基于J2EE的移动电商平台安全性研究

1. 绪论

电子商务是一种商业模式，它使得企业或个人能够通过网络进行交易。电子商务让企业在市场竞争上更加有优势，可以通过提供商品或服务得到更加快速高效的分销链。随着互联网的快速发展，电子商务这种媒介得以发展壮大。随着移动设备的普及，移动商务已经成为自己的市场。移动电商是“无线的电子商务”，它是指用户通过移动终端如手机、掌上电脑等设备访问网络，进行各种电子商务活动。近几年，移动电子商务平台数量在逐渐的增多，买家和卖家通过移动电子商务平台进行的交易额也在逐年呈现暴涨趋势。移动电商平台是建立在移动互联网基础上电子商务平台，像美团、饿了么、淘宝、拼多多都属于典型的移动电商平台。

J2EE能够快速、安全、可靠地构建基于分布式的移动电子商务系统，随着移动电子商务平台的不断发展，其体系结构也随着技术和时代的潮流发生了很大的变化，大致经历了三个阶段：

单点服务阶段：此阶段应用程序、数据库、文件都部署在一台服务器上，或者独立部署。

集群阶段：此阶段应用程序、数据库、文件分别部署，通过增加应用程序以及负载均衡的方式，将请求访问到不同的服务器上。

分布式阶段：此阶段将应用程序中各个业务应用都会使用到一些基本的业务服务，例如用户服务、订单服务、支付服务、安全服务。我们将这些服务从应用程序中抽离出来，利用分部式服务框架搭建分布式服务。

本文主要基于J2EE的移动电商平台安全性为研究目标。探讨目前基于J2EE的移动电商平台的系统架构、安全性机制、安全性问题及改进方案。

1. 基于JAVA的移动电商平台系统架构
2. JAVA的概述

Java是一门面向对象编程语言，不仅吸收了C++语言的各种优点，还摒弃了C++里难以理解的多继承、指针等概念，因此Java语言具有功能强大和简单易用两个特征。Java语言作为静态面向对象编程语言的代表，极好地实现了面向对象理论，允许程序员以优雅的思维方式进行复杂的编程。并且具有以下特点：

* 面相对象

把相关的数据和方法组织为一个整体来看待，从更高的层次来进行系统建模，更贴近事物的自然运行模式。

* 健壮性

吸收了C/C++语言的优点，但去掉了其影响程序健壮性的部份（如指针、内存的申请和释放等），提供了一个相对安全的内存管理和访问机制

* 跨平台

通过java语言编写的应用程序在不同的系统平台上都可以运行。只要在需要运行java应用程序的操作系统上，先安装一个java虚拟机（JVM）即可。由JVM来负责java程序在该系统中的运行。

1. 基于JAVA的移动电商平台系统架构

随着移动电子商务平台的不断发展，通过移动电子商务平台进行的交易额也在逐年呈现暴涨趋势。为了应对愈发暴涨的流量，网站的系统架构，从最初的单独部署应用服务，到现在大规模集群方式，技术的发展经过了一代又一代的革新。本文以阿里为参考，得出一个基于JAVA的移动电商平台系统架构如下：

