# 一、dubbo的SPI

SPI是一种服务发现机制，JDK自带有服务发现机制，作用是自动加载位于指定文件夹下的配置文件，并加载其中所配置的类，例如：自动加载classpath下的数据库驱动类等。

Dubbo也有一套自身的服务发现机制，相较于jdk的SPI，dubboSPI支持懒加载，AOP、IOC，以及对异常的正确报出。

dubbo的SPI依赖于ExtensionLoader类，这个类同样是通过指定文件夹下所配置的文件扩展加载组件。

## 1.扩展流程



## 2.相关注解

dubbo通过三个注解@SPI、@Adaptive、@Activate对接口的实现类进行动态扩展。

@SPI(String value):

标注在接口上，代表这个接口的实现类可被扩展，value是该扩展接口默认实现类缩写。

@Adaptive(String[] value):

若标注在一个实现类上(不能标注在接口上)，表明这个类是一个自适应的装饰类，在调用getAdaptiveExtension()时不会动态生成编译装饰类。

若标注在接口的方法上，表明该方法需要被动态生成，在调用getAdaptiveExtension()时会动态生成并编译一个装饰类(类名以“$Adaptive结尾”)，没有标记@Adaptive注解的方法默认抛出异常。

其中的value为动态生成类需要过滤url的字段，会将@SPI中value和@Adaptive中的value对url先后进行过滤。

@Activate

@Adaptive是获取一个适配类，而@Activate则可以获取一批适配类，在impl上加上@Adaptive注解在调用Extension. getActivateExtension(URL url, String key)可以获得url满足条件的实现类。

## 3.ExtensionLoader相关api

ExtensionLoader是dubboSPI的核心类，相关API会对标注@SPI接口的类进行动态扩展

### ①getExtensionLoader(Interface) 静态方法



### ②ExtensionLoader的<init>

<init>中会保存interface.class，并且初始化内部的objectFactory字段，调用的是

ExtensionLoader.getExtensionLoader(ExtensionFactory.class).getAdaptiveExtension()【1】方法来初始化objectFactory

AdaptiveExtensionFactory被@Adaptive标记并且满足SPI扩展，所以【1】方法获得的其实是AdaptiveExtensionFactory实例。

### ③AdaptiveExtensionFactory的作用

AdaptiveExtensionFactory是装饰类(同样也是单例类)，其内部组合了SpringExtensionFactory和SpiExtensionFactory(分别提供了Spring/SPI的IOC)

### ④getExtension(key) 实例方法

通过getExtensionLoader(Interface)获得ExtensionLoader的实例之后，可以调用getExtension(key)获取SPI配置的key对应interface的实现类。



【1】中主要分成四类：

1）被@Adaptive标记的类

2）含有一个参数为interface的构造函数的实现类（wrapper类）

3）被@Activate标记的类

4）其他类与3）中的类

### ⑤getAdaptiveExtension() 实例方法

如果调用该方法的扩展点接口，没有实现类被@Adaptive标记，则会为其自动生成一个装饰类，装饰类会对url进行筛选出符合条件的属性，完成getExtension()的调用。

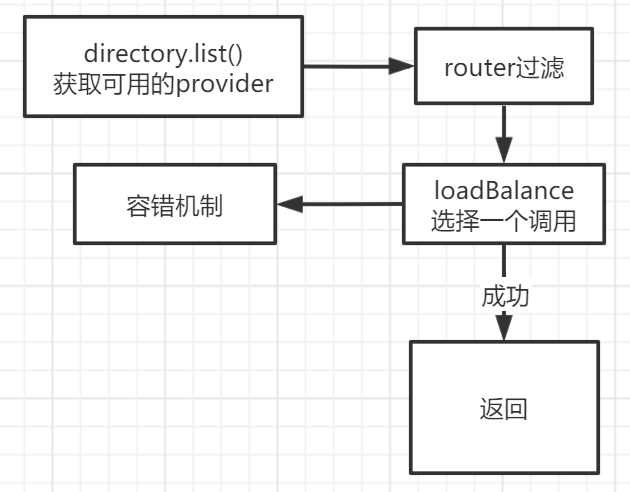


# 二、容器启动

Main.main(args)会自动启动配置在dubbo.properties中的容器，默认为spring容器。spring容器启动时依据spring的扩展机制，会对dubbo相关标签进行加载。

# 集群容错

集群容错全是在consumer执行的，包括路由选择(router)、负载均衡(loadBalance)、容错机制(cluster)，执行顺序如下所示



## 1.容错机制cluster

每个cluster可以伪装一个clusterInvoker给客户端调用，clusterInvoker中放了RegistryDirectory。

①Failover：失败重新loadbalance选择一个进行调用，可以设置retries次数选择重试的次数

②failfast：若失败，直接返回异常，不重试

③failsafe：若失败，返回null

④failback：若失败，返回空值，但是后台会调用一个线程池对错误重新定时调用

⑤forking：若失败，可以设置forks最大并行数

loadbalance选择出forks个provider放入线程池运行，默认同步调用，先运行完的线程将数据放入一个BlockingQueue，主线程调用blockingQueue.poll() timeout秒，返回先调用完的数据给用户。

⑥broadcast：广播调用所有可用的服务，若出现一个失败，则报错。不需要负载均衡

## 2.RegistryDirectory