Hadoop生态系统安全审查工具使用文档

目录

1. **引言** 
   1. **编写目的**

占位符

* 1. **技术支持**

占位符

1. **工具部署** 
   1. **环境要求**

工具需要的环境是Python 3.5以上。

工具需要的Python模块如下：

**Django:**

安装方法：pip install django==2.0.3

**Paramiko:**

安装方法：pip install paramiko

**Pexpect:**

安装方法：pip install pexpect

**Reportlab:**

安装方法：pip install reportlab

**Nmap:**

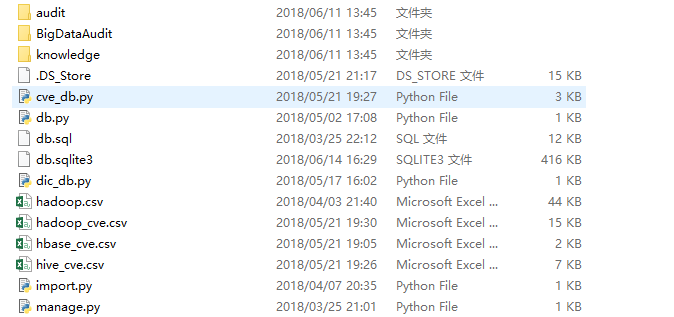
安装方法：pip install nmap

**Netaddr:**

安装方法：pip install netaddr

* 1. **启动说明**

在工具所在的Django工程根目录（包含manage.py文件的目录）下，启动命令行工具（Windows系统）或终端（Mac OS或Linux系统），输入并执行指令python manage.py runserver。



**图2.1 Django工程根目录**

启动浏览器，输入IP地址127.0.0.1:8000，打开工具主界面。



**图2.2 工具主界面**

1. **安全审查流程** 
   1. **工具使用基本流程**

本工具对Hadoop生态系统的安全特性检查提供了系统而全面的支持，包括了从检查前的检测预备，到外部渗透测试、安全机制检查，以及最后检查报告输出等各个步骤。

* + 1. 检测准备

在对Hadoop集群进行自动化检查之前，测试人员应尽可能的了解被测系统的相关信息。本工具针对这方面需求，提供了运行环境调查分析、用户基本信息维护等辅助功能，使得安全评估人员可以更好地在Hadoop集群的运行环境具体配置、用户信息情况等方面与测试系统的管理员进行交互。因此在使用本工具和检测目标系统之前，为保证安全审查有效实施，安全测试人员需要根据被测系统的实际情况，对其运行环境进行调查研究，并对工具的基础知识库进行相应的维护，例如维护安全特性知识库中的服务组件、安全配置、审查策略、潜在的用户弱密码及弱密码规则、漏洞补丁等。

* + 1. 外部渗透测试

在对目标对象检查前，首先要使用相关的工具对Hadoop集群运行环境进行扫描，以发现Hadoop生态系统及相关组件。Hadoop相关组件一般会开放服务端口来响应某些功能需求。因此服务发现可通过IP地址段和端口探测Hadoop相关组件。当获取了Hadoop集群的IP地址以后，测试工具可以与远程数据库进行通信，在没有合法用户名和密码的情况下，对集群进行渗透性探测。在渗透性测试过程中，本工具提供了对弱密码用户进行暴力破解的登录尝试。因此在运行本测试之前，建议用户对知识库中常见用户名及其密码库进行维护，以保证渗透测试的有效性。此外，本工具建立了Hadoop相关的CVE知识库，能够对Hadoop集群中各组件的版本进行检测并提醒用户相关漏洞可能存在及解决方案。

* + 1. 安全机制检查
    2. 检查报告输出

在进行完以上所有工作后，可以把目标系统的安全特性检测结果用统计图表、文档等形式输出并可供评估人员下载。

* 1. **实施前准备工作**

为了保证Hadoop安全审查的有效顺利进行，需要在使用工具进行审查之前，根据被测系统的具体情况，进行必要的准备工作，包括运行环境调查分析，工具安全知识库基本信息维护等工作：

* + 1. 运行环境调查分析

在对Hadoop生态系统本身进行安全审查之前，首先需要对运行环境进行调查分析，作为目标Hadoop系统风险评估的基本参考信息。

* + 1. 弱密码知识库维护

需要对用户个人信息与密码习惯进行调查，根据调查结果，通过本工具的知识库管理模块可以对弱密码字典与弱密码生成规则进行管理与维护，以保证弱密码渗透的有效性。

* + 1. 漏洞知识库维护

每隔一段时间需要对工具的CVE知识库进行更新与补充，通过本工具的知识库管理模块可以对漏洞知识库进行管理与维护，以保证漏洞检测的时效性。

* + 1. 审查策略构建与维护

需要有针对性地建立或维护相关检查策略。比如HIPAA——健康保险隐私及责任法案，主要是针对健康信息的存储数据库的安全设定的一系列相关标准，FISMA——联邦信息安全管理法案，主要是针对政府管理信息的存储数据库的安全设定的一系列相关标准等。

