可能会存在一些争议，因为Jigsaw项目新增的模块化系统，出于强封装性的考虑，对反射访问进行了限制。Jigsaw引入了所谓Open的概念，只有当被反射操作的模块和指定的包对反射调用者模块Open，才能使用setAccessible；否则，被认为是非法（illegal）操作。如果我们的实体类是定义在模块里面，我们需要在模块描述符中明确声明：

```
module MyEntities {  
    // Open for reflection  
    opens com.mycorp to java.persistance;  
}
```

因为反射机制使用广泛，根据社区讨论，目前，Java 9仍然保留了兼容Java 8的行为，但是很有可能在未来版本，完全启用前面提到的针对setAccessible的限制，即只有当被反射操作的模块和指定的包对反射调用者模块Open，才能使用setAccessible，我们可以使用下面参数显式设置。

```
--illegal-access={ permit | warn | deny }
```

## 2 动态代理

前面的问题到了动态代理，我们一起看看，它到底是解决什么问题？

首先，它是一个代理机制。如果熟悉设计模式中的代理模式，我们会知道，代理可以看作是对调用目标的一个包装，这样我们对目标代码的调用不是直接发生的，而是通过代理完成。其实很多动态代理场景，我认为也可以看作是装饰器（Decorator）模式的应用，我会在后面的专栏设计模式主题予以补充。

通过代理可以让调用者与实现者之间解耦。比如进行RPC调用，框架内部的寻址、序列化、反序列化等，对于调用者往往是没有太大意义的，通过代理，可以提供更加友善的界面。

代理的发展经历了静态到动态的过程，源于静态代理引入的额外工作。类似早期的RMI之类古董技术，还需要rmic之类工具生成静态stub等各种文件，增加了很多繁琐的准备工作，而这又和我们的业务逻辑没有关系。利用动态代理机制，相应的stub等类，可以在运行时生成，对应的调用操作也是动态完成，极大地提高了我们的生产力。改进后的RMI已经不再需要手动去准备这些了，虽然它仍然是相对古老落后的技术，未来也许会逐步被移除。

这么说可能不够直观，我们可以看JDK动态代理的一个简单例子。下面只是加了一句print，在生产系统中，我们可以轻松扩展类似逻辑进行诊断、限流等。

```
public class MyDynamicProxy {  
    public static void main (String[] args) {  
        HelloImpl hello = new HelloImpl();  
        MyInvocationHandler handler = new MyInvocationHandler(hello);  
        // 构造代码实例  
        Hello proxyHello = (Hello) Proxy.newProxyInstance(HelloImpl.class.getClassLoader(), HelloImpl.class.getInterfaces(), handler);  
        // 调用代理方法  
        proxyHello.sayHello();  
    }  
}  
  
interface Hello {  
    void sayHello();  
}  
  
class HelloImpl implements Hello {  
    @Override  
    public void sayHello() {  
        System.out.println("Hello World");  
    }  
}  
  
class MyInvocationHandler implements InvocationHandler {  
    private Object target;  
    public MyInvocationHandler(Object target) {  
        this.target = target;  
    }  
    @Override  
    public Object invoke(Object proxy, Method method, Object[] args)  
        throws Throwable {  
        System.out.println("Invoking sayHello");  
        Object result = method.invoke(target, args);  
        return result;  
    }  
}
```

上面的JDK Proxy例子，非常简单地实现了动态代理的构建和代理操作。首先，实现对应的InvocationHandler；然后，以接口Hello为纽带，为被调用目标构建代理对象，进而应用程序就可以使用代理对象间接运行调用目标的逻辑，代理为应用插入额外逻辑（这里是println）提供了便利的入口。

从API设计和实现的角度，这种实现仍然有局限性，因为它是以接口为中心的，相当于添加了一种对于被调用者没有太大意义的限制。我们实例化的是Proxy对象，而不是真正的被调用类型，这在实践中还是可能带来各种不便和能力退化。

如果被调用者没有实现接口，而我们还是希望利用动态代理机制，那么可以考虑其他方式。我们知道Spring AOP支持两种模式的动态代理，JDK Proxy或者cglib，如果我们选择cglib方式，你会发现对接口的依赖被克服了。

cglib动态代理采取的是创建目标类的子类的方式，因为它是子类化，我们可以达到近似使用被调用者本身的效果。在Spring编程中，框架通常会处理这种情况，当然我们也可以[显式指定](#)。关于类似方案的实现细节，我不再详细讨论了。

