CLI検索を使用して大量のデータをエクスポートする方法

適用性

Splunkエンタープライズ

説明

ステップバイステップガイド

大量のデータをエクスポートするには、検索ヘッドのCLIプロンプトから次の行に沿って検索を使用できます。

```
splunk search "index=pan_log sourcetype=pan:* src_ip!=<Internal_IP_Scheme> dest_ip!=<Int
| fillnull value=null
| dump basefilename=<insert_your_chosenfilename> compress=5 fields=_time,app,app:subcate
dest_ip,dest_port,dest_translated_ip,dest_user,dest_zone,direction,duration,filename,htt
packets_out,sig,signature,sourcetype,src_ip,src_translated_ip,src_user,target_id,threat:
```

上記の検索の詳細は次のとおりです。

「pan_log」インデックスでpan:*ソースタイプを検索し、IP範囲を除外し、NULLを埋めてdumpコマンドにパイプします。その後、次の処理に進みます。

このコマンドは、指定された検索クエリを実行し、「$SPLUNK_HOME/var/run/splunk/dispatch/<sid>/dump」にあるローカルディスクに検索結果をホットエクスポートします。 入力イベントの特別なフィールド_dstpathを認識します。このフィールドがsetの場合、最終的な宛先パスを計算するためにdstに付加されるパスとして使用されます。

このプロセスの主な欠点は、結果のTTLが2分であることです。したがって、プロセスによって作成される*.gzファイルを取得する前に、これらの結果が再度取得されないことを確認するために、保存トークンファイルを作成する必要があります。以下の「関連」リンクを参照してください。

関連：

Windowsで拡張子のないファイルを作成する方法

調査

この問題の調査中に、完全なデータセットを取得していないことが判明しました。

このデータを取得して予想よりも速くカリングしたことが原因であると判断しました。調査の結果、これは正常な動作であることがわかりました。要約インデックス検索およびその他の検索は、2分間の検索アーチファクトTTLで実行できます。このため、<sid>ディレクトリに保存トークンファイルを作成します。Windowsマシンの場合は、この記事の説明セクションのリンクを参照してください。*nixを使用している場合は、$SPLUNK_HOME/var/run/splunk/dispatch/<sid>に移動して、コマンド「touch save」を発行します。

これが完了すると、アーティファクトの修復の問題を回避できました[***保存ファイルを忘れずに削除してください!***]

その後、もう少しトラブルシューティングを進めて、定義されたフルデータセットを生成する実用的な時間枠を取得します。データセット内のイベント数が極端に多いため、最終的に24時間の検索を実行して1日ずつエクスポートする必要がありました。

解決方法

24時間の時間帯に対して実行する必要があるデータセットを絞り込むと、24時間のPANロギングの完全なダンプを生成できます。別のTSEは最近、この同じ方法を使用して過去の事例に戻り、具体的な情報を入手できるようになりました。

関連リンク

Splunkのマニュアル：ダンプコマンドマニュアル： 検索リファレンスダンプ

翻訳について
シスコは世界中のユーザにそれぞれの言語でサポート コンテンツを提供するために、機械と人による翻訳を組み合わせて、本ドキュメントを翻訳しています。ただし、最高度の機械翻訳であっても、専門家による翻訳のような正確性は確保されません。シスコは、これら翻訳の正確性について法的責任を負いません。原典である英語版（リンクからアクセス可能）もあわせて参照することを推奨します。