* 1. **外部渗透测试流程**

占位符

* 1. **安全机制审查流程**

占位符

1. **服务发现**

占位符

1. **弱密码渗透**

占位符

1. **版本漏洞检测**

占位符

1. **安全元数据查看**

安全元数据查看是从Hadoop 集群中获取到安全信息并且将安全信息直观地展示给审查人员。审查人员通过安全元数据查看可以对待审查的Hadoop 集群有更加深入的了解。2017 年颁布的《网络产品和服务安全审查办法》的第十二条中提到，在安全审查的过程中需要考虑到透明性的问题。安全元数据查看是使得集群安全信息透明化的重要方式。

工具定义的Hadoop 集群安全元数据分为四类。第一类是Hadoop 集群基础设施安全信息，第二类是Hadoop 集群平台安全信息，第三类是Hadoop 集群用户授权信息，第四类是Hadoop 集群安全审计与运行日志信息。

* 1. **选择待测目标**

在开始安全元数据查看前，审查人员需要选择待测节点。

**图7.1 未登录集群状态 图7.2 已登录集群状态**

图7.1显示的是工具处于未登录集群的状态，图7.2显示的是工具处于已登录集群的状态。当工具处于未登录集群状态下，审查人员无法进行正常的安全元数据查看操作。因此需要点击图7.1中的登录来出现图7.3所示界面。



**图7.3 集群信息输入界面**

图7.3是登录集群界面，审查人员需在对话框中输入待测节点的IP信息，用户名信息和密码信息后，点击登录按钮。若输入的信息正确，则工具进入如图7.2所示的状态，即可进行后续审查工作。

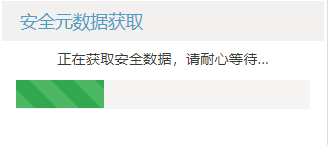
* 1. **Hadoop集群基本元数据信息查看**

安全元数据审查的主界面如图7.4所示，红色框内可选择安全元数据查看的类别。Hadoop集群基本元数据信息对应的是Hadoop 集群基础设施安全信息，在工具中体现为选择集群基本信息标签。



**图7.4 集群基本元数据信息查看界面**

当工具处于登录状态下时，审查人员可点击页面上的开始按钮开始安全元数据查看，等待界面如图7.5所示。



**图7.5 安全元数据获取等待界面**

集群基本元数据信息以表格的形式显示。



**图7.6 集群基本信息结果**

* 1. **Hadoop集群运行状态信息查看**

安全元数据审查的主界面如图7.7所示，红色框内可选择安全元数据查看的类别。Hadoop集群运行状态信息属于Hadoop 集群基础设施安全信息，在工具中体现为选择集群状态信息标签。



**图7.7 集群运行状态信息查看页面**

当工具处于登录状态下时，审查人员可点击页面上的开始按钮开始安全元数据查看，等待界面如图7.5所示。

集群运行状态信息以文字的形式显示。



**图7.8 集群运行状态信息结果**

* 1. **Hadoop集群服务级授权信息查看**

安全元数据审查的主界面如图7.7所示，红色框内可选择安全元数据查看的类别。Hadoop集群服务级授权信息属于Hadoop 集群用户授权信息，在工具中体现为选择集群服务级授权信息标签。



**图7.9 集群服务级授权信息查看页面**

当工具处于登录状态下时，审查人员可点击页面上的开始按钮开始安全元数据查看，等待界面如图7.5所示。

集群服务级授权信息以表格的形式显示。

* 1. **Hadoop集群日志信息查看**

安全元数据审查的主界面如图7.7所示，红色框内可选择安全元数据查看的类别。Hadoop集群日志信息属于Hadoop 集群安全审计与运行日志信息，分为日志名称信息和日志路径信息，在工具中体现为选择集群日志信息标签。



**图7.10 集群日志信息查看页面**

当工具处于登录状态下时，审查人员可点击页面上的开始按钮开始安全元数据查看，等待界面如图7.5所示。

集群日志信息以表格的形式显示。

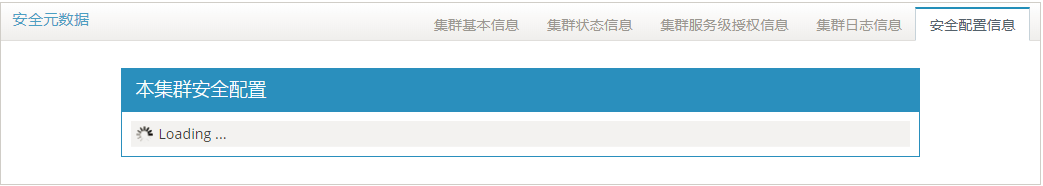
* 1. **Hadoop集群配置信息查看**

安全元数据审查的主界面如图7.7所示，红色框内可选择安全元数据查看的类别。Hadoop集群配置信息属于Hadoop 集群平台安全信息，在工具中体现为选择安全配置信息标签。