那我们在开发中如何选择呢？我来简单对比下两种方式各自优势。

JDK Proxy的优势：

- 最小化依赖关系，减少依赖意味着简化开发和维护，JDK本身的支持，可能比cglib更加可靠。

- 平滑进行JDK版本升级，而字节码类库通常需要进行更新以保证在新版Java上能够使用。
- 代码实现简单。

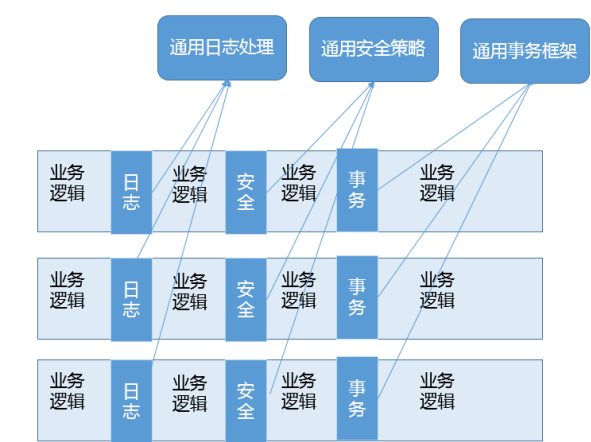
基于类似cglib框架的优势：

- 有的时候调用目标可能不便实现额外接口，从某种角度看，限定调用者实现接口是有些侵入性的实践，类似cglib动态代理就没有这种限制。
- 只操作我们关心的类，而不必为其他相关类增加工作量。
- 高性能。

另外，从性能角度，我想补充几句。记得有人曾经得出结论说JDK Proxy比cglib或者Javassist慢几十倍。坦白说，不去争论具体的benchmark细节，在主流JDK版本中，JDK Proxy在典型场景可以提供对等的性能水平，数量级的差距基本上不是广泛存在的。而且，反射机制性能在现代JDK中，自身已经得到了极大的改进和优化，同时，JDK很多功能也不完全是反射，同样使用了ASM进行字节码操作。

我们在选型中，性能未必是唯一考量，可靠性、可维护性、编程工作量等往往是更主要的考虑因素，毕竟标准类库和反射编程的门槛要低得多，代码量也是更加可控的，如果我们比较下不同开源项目在动态代理开发上的投入，也能看到这一点。

动态代理应用非常广泛，虽然最初多是因为RPC等使用进入我们视线，但是动态代理的使用场景远远不仅如此，它完美符合Spring AOP等切面编程。我在后面的专栏还会进一步详细分析AOP的目的和能力。简单来说它可以看作是对OOP的一个补充，因为OOP对于跨越不同对象或类的分散、纠缠逻辑表现力不够，比如在不同模块的特定阶段做一些事情，类似日志、用户鉴权、全局性异常处理、性能监控，甚至事务处理等，你可以参考下面这张图。



AOP通过（动态）代理机制可以让开发者从这些繁琐事项中抽身出来，大幅度提高了代码的抽象程度和复用度。从逻辑上来说，我们在软件设计和实现中的类似代理，如Facade、Observer等很多设计目的，都可以通过动态代理优雅地实现。

今天我简要回顾了反射机制，谈了反射在Java语言演进中正在发生的变化，并且进一步探讨了动态代理机制和相关的切面编程，分析了其解决的问题，并探讨了生产实践中的选择考量。

一课一练

关于今天我们讨论的题目你做到心中有数了吗？留一道思考题给你，你在工作中哪些场景使用到了动态代理？相应选择了什么实现技术？选择的依据是什么？

请在留言区写写你对这个问题的思考，我会选出经过认真思考的留言，送给你一份学习鼓励金，欢迎你与我一起讨论。

你的朋友是不是也在准备面试呢？你可以“请朋友读”，把今天的题目分享给好友，或许你能帮到他。



肖一林

2018-05-17

提一些建议：应该从两条线讲这个问题，一条从代理模式，一条从反射机制。不要老担心篇幅限制讲不清问题，废话砍掉一些，深层次的内在原理多讲些（比如asm），容易自学的扩展知识可以用链接代替

代理模式（通过代理静默地解决一些业务无关的问题，比如远程、安全、事务、日志、资源关闭.....让应用开发者可以只关心他的业务）

静态代理：事先写好代理类，可以手工编写，也可以用工具生成。缺点是每个业务类都要对应一个代理类，非常不灵活。

动态代理：运行的自动生成代理对象。缺点是生成代理代理对象和调用代理方法都要额外花费时间。

JDK动态代理：基于Java反射机制实现，必须要实现了接口的业务类才能用这种办法生成代理对象。新版本也开始结合ASM机制。

cglib动态代理：基于ASM机制实现，通过生成业务类的子类作为代理类。

Java 发射机制的常见应用：动态代理（AOP、RPC）、提供第三方开发者扩展能力（Servlet容器，JDBC连接）、第三方组件创建对象（DI）.....