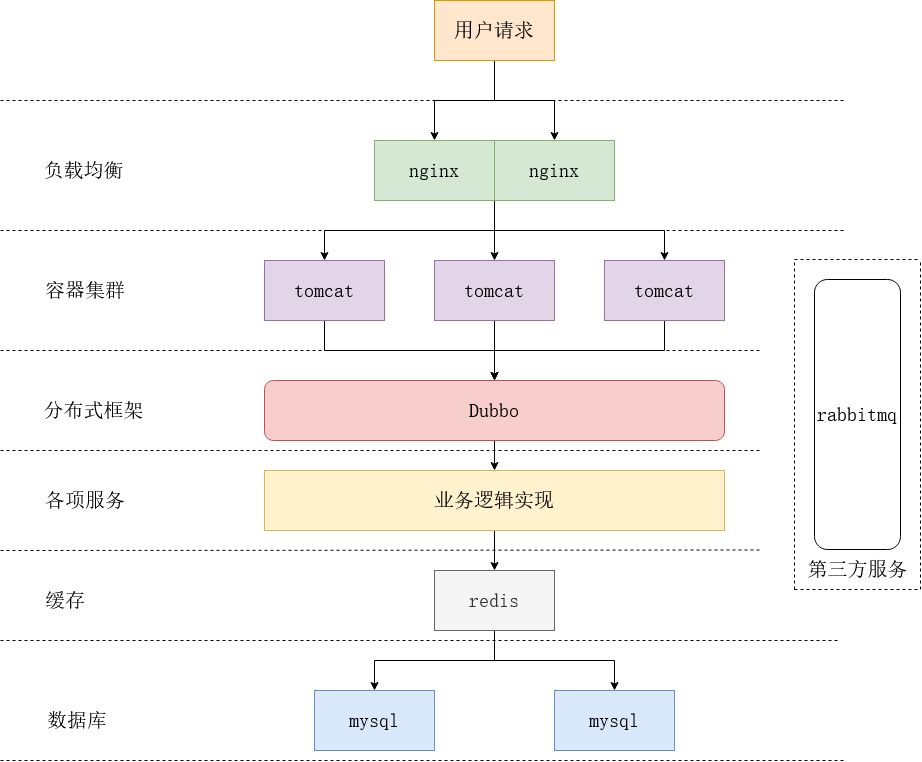


图1 基于JAVA的移动电商平台系统架构

从系统架构来看，一个基于JAVA的移动电商平台按照各组成部分可以分为以下几个部分:

* 负载均衡

负载均衡是指有效地将传入的网络流量分布到一组后端服务器上，也称为服务器池。移动电商平台必须为用户提供成百上千的并发请求，并以快速和可靠的方式返回正确的文本、图像、视频或应用程序数据。受限与现代科学技术，一台服务器并不能承载成百上千的并发请求，为了以成本有效的方式处理这些请求，现代计算最佳实践通常是添加更多的服务器来解决。

通过负载均衡，可以将用户请求分发到提供相同功能的不同服务器，降低服务器因并发量过大而导致容器奔溃的几率。也可以将CSS，JS，HTML，图片等静态资源从容器中分离出来，使得单个容器更易维护，运行效率更高。

现在使用最广泛的负载均衡工具为nginx。

* 容器集群

Servlet容器是指使用JAVA编写的，一个符合Servlet容器规范的程序，该程序封装了包含HTTP的网络协议规范，使得开发过程中，不用关注网络部分，从而能更好的专注于业务开发。Servlet容器的作用是负责处理客户请求，当客户请求来到时，容器获取请求，然后调用某个Servlet，并把服务的执行结果返回给客户。

容器集群，在一个移动电商平台中，单个容器并不能承载如此庞大的并发量，为了加强整体架构的刚性，承载更大的并发量，在一个移动电商平台中，通常采用多个容器，形成一种容器集群的模式。

* 分布式框架

在一个移动电商平台中，为了拆分业务，更好的使用容器，单个容器并不会拥有太多的功能，随着业务的发展，容器的增多，容器间相互通信与管理成为了一个很大的难题。分布式框架并不是一个程序，它是一段预先编写好的代码，封装了容器与容器之间相互通信的协议，提供了丰富的API，使得开发过程中，不用过分关注容器与容器之间的通信，每个开发人员可以只关注自己的部分业务。

一个分布式框架通常包括: 服务发现注册、配置中心、消息总线、负载均衡、断路器、数据监控等功能。

* 服务

这里的服务是值由开发人员编写的程序逻辑，一个移动电商平台常见的服务有：登录服务、购物车服务、交易服务、对账服务等。开发人员进行开发时，通常会分为三层：视图层、业务层与持久层。视图层负责页面展示与调整、业务层负责业务模块的逻辑应用设计、持久层负责数据持久化工作。