**图7.11 集群配置信息查看页面**

当工具处于登录状态下时，审查人员可点击页面上的开始按钮开始安全元数据查看，等待界面如图7.12所示。



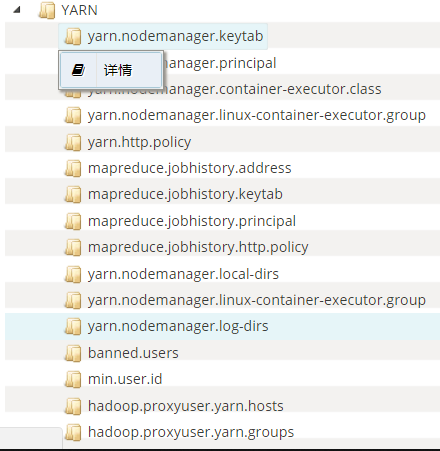
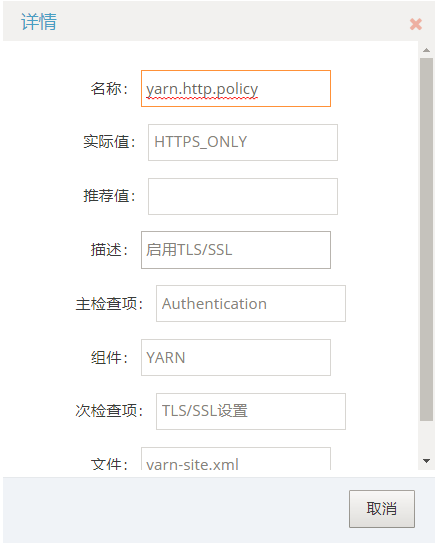
**图7.12 集群配置信息获取等待页面**

集群日志信息以树形结构的形式显示。



**图7.13 集群安全配置信息结果图**

展开树形结构可以看到对应组件的具体配置项名称，在配置项名称上（图中红色框的范围内）右键可弹出菜单，点击详情按钮可查看配置项的具体信息。

**图7.14 右键弹出菜单 图7.15 安全配置详细信息**

1. **安全评估**

安全评估部分主要分为安全合规性评估和安全冲突检测两个部分。

安全合规性评估的主要内容是对Hadoop 集群启用的安全机制进行分析。安全合规性评估发现并展示集群中存在问题或者缺失的安全机制，并且基于Hadoop 集群安全机制的分析结果给出安全建议。

安全冲突检测概念的提出源于Hadoop 生态系统在结构上的特点。部署了Hadoop 生态系统的系统往往是分布式集群的形式，在集群的每个节点上都存在着配置文件。当管理者对集群的某项配置进行修改时，所有包含该配置所在的配置文件的集群内节点都应该修改该项配置。不同节点不一致的配置会影响到该配置项涉及的安全机制的启用，因此我们需要安全冲突检测的概念，用于审查不同节点上同一配置项的配置值是否相同。此外，对于存储结构而言，其逻辑存储和物理存储可以分别从属于不同组件，然而不同组件的用户授权情况可能存在不一致的情形。为避免授权不一致导致可能的数据泄露，我们也需要安全冲突检测的概念。因此，安全冲突检测分为节点配置一致性审查和组件授权一致性审查。

在执行安全评估前，需要先输入主机用户名密码信息。点击左上方登录后出现如下界面，输入IP、用户名和密码。



**图8.1 输入用户名密码图**

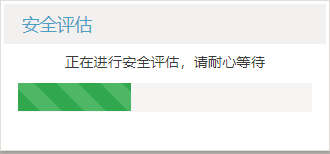
* 1. **安全合规性评估**

安全评估的主界面如下图所示。



**图8.2 安全评估主界面**

安全合规性评估涉及审查策略和组件的选取。页面中间可以选择审查策略和待检测的组件。除了工具预置的几种审查策略外，还可以点击自定义策略打开策略维护界面。关于策略维护的具体操作将在文档的第9章中具体说明。选择策略和组件后点击下方安全合规性评估按钮即可开始安全合规性评估。



**图8.3 安全评估等待界面**

安全合规性评估的结果将以弹出页面的形式显示，显示为表格。单击右下角关闭按钮可关闭页面。若不存在不合规的配置，则弹出页面显示提示文字，单击右下角关闭按钮可关闭界面。