我水平比较菜，希望多学点东西，希望比免费知识层次更深些，也不光是为了面试，所以提提建议。

作者回复

2018-05-18

谢谢反馈，类似ASM这种字节码操纵是有单独章节覆盖的，前面基础篇有个整体印象，免得陷入细节；Java内部动态生成还有其他领域，比如Lambda实现机制，个人认为一起分析会连贯一些

公号-Java大后端

2018-05-17

反射与动态代理原理

1 关于反射

反射最大的作用之一就在于我们可以不在编译时知道某个对象的类型，而在运行时通过提供完整的“包名+类名.class”得到。注意：不是在编译时，而是在运行时。

功能：

- 在运行时能判断任意一个对象所属的类。
- 在运行时能构造任意一个类的对象。
- 在运行时判断任意一个类所具有的成员变量和方法。
- 在运行时调用任意一个对象的方法。

说大白话就是，利用Java反射机制我们可以加载一个运行时才得知名称的class，获悉其构造方法，并生成其对象实体，能对其fields设值并唤起其methods。

应用场景：

反射技术常用在各类通用框架开发中。因为为了保证框架的通用性，需要根据配置文件加载不同的对象或类，并调用不同的方法，这个时候就会用到反射——运行时动态加载需要加载的对象。

特点：

由于反射会额外消耗一定的系统资源，因此如果不需要动态地创建一个对象，那么就不需要用反射。另外，反射调用方法时可以忽略权限检查，因此可能会破坏封装性而导致安全问题。

2 动态代理

为其他对象提供一种代理以控制对这个对象的访问。在某些情况下，一个对象不适合或者不能直接引用另一个对象，而代理对象可以在两者之间起到中介的作用（可类比房屋中介，房东委托中介销售房屋、签订合同等）。

所谓动态代理，就是实现阶段不用关心代理谁，而是在运行阶段才指定代理哪个对象（不确定性）。如果是自己写代理类的方式就是静态代理（确定性）。

组成要素：

(动态)代理模式主要涉及三个要素：

其一：抽象类接口

其二：被代理类（具体实现抽象接口的类）

其三：动态代理类：实际调用被代理类的方法和属性的类

实现方式：

实现动态代理的方式很多，比如 JDK 自身提供的动态代理，就是主要利用了反射机制。还有其他的实现方式，比如利用字节码操作机制，类似 ASM、CGLIB（基于 ASM）、Javassist 等。

举例：常可采用由JDK提供的动态代理接口InvocationHandler来实现动态代理类。其中invoke方法是该接口定义必须实现的，它完成对真实方法的调用。通过InvocationHandler接口，所有方法都由该Handler来进行处理，即所有被代理的方法都由InvocationHandler接管实际的处理任务。此外，我们常可以在invoke方法实现中增加自定义的逻辑实现，实现对被代理类的业务逻辑无侵入。