* 缓存

缓存就是将程序或系统经常要调用的对象存在内存中，一遍其使用时可以快速调用，不必再去查询数据库或者第三方服务。在一个移动电商平台中，往往存在着大量的热点数据，这些数据查询的频率很高，但是并不需要实时更新，如商品详情，商品浏览数，以及商品评论等。对这类数据采用缓存，能有效的降低数据库的访问频率，减少由数据库瓶颈对整个移动电商平台的影响。

* 数据库

数据库是一个长期存储在计算机内的、有组织的、有共享的、统一管理的数据集合。在一个移动电商平台中，单个数据库不会存储所有业务的数据，而是会根据业务进行划分，单个数据库只存储与单个容器相关的业务数据。如果需要表与表之间的关联查询，往往是通过业务逻辑来实现的。

* 第三方服务

第三方服务是指由第三方实现的程序，它封装了一系列的功能，只要按照程序的API给与入参，程序就能使用相关功能，节约了开发者开发相关功能的时间。目前在移动电商平台中用的比较多的第三方服务有：消息队列服务(rabbitmq),工作流服务(actoviti),搜索引擎(solr)。

1. 基于JAVA的移动电商平台的安全性机制
2. JAVA安全机制

java通过的沙箱安全模型保证了其安全性， java提供的沙箱组件有以下几种：

1. 类装载器结构；
2. 防止恶意代码去干涉善意的代码。

这是通过为不同类加载器提供不同的命名空间来实现的，在java虚拟机中，在同一个命名空间内的类可以直接进行交互，而不同的命名空间中类甚至不能觉察彼此的存在，除非显式地提供允许它们交互的机制。

1. 守护了被信任的类库的边界

虚拟机通过使用不同的类装载器装载可靠的包和不可靠的包，即所谓的双亲委派模式：在某个特定的类装载器试图以常用方式装载类型之前，它会先默认将这个任务“委派”给它的双亲，这个双亲再依次请求自己的双亲来装载这个类型。这个委派的过程一直向上继续，直到达到启动类加载器（boostrp classloader），如果一个类加载器的双亲类加载器有能力装载这个类型，则这个类加载器返回这个类型，否则这个类装载器试图自己来装载这个类型。

1. 将代码归入某类（称为保护域），该类确定了代码可以进行哪些操作。
2. class文件检验器；

和类加载器一起，class文件检验器保证了装载的class文件内容中有正确的内部结构，并且这些class文件相互间协调一致，如果class文件检验器在class文件中发生了问题，它将抛出异常。

java虚拟机的class文件检验器在字节码执行之前，必须完成大部分检验工作。class文件检验器需要经过四趟独立的扫描来完成其工作。

第一趟扫描：class文件结构检查。对每一段被当做类型导入的字节序列，class文件检验器都会确认其是否符合class文件格式。

第二趟扫描：类型数据的语义检查。检验一些java语言应该在编译时遵守的强制规则。

第三趟扫描：字节码验证。它确保局部变量在赋值之前不可访问，类的字段中总必须赋予正确类型的值等等。

第四趟扫描：符号引用验证。在这趟检查中，java虚拟机将追踪那些引用，从被验证的class文件到被引用的class文件，确保这个引用是正确的。

1. 内置于java虚拟机（及语言）的安全特性；

java虚拟机装载了一个类，并且对它进行了第一到第三趟的class文件检验，这些字节码就可以被运行了。除了对符号引用的检验（class文件检查的第四趟扫描），java虚拟机在执行字节码时，还进行了一些内置的安全机制的操作。这些机制大多数是java类型安全的基础：

1.类型安全的引用转换；

2.结构化的内存访问（无指针算法）；

3.自动垃圾收集；

4.数组边界检查；

5.空引用检查。

1. 安全管理器及java API。

安全管理器定义了沙箱的外部边界，保护虚拟机的外部资源不被虚拟机内运行的恶意或者有漏洞的代码侵犯。javaAPI在进行一个可能不安全的操作前，总是会检查安全管理器，所以javaAPI不会在安全管理器建立的安全策略下执行被禁止的操作。