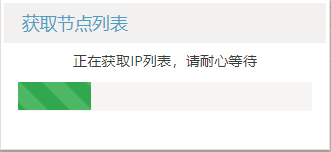


**图8.4 安全合规性评估结果界面**

* 1. **节点间配置一致性审查**

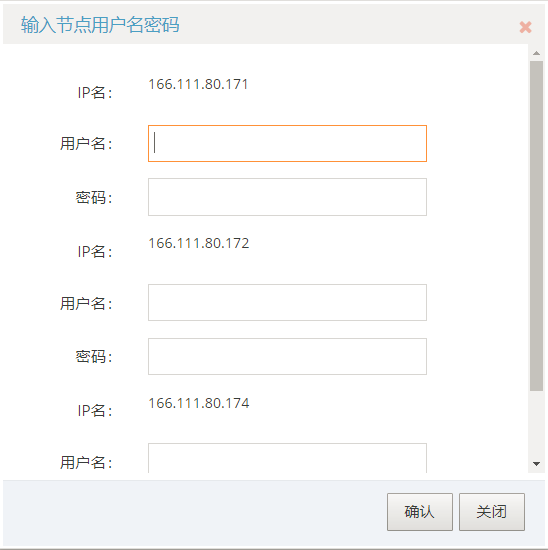
节点间配置一致性审查属于安全冲突检测的一部分。节点间配置一致性审查也需要选取审查策略以及待测的组件。节点间配置一致性审查的主界面如图8.2所示。

与安全合规性评估不同的是，节点间配置一致性审查在选择了审查策略和待测组件后，需要点击下方节点配置一致性审查按钮。点击后工具开始获取集群的节点列表，等待界面如下。



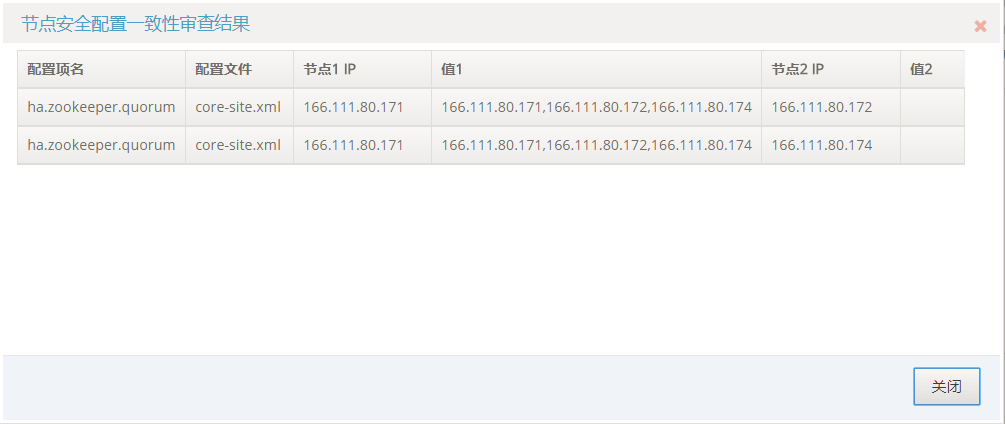
**图8.5 获取节点列表等待界面**

获取节点列表完毕后，会出现包含集群节点IP信息的页面。审查人员需要输入对应节点的用户名密码信息，输入界面如下。



**图8.6 输入节点用户名密码界面图**

输入正确的用户名密码后，点击下方确认按钮，会出现如图8.3的等待页面。节点配置一致性审查的结果将以弹出页面的形式显示，显示为表格。单击右下角关闭按钮可关闭页面。若不存在不一致的配置，则弹出页面显示提示文字，单击右下角关闭按钮可关闭界面。



**图8.7 节点间配置一致性审查结果页面**

* 1. **组件间授权一致性审查**

组件间授权一致性审查属于安全冲突检测的一部分。组件间授权一致性审查不需要选取审查策略以及待测的组件，默认检查Hadoop的两个存储组件HDFS和HBase。组件间授权一致性审查的主界面如图8.2所示。

尽管页面上存在策略和组件选项，这不会影响组件间授权一致性审查的结果。开始审查无需选择策略和组件，只需点击下方组件授权一致性审查按钮。此后开始获取两个组件的用户授权信息，会出现如图8.3的等待页面。组件授权一致性审查的结果将以弹出页面的形式显示，显示为表格。单击右下角关闭按钮可关闭页面。若不存在不一致的授权，则弹出页面显示提示文字，单击右下角关闭按钮可关闭界面。



**图8.8 组件间授权一致性审查结果页面**

1. **知识库使用与管理**

占位符