HTML Comment Scrubber iRule

ichiro, 2008-29-10

はじめに:
HTML文にコメントを利用するケースは多いですが、ブラウザでアクセスする際にコメントが表示されないため、コメントの存在自体を忘れがちです。しかし社内情報やサイトに関する機密情報など、一般ユーザに見せたくないものがコメントとして流れ、情報漏えいにつながるケースがあります。(例えばコメントの内容そのものが機密情報ではなくても、サイトへの不正アクセスのヒントとなるケースがあります)

そのため、HTMLコメントを一切出さないことがセキュリティポリシーとして一つの選択肢になりますが、全てのコンテンツをきちんとチェックしてコメントを削除するのは大変な作業であり、削除ミスや見逃しが起こる原因となります。

今月のiRuleでは、HTMLコメントを探し出し、存在した場合に削除するiRuleを紹介します。

詳細はこちらでも確認できます
http://devcentral.f5.com/Wiki/default.aspx/iRules/HtmlCommentScrubber.html

メリット:
HTMLコメントを自動で探し出して削除することで、機密情報の漏洩を防ぐ

タイトル:
HTML Comment Scrubber iRule

設定概要:
* HTTP::collectコマンドにてサーバからのレスポンスを一旦BIG-IP内にバッファ
* regexpコマンドで<! -- ... -- >のパターンを探しだす
* 上記の文字列を全体的に" "(空白)で置き換える(タグを含むコメント文が空白となる)

動作概要:
このルールはサーバから返されたHTMLコンテンツを全て読み取って、<! -- コメント -- >のHTML文が見つかった場合、コメントのテキストを空白(0x20)で置き換えます。

コメントそのものも、iRules的に可能なロジックですが、コンテンツサイズに影響するようなアクションをとるとChunkedタイプのレスポンスに対してBIG-IPのRechunk機能を使う必要がなくなります。

また、先月のiRuleでRegexpコマンドは負荷が大きいため、代わりにScanコマンドで文字列を探す方法を紹介しましたが、今回のように、文字列を置き換えるにはregexpが簡単でわかりやすい方法になります。

iRuleのregexpコマンドでは下記のようなロジックを使用します:

1. HTTP::payloadに対して<! -- ... -- >の文字列を探す
2. 上記パターンで見つかった文字列を、全て(-allの引数)探し出す
3. 見つかった文字列に対して、HTTP::payload内のインデックス情報(ペイロード内で見つかった文字列の最初と最後の文字のポジション)を記録(-indicesの引数)

インデックス情報をTCLリストとして返す(-inlineの引数)次に、上記のインデックス情報を使って、HTTP::payload replaceコマンドによってコメント文字列の始まるポジションから終わるポジションまで空白で置き換える。

使い方:
* HTTPのVirtual ServerにiRuleを関連付ける

【iRule定義】

```
when HTTP_REQUEST {
  # レスポンスがChunkedにならないようにリクエストを変換
  if { [HTTP::version] eq "1.1" } {
    if { [HTTP::header is_keepalive] } {
      HTTP::header replace "Connection" "Keep-Alive"
    }
    HTTP::version "1.0"
  }
}
when HTTP_RESPONSE {
  if { [HTTP::header exists "Content-Length"] } {
    set content_length [HTTP::header "Content-Length"]
  } else {
    set content_length 1000000
  }
  if { $content_length > 0 } {
    HTTP::collect $content_length
  }
}
when HTTP_RESPONSE_DATA {
  # HTMLコメントを見つけ出す


  set indices [regexp -all -inline -indices {<!--(?:[^-]|--|[^-]-[^-])*?--\s*?>} [HTTP::payload]]
  # コメントを空白で置き換える
  #log local0. "Indices: $indices"


  foreach idx $indices {
    set start [lindex $idx 0]
    set len [expr {[lindex $idx 1] - $start + 1}]
    log local0. "Start: $start, Len: $len"
    HTTP::payload replace $start $len [string repeat " " $len]
  }
}
```

F5ネットワークスジャパンでは、サンプルコードについて検証を実施していますが、お客様の使用環境における動作を保証するものではありません。実際の使用にあたっては、必ず事前にテストを実施することを推奨します。

F5 Networks, Inc. | 401 Elliot Avenue West, Seattle, WA 98119 | 888-882-4447 | f5.com

F5 Networks, Inc.
Corporate Headquarters
info@f5.com

F5 Networks
Asia-Pacific
apacinfo@f5.com

F5 Networks Ltd.
Europe/Middle-East/Africa
emeainfo@f5.com

F5 Networks
Japan K.K.
f5j-info@f5.com