1. 数据传输安全机制

数据传输过程中，安全性十分重要，尤其是对于一些敏感的用户信息，其安全性保证更应当被重视。保证传输数据的安全性，主要有以下方法：

1. 数据加密

数据加密一直是数据安全的重要部分，数据加密也就是利用技术手段对重要数据进行加密，然后把加密过后的数据进行传输，到达目的地后再用相同或不同的手段进行解密。对数据加密的技术分为两类,即对称加密技术和非对称加密技术。

* 对称加密

对称加密的特点是加密与解密时采用相同的密钥。对称加密算法简单快捷,密钥较短,且破译困难。对称加密以数据加密算法(DES, Data Encryption Standard)为典型代表。对称加密在使用中会存在以下几个问题：

1.通讯双方在首次通讯时需要一个安全的通道协商一个共同的密钥、

2.密钥的数目难于管理、

3.对称加密算法一般不能提供信息完整性的鉴别、

4.对称密钥的管理和分发工作是一件具有潜在危险的和烦琐的过程。

* 非对称加密

非对称加密算法需要有两个密钥，公开密钥和私有密钥。一般来说，公钥并不需要对任何人保密，是公开的。私钥并不公开，只有其所有者知道。如果用公开密钥对数据进行加密,只有用对应的私有密钥才能解密如果用私有密钥对数据进行加密,那么只有用对应的公开密钥才能解密。因为加密和解密使用的是两个不同的密钥,所以这种算法叫作非对称加密算法。

非对称加密算法的保密性比较好,它消除了最终用户交换密钥的需要,但加密和解密花费时间长、速度慢,它不适合于对文件加密而只适用于对少量数据进行加密。

1. 数字签名

数据加密只解决了传输数据机密性的问题，如果需要确认发信者的身份，则需要使用数据签名的技术。数字签名是建立在公开密钥体制基础上,它是公开密钥加密技术的另一类应用。它的主要方式是，报文的发送方从报文文本中生成消息摘要，发送方用自己的私有密钥对这个消息摘要进行加密来形成发送方的数字签名，然后,这个数字签名将作为报文的附件和报文一起发送给报文的接收方。报文的接收方首先对接收到的原始报文生成消息摘要，接着再用发送方的公开密钥来对报文附加的数字签名进行解密。如果两个值相同、那么接收方就能确认该数字签名是发送方的。

使用数字签名，能确认以下两点：第一，信息是由签名者发送的；第二，信息自签发后到收到为止未曾作过任何修改。

1. SSL协议

SSL（安全套接字）协议利用对称加密技术和公开密钥加密技术，在两点之间交换私钥建立了安全的连接。该协议提供了以下的安全服务：

1.认证用户和服务器，确保数据发送到正确的客户机和服务器；

2.加密数据以防止数据中途被窃取；

3.维护数据的完整性，确保数据在传输过程中不被改变。

1. 访问控制安全策略

访问控制是给出一套方法，将系统中的所有功能标识出来，组织起来，托管起来，将所有的数据组织起来标识出来托管起来， 然后提供一个简单的唯一的接口，这个接口的一端是应用系统一端是权限引擎。权限引擎所回答的只是：谁是否对某资源具有实施 某个动作（运动、计算）的权限。返回的结果只有：有、没有、权限引擎异常了。

1. 基于角色的访问控制模型

传统的强制访问控制和自主访问控制由于各自的缺点均不能满足移动电商平台对访问控制的需求。在基于角色的访问控制模型中，将资源与角色相关联，角色与用户关联，将角色设置成资源的集合，从而使得整个模型系统设计的简易性和安全控制的灵活性。同时允许角色进行重用、继承和派生操作。这样，不仅可以对资源进行访问控制，而且派生角色可以继承父角色的访问控制设置，这对于信息量巨大、内容更新变化频繁的移动电商平台非常有益，可以减轻由于信息资源的派生、演化和重组等带来的分配、设定角色权限等的工作量。