郭峰

2018-05-17

没涉及到原理，proxy到底是如何实现的，运行时拦截？cglib是编译时对类进行织入？要是更细一些就好了

刘方杰

2018-05-18

哎，阅读越来越困难了，我离核心是不是太远了。

jimforcode

2018-05-17

好像和啥原理没啥关系吧，总结来说就是jdk 自身的反射机制或用第三方库，哪哪看到的都这样说，一笔带过

作者回复

2018-05-17

谢谢反馈，字节码操作、运行时拦截、加载期编织、Java agent等，会和Aop单独介绍，那些内容不是几句话说得完

jimforcode

2018-05-17

比较期待大神讲讲动态代理的原理是什么，对性能会带来什么影响，有没有什么问题需要规避，谢谢

曹铮

2018-05-17

先回答问题：  
99%的Java程序员应该都间接使用了AOP。自己项目里直接编写的，比如调用追踪，通用日志，自动重试，反射和AOP真是双刀剑效果拔群的技术。从MVC开始约定胜过配置的开发现念大行其道，ORM自动映射，plugin模式，到现在的spring + boot +cloud的declarative编程大量基于此实现，可以说没有反射和AOP就没有Java的今天。反面就是，自己想进行定制化的改造封装真挺苦涩  
再提问题：  
1. 听到过个说法，反射慢因为一是获取Field，Method等要进行元数据的查找，这里有字符串匹配操作。二是Invoke时，要进行一些安全性上的检查。这种说法对么？JVM在解释执行的时候就不做一些操作内存操作的检查了么？如果不对，那原因是什么？还有没有其他？  
2. 以前写C#的，里面可以拼表达式树，运行时生成一个函数（不需要有对象），理论上性能是和手写代码一样的，但可以缓存起来。这解决的是手写中间代码太难的问题。请问Java有这种类似的功能嘛？

作者回复

2018-05-18

1. 基本如此；反射是相对保证类型安全的，我觉得要比较也是和methodhandle之类对比，那个更是接近jvm内部的黑盒，性能更好  
2. 你是说lambda？也是需要jvm生成call site，然后invokedynamic之类调用，所以首次调用开销明显，C#不了解，不过动态生成的感觉都是如此吧；这东西目前没有cache，如果你说的是存储在文件系统；未来，嘿嘿.....

这些太零碎，说过了会有单独章节介绍，不然没基础的就晕了，还用不上

ai~ wen

2018-05-21

对于我这初学者，读着就懵了

TWO STRINGS

2018-05-17

老师可以在分享结束时推荐一些好的文章，书籍甚至演讲之类的么？

作者回复

2018-05-18

没问题，喜欢底层，去查查JavaOne，FOSDEM，jvmsummit等

灰飞灰猪不会灰飞.烟灭

2018-05-17

cglib是怎么实现对目标对象的拦截的呢？

作者回复

2018-05-18

计划单独介绍

言志

2018-05-17

采用cglib方式的动态代理还有个缺点，不能应用到被代理对象的final方法上。  
我在多数数据项目中自动切换数据源功能用到了

云学

2018-06-09

看了好多篇文章，总体感觉是比较累，无论读者是否具有java背景，都应该让他看懂，而不是越看越糊涂，疑问反而更多了

作者回复

2018-06-10

非常感谢，读者基础不同，我尽量兼顾并增加基础的介绍，因为也有反馈希望可以更全面、深入...  
有好的建议请不吝赐教

bigfish

2018-07-08

本来资质愚笨，看不懂很多东西的原理,想进来学习一下，jdk动态代理的原理和cglib代理原理等一些原理性的东西（其他章节也是如此），发现听到原理性的东西不多都是一带而过,其实您做的课件针对很多点都是一带而过。听到某个名词一下来了兴趣继续一听结束了,我们都知道Java很大很多可以研究的，其实我们想听的很多是一些点的原理，讲完原理在结合实际应用阐述一下,也许我们就会有些豁然明朗的感觉,希望能理解一下！！

作者回复

2018-07-11

谢谢反馈，后面类加载章节介绍了两者底层机制，照顾下不同基础的读者

蒙奇D路飞

2018-05-28

感觉细节层面缺少具体描述，希望后续对底层原理的描述更细致些~

作者回复

2018-05-28

谢谢反馈，有章节介绍类似字节码操纵之类底层技术，照顾不同基础

雷贤

2018-05-22

何为反射？  
反射是指计算机程序在运行时可以访问、检测和修改它本身状态或行为的一种能力。比喻来说，反射就是程序在运行的时候能够“观察”并且修改自己的行为。文章中提到了“自省”它与反射是有区别的，“自省”机制仅指程序在运行时对自身信息（元数据）的检测，而反射机制不仅包括要能在运行时对程序自身信息进行检测，还要求程序能进一步根据这些信息改变程序状态或结构。  
通过反射的运行时可以访问、检测和修改自身状态的特性，也就出现了动态代理。  
一个小自对反射的使用场景（受限于业务能力）  
在一个程序中两个类A、B，整个程序的运行都由B类来承担，A类的所要做的工作是满足B的要求，如果在没有反射这种机制的条件下实现这个程序就会，既要吧B类中的对整个程序运行逻辑进行编写，还要回到A类中对B的工作进行人工手动辅助，这时只要B的任何改动都会影响到A。  
若在有了反射或代理之后A的代码就可以自己去访问、检测B，从而自动的修改A自身的状态来辅助B。

