如何将正则表达式(regex)与grep结合使用来搜索日志？

目录

问题

如何将正则表达式(regex)与grep结合使用来搜索日志？

环境

思科网络安全设备
思科邮件安全设备
思科安全管理设备

解决方案

正则表达式(regex)可以是与“grep”命令一起使用以搜索设备上可用的日志（如访问日志、代理日志等）时的强大工具。使用CLI命令“grep”时，我们可以根据网站、URL的任何部分或用户名搜索日志，以命名一些。

以下是一些常见场景，您可以使用带有grep的regex来帮助进行故障排除。

情形 1：在访问日志中查找特定网站

最常见的情况是尝试在思科网络安全设备(WSA)的访问日志中查找向网站发出的请求。

例如:
通过SSH连接到设备。一旦您出现提示符，我们可以键入“grep”命令列出可用日志。

CLI> grep
输入要“grep”的日志编号。 []> 1（在此处选择访问日志的编号）
将正则表达式输入“grep”。 []>**网站\.com**

方案 2：尝试查找特定文件扩展名或顶级域

我们可以使用“grep”命令在URL或顶级域(.com、.org)中查找特定文件扩展名(.doc、.pptx)。

例如:
要查找以.crl结尾的所有URL，我们可以使用以下正则表达式：**\.crl$**

要查找包含文件扩展名.pptx的所有URL，我们可以使用以下正则表达式：**\.pptx**

情形 3：尝试查找网站的特定块

搜索特定网站时，我们可能也在搜索特定HTTP响应。

例如:
如果要搜索domain.com的所有TCP_DENIED/403消息，可以使用以下正则表达式：**tcp_denied/403.*domain\.com**

场景 4：在访问日志中查找计算机名称

使用NTLMSSP身份验证方案时，我们可能会遇到用户代理（Microsoft NCSI是最常见的）在身份验证时错误地发送计算机凭据而不是用户凭据的实例。要跟踪导致此情况的URL/用户代理，我们可以使用带有“grep”的regex隔离身份验证时发出的请求。

如果我们没有使用的计算机名称，我们可以使用“grep”并查找在使用以下正则表达式进行身份验证时用作用户名的所有计算机名称：**\$@**

一旦我们有了发生这种情况的行，我们就可以使用以下正则表达式对特定计算机名称进行“grep”：**机器名\$**

出现的第一个条目应该是用户使用计算机名称而不是用户名进行身份验证时发出的请求。

场景 5：在访问日志中查找特定时间段

默认情况下，访问日志订阅将不包含显示可读日期/时间的字段。如果要检查特定时间段的访问日志，我们可以执行以下步骤：

从http://www.onlineconversion.com/unix_time.htm等站点查找UNIX时间戳。一旦您有时间戳，您就可以在访问日志中搜索特定时间。

例如:
Unix时间戳1325419200等于01/01/2012 12:00:00。

我们可以使用以下regex条目在2012年1月1日12:00时左右搜索访问日志：13254192

场景 6：搜索严重或警告消息

我们可以使用正则表达式在任何可用日志（如代理日志或系统日志）中搜索严重或警告消息。

例如：
要在代理日志中搜索警告消息，可以输入以下正则表达式：

1. **CLI>** grep
2. 输入要“grep”的日志编号。
   []> 17（在此处为代理日志选择#）
3. 将正则表达式输入“grep”。
   []> **警告**

其他有用链接：

正则表达式 — 用户指南