1. 服务接口的安全设计

在一个移动电商平台中，资源的形式通常是以接口的形势存在的，它需要相应的参数，程序执行完成后，返回对应的输出。基于角色的访问控制模式仅仅能保证用户是否能访问接口，但是却没有对接口的入参进行校验，因此，如果想要更高的安全性，往往还需要对服务接口进行以下设计：

* 时间戳机制

因为移动电商平台的数据经常在公网上传输，容易被抓包，虽然经过加密，加签处理，攻击者拿不到真实数据，但是攻击者仍然可以基于此，进行恶意请求。这时候可以使用时间戳机制，在每次请求中加入当前的时间，服务器端获取当前时间和消息中的时间比较，看看是否在一个固定的时间范围内，这样恶意请求的数据包是无法更改里面时间的，所以不在固定时间范围内的请求就视为非法请求了；

* 限流机制

因为移动电商平台对外开发，攻击者能合法的拿到账号数据，并且对移动电商平台进行正常请求。但是，如果出现频繁极高重复调用接口的情况，这个时候就需要限流。限流机制基于计数器方法，即当一定时间内，用户请求数超过设定的阈值时，进行限流，拒绝该次接口调用的请求。

* 黑名单机制

如果一个账号多次进行非法操作，比如多次请求访问不属于自己角色接口时，此时考虑黑名单机制。黑名单机制通常使用标记法来实现，即当一个用户进行多次非法操作时，更改该用户状态，下次用户在进行请求时，验证状态，如果状态异常，则拒绝该次请求，或者进行账号封停操作。

* 数据合法性校验

数据合法性校验对于移动电商平台尤为重要，数据合法性校验是指对请求的入参精选校验，常规的校验有身份证长度，电话号码长度等；安全性的校验有特殊字符，特殊标签等。常见的攻击手段往往都是通过使用非法的请求参数对系统进行攻击。比如XSS攻击，就是使用非法字符，向页面中植入一段脚本，使得其他用户访问这个页面时，信息泄露至脚本处。严重的情况可导致整个移动电商平台奔溃，无法正常服务。

1. 基于JAVA的移动电商平台安全性风险
2. 单点服务故障

单点服务故障是指在整个移动电商平台的系统架构中，容器集群中的单个容器无法提供正常服务，即接口不能提供正确的返回结果。对于移动电商平台来说，一次完整的调用需要多个容器协同工作，一个容器故障，无法正常提供服务，就可以导致这次调用链的失败。如果没有使用有效的容错机制，请求链接不能有效地关闭，错误的请求会积压到调用方，从而降低整个移动电商平台的吞吐量。

1. 导致单点服务故障的原因

* 配置错误

配置文件是平台中一些需要预定义变量的集合。配置文件编写错误即预定义的变量与正确变量不符。配置错误通常包括容器配置文件编写错误，或程序配置文件编写错误。容器配置配件编写错误会导致容器无法正常启动，对外提供服务；或能正常提供服务，但是容器里的资源接口无法正常访问。容器与容器之间的通信主要通过分布式框架的注册中心与分布式框架协议。注册中心负责服务的注册与发现。如果分布式配置文件编写错误，会导致程序无法正常启动，或程序正常启动，但是调用方无法找到被调用方的情况。

此外，程序中还有一些第三方服务，或者程序框架的配置文件。如果这些配置文件出错，均会导致单点服务故障。

* 程序异常

在JAVA里，异常被分为检查异常和非检查异常。检查异常是指编写程序时，class文件检验器校验class文件，并非提示的异常，这类异常会在编写程序期间处理并解决。非检查异常是指程序的瑕疵或逻辑错误，并且在运行时无法恢复，class文件检验器无法校验，在异常发生时，由JAVA虚拟机捕捉异常，并输出到控制台。

非检查异常无法在编译期解决，它对于整个移动电商平台有着更大的危害。导致非检查异常的原因多种多样，有一种非检查异常是在程序编写递归服务时，递归服务没有定义结束递归的边界，从而无限调用，产生内存异常。还有一种非检查异常是空指针异常，这里的指针并不是C语言中的指针，它是在JAVA虚拟机中，指向内存区中的对象的指针。程序运行时，该句柄指向的对象为空时，调用这个空对象的方法或属性时，就会抛出空指针异常。

