Que é formato do mbox de UNIX (caixa postal)?

Índice

Introdução

Este documento descreve o formato da caixa postal de Unix (mbox) e como se relaciona para se usar na ferramenta de segurança do email de Cisco (ESA).

Que é formato do mbox de UNIX (caixa postal)?

O formato do mbox de UNIX é usado por AsyncOS quando as mensagens são arquivadas e entrou a ação do log() do filtro da mensagem. Do “a mensagem arquivo” é uma proteção anti-vírus (Sophos e McAfee), avançada adicional da opção de configuração para o Anit-Spam de Ironport (IPA), do malware (AMP), e Graymail no ESA.

O formato de Mbox é (isto é,) um formato do arquivo não binário ASCII-formatado que possa conter zero ou mais mensagens do correio. As mensagens são concatenadas no arquivo do mbox e podem ser erguidas distante basearam em cordas específicas no arquivo. Este formato é idêntico com a mensagem porque são transferidos entre gateways complacentes do correio do RFC 2821.

Cada mensagem no formato do mbox começa com uma linha de que comece com a corda “” (caracteres ASCII F, r, o, m, e espaço). “” Das linhas são seguidos por diversos mais campos: envelope-remetente, data, e (opcionalmente) mais dados.

O primeiro campo após “” da corda é o envelope-remetente da mensagem. O dependente em cima de que o aplicativo cria o arquivo do mbox, o envelope-remetente pôde esta presente enquanto uma caixa postal ou real puderam ser um outro caráter ou corda. O mais geralmente, você encontrará que “-” (único traço do caráter) substitui o envelope-remetente se o envelope-remetente real não é disponível ou não conhecido. O campo da data introduzido pelo ESA é no formato padrão do asctime() de UNIX e é sempre 24 caráteres de comprimento. Em alguns arquivos do mbox redigidos pelas aplicações NON-AsyncOS, a informação adicional segue a data. Estes três campos são separados por um espaço único.

Está aqui um exemplo de um arquivo do mbox com uma única mensagem nele:

```
From Adam@Outside.COM Sun Oct 17 12:03:20 2004
Received: from mail.outside.com (192.35.195.200)
by smtp.alpha.com with ESMTP; 17 Oct 2004 12:03:20 -0700
X-IronPort-AV: i="3.85,147,1094454000";
v="EICAR-AV-Test'0'v";
d="scan'208"; a="86:adNrHT37924848"
X-IronPort-RCPT-TO: alan@mail.example.com
From: Adam@Outside.COM
To: Alan Alpha <Alan@mail.example.COM>
```

```
Subject: Exercise 7a Anti-Virus Scanning
Reply-To: Adam Alpha <adam@outside.com>
Date: Sun, 17 Oct 2004 12:02:39 -0700
MIME-version: 1.0
Content-type: multipart/mixed; boundary="IronPort"

--IronPort
Content-type: text/plain; format=flowed; charset=us-ascii
Content-transfer-encoding: 7bit

Blah blah blah blah blah
Blah blah blah blah blah
Blah blah blah blah blah
...
--IronPort
Content-type: text/plain
Content-transfer-encoding: 7bit
Content-disposition: inline

X5O!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-
FILE!$H+H*">X5O!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-FILE!$H+H*

--IronPort--
```

Quando os arquivos mbox-formatados são analisados gramaticalmente, é desejável não ler demasiada semântica no “” da linha que separa mensagens. Porque muitas utilidades diferentes redigem arquivos do mbox, há uma variação considerável nestas linhas. Contudo, “” da linha pode sempre ser usado como uma linha do separador da mensagem a fim indicar confiantemente que uma mensagem nova começou no arquivo do mbox. Em tudo, há aproximadamente 20 formatos conhecidos para as cordas após “” do separador da mensagem, que faz geralmente muito difícil o analisar gramaticalmente.

Depois que “” da linha é um mensagem de Email no formato do RFC 2822, com uma série de encabeçamentos do corpo da mensagem seguidos por uma linha em branco seguida pelo índice adicional do corpo da mensagem.

A fim assegurar-se de que as mensagens estejam separadas corretamente, as linhas de que comece com a corda “” prepended sempre por um único “**>**”. Várias variações diferentes das linhas do punho de arquivos do mbox que começam com o “**>**From” diferentemente. Nas implementações precoces dos aplicativos que redigiram arquivos do mbox, estas linhas elas mesmas não foram citadas. Os arquivos de registro de AsyncOS prepend sempre “**>**” às linhas por que comece com o uns ou vários “**>**” os caráteres seguidos “de”.

Está aqui um exemplo de um arquivo do mbox que contenha uma mensagem que tenha as linhas de que contenha começar amarra “”, “**>**From” e “**>>>>**From” nele:

```
From jtrumbo@example1.com Sun Dec 12 12:27:33 2004
X-IronPort-RCPT-TO: trumbo@example1.com
From: jtrumbo@example1.com
To: trumbo@example2.com
Subject: Quote this, if you dare
Date: Sun, 12 Dec 2004 12:28:00 -0700

The following line is just From
>From A From Line

The following line has quoted >From
>>From A >From Line
```

```
The following line has many >>>>From
>>>>>From This line has 4 > characters before From

And this is the last line
```

A extremidade de uma mensagem em um erro de arquivo do mbox é sinalizada tradicionalmente por uma linha em branco. Contudo, isto não está sempre atual (embora AsyncOS o coloca lá). Quando um arquivo do mbox-formato é analisado gramaticalmente, você deve sinalizar a extremidade de uma mensagem pelo começo de uma mensagem nova (suprima da linha em branco se uma esta presente) ou para o fim do arquivo.

Uma outra variação no formato do mbox chamado para o comprimento da mensagem a ser sinalizada em um campo do “Índice-comprimento” dentro do cabeçalho da mensagem. Esse formato não se usou “” da linha citar. AsyncOS não usa este formato e não introduz um campo do Índice-comprimento.