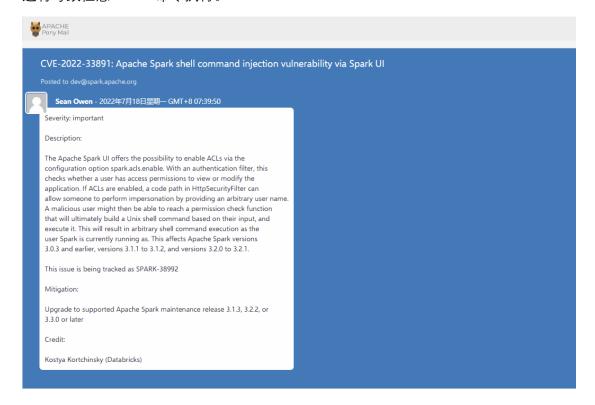
CVE-2022-33891_Apache_Spark_shell 命令注入漏洞

漏洞描述

7月18号,Apache 发布安全公告,修复了一个 Apache Spark 中存在的命令注入漏洞。漏洞编号:CVE-2022-33891,漏洞威胁等级:高危。 Apache Spark UI 提供了通过配置选项 Spark .acl .enable 启用 acl 的可能性。使用身份验证过滤器,这将检查用户是否具有查看或修改应用程序的访问权限。如果启用了 acl, HttpSecurityFilter中的代码路径可以允许某人通过提供任意用户名来执行模拟。 恶意用户可能能够访问权限检查功能,该功能最终将根据他们的输入构建一个 Unix shell 命令并执行它。这将导致任意 shell 命令执行。



相关介绍

Apache Spark 是美国阿帕奇(Apache)软件基金会的一款支持非循环数据流和内存计算的大规模数据处理引擎。Spark 优点在于能更好地适用于数据挖掘与机器学习等需要迭代的 MapReduce 的算法。

利用范围

Spark Core - Apache <=3.0.3

```
3.1.1 <= Spark Core - Apache <=3.1.2
```

3.2.0 <= Spark Core - Apache <=3.2.1

漏洞分析

环境搭建

在宣网下载 Apache Spark 3.2.1 版本进行漏洞复现分析

Index of /dist/spark/spark-3.2.1

	Name	Last modified	<u>d</u>	Size	<u>Description</u>
4	Parent Directory			_	
Ð	SparkR 3.2.1.tar.gz	2022-01-20 21	1:37	341K	
	SparkR 3.2.1.tar.gz.asc	2022-01-20 21	1:37	833	
Ĭ	<u>SparkR 3.2.1.tar.gz.sha512</u>	2022-01-20 21	1:37	207	
	pvspark-3.2.1.tar.gz	2022-01-20 21	1:37	268M	
	pyspark-3.2.1.tar.gz.asc	2022-01-20 21	1:37	833	
	pyspark-3, 2, 1, tar, gz, sha512	2022-01-20 21	1:37	210	
	spark-3.2.1-bin-hadoop2.7.tgz	2022-01-20 21	1:37	260M	
	spark-3.2.1-bin-hadoop2.7.tgz.asc	2022-01-20 21	1:37	833	
	spark-3.2.1-bin-hadoop2.7.tgz.sha512	2022-01-20 21	1:37	268	
	spark-3.2.1-bin-hadoop3.2-scala2.13.tgz	2022-01-20 21	1:37	293 M	
	spark-3, 2, 1-bin-hadoop3, 2-scala2, 13, tgz, asc	2022-01-20 21	1:37	833	
	<u>spark-3, 2, 1-bin-hadoop3, 2-scala2, 13, tgz, sha512</u>			187	
	spark-3, 2, 1-bin-hadoop3, 2, tgz	2022-01-20 21		287M	
	spark-3, 2, 1-bin-hadoop3, 2, tgz, asc	2022-01-20 21			
	spark-3.2.1-bin-hadoop3.2.tgz.sha512	2022-01-20 21		268	
	spark-3.2.1-bin-without-hadoop.tgz	2022-01-20 21		200M	
	spark-3.2.1-bin-without-hadoop.tgz.asc	2022-01-20 21		833	
	spark-3.2.1-bin-without-hadoop.tgz.sha512	2022-01-20 21		288	
	spark-3.2.1.tgz	2022-01-20 21		27M	
F	spark-3.2.1.tgz.asc	2022-01-20 21			
	<u>spark-3, 2, 1, tgz, sha512</u>	2022-01-20 21	1:37	140	