1. 单点服务故障的容错机制

对于单点服务的故障，一般有以下的容错机制：

* 重试机制

重试机制是当发生单点服务故障时，调用方或者被调用方居然进行重新调用的情况。但是并不是所有情况都需要进行重试，当被调用方出现异常时则需要进行重试，如：程序异常、调用超时、网络中断等。符合程序安全机制的部分则不进行重试，如：数据不合法、用户处于黑名单中、消息中的时间戳不符合等。在分布式框架中通常有着自己的重试机制，在dubbo中可以设置retries=1，timeout=500，即调用失败只重试1次，超过500ms调用仍未返回则调用失败。

* 服务熔断

服务熔断是当该单点服务多次发生故障，多次重试后均返回错误异常，此时对该服务进行熔断。服务熔断的作用可类比成家用保险丝，当服务发生不可用的情况时，为防止系统架构雪崩，暂停该服务的应用。但是与家用保险丝不同的是，服务熔断会间隔一段时间进行服务调用测试，如果服务正常，则恢复该服务应用，如果服务异常，则继续保存熔断状态，等待下一次的服务调用测试。

1. 缓存失效

缓存使用流程：服务先从缓存中取数据，取到数据直接返回结果，取不到数据时从数据库中取，数据库取到更新缓存，并返回结果，数据库也没取到，直接返回空结果。当缓存失效时，用户请求无法使用缓存，导致请求压力直接积压到数据库，从而导致整个系统失效。以下是缓存失效的原因以及常用的处理方案。

1. 缓存穿透

缓存穿透，是指用户请求缓存和数据库中都没有的数据。在使用数据库时，会采用一个自增长的正整数作为主键，缓存中获取数据的依据通常含有主键，以正整数主键为例，正整数主键从1开始，步长为1，那么此时多次请求主键-1时，或者超过数据库整数类型长度时，这时就会发生缓存穿透，请求直达数据库，并且数据库中无相应数据。

因为缓存穿透是缓存以及数据库中都没有对应数据，对于缓存有以下处理方法：

第一种处理方法：将数据库中所有数据导入至缓存，更新数据库时同时更新缓存，直接减去查询数据库的步骤。对于这种方法来说，少量的配置类型的数据是可以的，但是对于移动电商平台来说，数据量过于庞大，不可能将内存作为第二数据库，因此该方案在移动电商平台中无法应用。

第二种处理方法：查询数据库时，如果数据库中没有数据，就将一个空对象插入到数据库中，下次再有同样的请求，则将缓存中的空对象返回。这种处理方法的缺陷在于，可以执行缓存穿透的主键极多，使用这种方法，缓存中会存在着大量的无用的空对象，使用这些空对象，有可能也会造成程序的异常。对于此，人们又对这种方法做了一些改进，即使用布隆过滤器。布隆过滤器的工作方式是：定义一个数组，每个位置都保存为0；定义一系列HASH函数；当键入缓存数据时，对缓存数据的键进行HASH运算，将数列中对应的元素置为1；获取数据时，对缓存数据的键进行HASH运算，如果数据中的元素为0，则说明数据不存在。因为受限于队列大小，存在HASH碰撞的可能性，即使当数据中的元素为1时，数据也不一定存在，依然要从数据库中获取。但是使用布隆过滤器，能大大的减少了从数据库中获取数据的几率。

1. 缓存雪崩

缓存雪崩，是指缓存中数据大批量同时过期，导致大量数据直接请求数据库，从而导致系统失效。在移动电商平台正式服务前，会把热点数据键入缓存中，从而减少正式上线后直接查询数据库的频率，如果这部分热点数据设置了同样的缓存过期时间，那么在缓存将会同时过期，发生缓存雪崩。对于缓存雪崩，解决方法也比较简单，每个缓存设置不同的过期时间即可，或根据需要，让该缓存永不过期。

1. 总结

参考文献

［1］快消品移动电商…