作者在评论中提到实现后续单独说明，我想到时就可以知道：代码是如何实现程序在运行时的访问、检测和行为的修改了。

1024

2018-05-17

为什么JDK Proxy是基于反射实现的呢？这其中有啥考量呢？为什么不基于其他呢？

作者回复

2018-05-18

我的理解，核心功能需要最小依赖关系，性能也不错

Allen

2018-06-27

文章可以推荐一些关于java的一些前沿技术的网站，想提高一下技术敏感性，请不吝赐教！

沈琦斌	2018-06-26
老师，我有个疑问。一些ORM框架比如mybatis，我们在使用过程中已经指定如何进行映射的时候，为什么不直接new一个类对象，然后调用setter方法赋值，而要用动态代理呢？这样设计的好处是什么？水平比较菜，还请谅解，谢谢	
作者回复	2018-06-28
我理解是，不需要依赖目标类的实现，框架使用限制少些，例如，不再要求实现了setter	
沈琦斌	2018-06-26
老师，我有个疑问。一些ORM框架比如mybatis，我们在使用过程中已经指定如何进行映射的时候，为什么不直接new一个类对象，然后调用setter方法赋值，而要用动态代理呢？这样设计的好处是什么？水平比较菜，还请谅解，谢谢	
作者回复	2018-06-28
已回答	
chris	2018-06-13
之前项目中用到javaassist，在通用业务流程中埋了一个扩展点，通过xml配置具体实现类来做定制业务处理，大概是类加载时会把对应制代码生成字节码织入进去，老师可以讲讲javaassist原理吗？	
作者回复	2018-06-13
我后面会介绍字节码操纵相关，还有类似jdk里怎么生成lambda的代码等，不过jdk是用的asm	
Jeffrey	2018-06-12
感觉还是书看少了，如果看过书，有过业务经验，再来看作者的分享，兴许会有另一种收获。	
Geek_c1b553	2018-05-30
mysql主从库通过aop动态切换	
George	2018-05-25
一个类的属性copy到另外一个类，使用cglib，先完成属性的拷贝	
岁月如歌	2018-05-22
代理的作用，使调用着和实现着之间解藕	
岁月如歌	2018-05-22
反射的作用：发射使得我们可以在运行的时候再去加载类，创建对象，并能访问到类中的私有属性和方法	
岁月如歌	2018-05-22
JAVA是静态的强类型，也就是说JAVA的类型在编译时确定（静态语言），不同类型之前的变量转换需要强制转换（强类型）	
岁月如歌	2018-05-22
不同类型变量赋值时，是否需要强制转换来区分语言是强类型还是弱类型	
岁月如歌	2018-05-22
语言类型信息在什么时候检查来区分语言是动态类型还是静态类型	
iiiDragon	2018-05-21
老师，多些代码示例。看的时候懂，看完又啥不懂	
Allife	2018-05-20
听不懂啊，怎么办.....该从哪里补课	
作者回复	2018-05-21
具体哪些方面？反射还是动态代理？动态代理可以看看AOP的书籍或文章	
字节码操纵，不用太担心，我本章没有写太多，就是担心基础不一样，搞糊涂了，毕竟真正开发中做这个是小众的，大多数不会直接用	
龙猫9527	2018-05-18
我们使用spring中的aspect来成权限控制，通过切点表达式和注解以更加细粒度的方式控制。	
作者回复	2018-05-18
AspectJ 我理解是确实能力更全面一些；比如，它可以compile-time，binary，还有用Javaagent作load-timeweaving；cglib 更是runtime weaving	
王磊	2018-05-18
反射加载的类和常规通过new对象加载的类，在资源消耗上有什么不同？我理解他们的效果是相同的，只是加载方式不同；加载到内存的区域也是相同的，metaspace。理解对吗？	
作者回复	2018-05-19
我理解没区别	
王磊	2018-05-18
关于如下的原文，有些疑惑。两种方式的实现是这么泾渭分明吗，例如jdk代理根据接口生成实现类，不也操作字节码了吗，为什么只强调如cglib：反过来，cglib既然能通过指定类生成子类，加载父类如果不是反射，那用的是什么？希望老师把这里澄清一下。	
实现动态代理的方式很多，比如JDK自身提供的动态代理，就是主要利用了上面提到的反射机制。还有其他的实现方式，比如利用传说中更高性能的字节码操作机制，类似ASM、cglib（基于	

ASM) 、 Javassist 等。’	
作者回复	2018-05-19
纠结于细节无助于说明问题，个人认为最好看主要路径实现机制	
jutsu	2018-05-17
老师的讲解和大神的评论学习很多	