漏洞触发的关键在于是否启用ACL,使用身份验证过滤器。

启用 ACL 的两种方式: 1、通过设置选项 spark.acls.enable 启用 。2、运行 sparkshell 时,通过-c 参数启动。

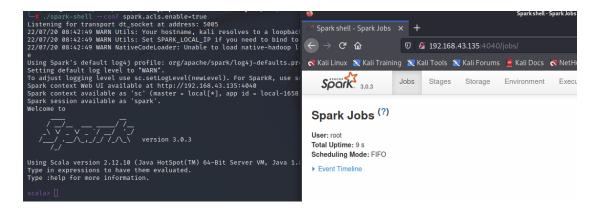
为更好分析漏洞,在运行 spark-shell 前,需在其中进行远程调试配置

export SPARK_SUBMIT_OPTS="-

agentlib:jdwp=transport=dt_socket,server=y,suspend=n,address=5005"

配置完成后运行 spark-shell, 并开启 ACL

./spark-shell --conf spark.acls.enable=true



漏洞原理

根据 diff 分析

```
v 💠 5 💶 core/src/main/scala/org/apache/spark/security/ShellBasedGroupsMappingProvider.scala 📮
                                           @@ -30,6 +30,8 @@ import org.apache.spark.util.Utils
 30
                       30
                                            \verb|private[spark]| class ShellBasedGroupsMappingProvider extends GroupMappingServiceProvider (and the private of the private 
 31
                       31
 32
                       33 + private lazy val idPath = Utils.executeAndGetOutput("which" :: "id" :: Nil).stripLineEnd
 33
 34
                                                 val userGroups = getUnixGroups(username)
                                                      logDebug("User: " + username + " Groups: " + userGroups.mkString(","))
                                          @@ -38,8 +40,7 @@ private[spark] class ShellBasedGroupsMappingProvider extends GroupMappingService
 38
 39
                                                 // shells out a "bash -c id -Gn username" to get user groups
 HyukjinKwon marked this conversation as resolved.
                                                                                                                                                                                                                                                                                                      - Show resolved
                      private def getUnixGroups(username: String): Set[String] =

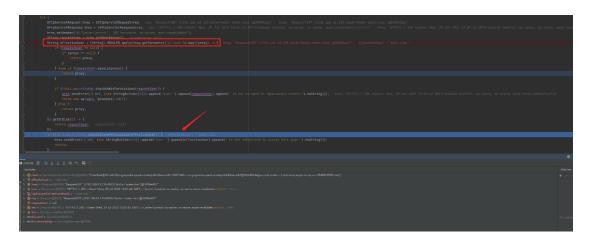
val cmdSeq = Seq("bash", "-c", "id -Gn " + username)
 41
 42
                         - Utils.executeAndGetOutput(cmdSeq).stripLineEnd.split(" ").toSet
                   44 + Utils.executeAndGetOutput(idPath :: "-Gn" :: username :: Nil).stripLineEnd.split(" ").toSet
 44
                        45
 45
                       46
```

如上所示,使用命令拼接且没有做任何处理,而在修复的版本中直接删除了 ShellBasedGroupsMappingProvider 函数中对 bash 的调用

动态分析

了解漏洞原理之后,就该考虑如何触发漏洞。

在 Apache spark 启用 ACL 后,会通过 HttpSecurityFilter 这个 filter 进行权限的校验 首先将断点打在 org.apache.spark.ui.HttpSecurityFilter#doFilter 函数处



在进入 doFilter 函数之后,首先会提取参数"doAs"的值,然后赋值给 effectiveUser, 进入 org.apache.spark.SecurityManager#checkUlViewPermissions 函数

```
politic booleen (contributivestation) (as for a street of the contribution) append(contribution) append(contributi
```

后续跟进一系列函数进行处理

```
private boolean isUserInACl(final String user, final Set/Strinp> acUsers, final Set/Strinp> acUsers, final Set/Strinp> acUsers, final Set/Strinp> acUsers "sotre-cloc" acutivers: "Set(root)" acutivers: "Set(
```

进入 org.apache.spark.security.ShellBasedGroupsMappingProvider#getGroups 函数时, username 为传入参数

随后进入

org.apache.spark.security.ShellBasedGroupsMappingProvider#getUnixGroups 函数

在这里 username 进行了拼接处理,因为我们传入的 username 参数可控,便形成了命令注入。

后续将通过 executeAndGetOutput 函数直接触发传入的命令,造成命令执行。



漏洞复现

通过反单引号和参数"doAs"成功命令注入



修复建议

建议受影响的用户升级到安全版本: Apache Spark 3.1.3、3.2.2 或 3.3.0 或更高版本。