Wat is UNIX mbox (postvak) formaat?

Inhoud

Inleiding

Dit document beschrijft Unix brievenbus (vakje) formaat en hoe het op gebruik op de Cisco e-mail security applicatie (ESA) betrekking heeft.

Wat is UNIX mbox (postvak) formaat?

UNIX mbox formaat wordt gebruikt door AsyncOS wanneer er berichten worden gearchiveerd en inlogd in de actie Blog() van het bericht filter. "Archive Message" is een aanvullende configuratieoptie voor Ironport Anit-spam (IPAS), Anti-virus (Sofo's en McAfee), Advanced Malware Protection (AMP) en Graymail op het ESA.

Mbox formaat is een ASCII-formaat (dat wil zeggen niet binair) dat nul of meer e-mailberichten kan bevatten. Berichten worden aaneengezet in het veldbestand en kunnen apart worden geprikt op basis van specifieke strings in het bestand. Dit formaat is identiek aan het bericht aangezien ze worden overgebracht tussen RFC 2821 conforme postpoorten.

Elk bericht in box formaat begint met een lijn die met de string "From" (ASCII tekens F, r, o, m en space) begint. 'Van'-regels worden meerdere velden gevolgd: envelop-zender, datum en (optioneel) meer gegevens.

Het eerste veld na de 'From' string is de enveloppe-zender van het bericht. Afhankelijk van welke toepassing het box bestand maakt, kan de envelop-zender aanwezig zijn als een echte postvak of kan het een ander teken of string zijn. Meestal merkt u dat een "-" (enkele tekenstreepje) de map-sender vervangt als de eigenlijke enveloppe-sender niet beschikbaar of niet bekend is. Het door de ESA ingevoegd datumveld is in het UNIX asctime() formaat en is altijd 24 tekens lang. In sommige doosbestanden die door niet-AsyncOS implementaties zijn geschreven, volgt verdere informatie de datumstempel. Deze drie velden worden van elkaar gescheiden door één ruimte.

Hier is een voorbeeld van een postvak bestand met één bericht erin:

```
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
```

```
--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--
```

Wanneer box-geformatteerde bestanden worden geparsed, is het beter niet te veel semantiek in de "From" regel te lezen die berichten scheidt. Omdat veel verschillende nutsbedrijven mbox-bestanden schrijven, is er een aanzienlijke variatie in deze regels. Echter, "Van" lijn kan altijd worden gebruikt als berichtscheidingslijn om betrouwbaar aan te geven dat een nieuw bericht is begonnen in het veldbestand. In totaal zijn er ongeveer 20 bekende formaten voor de strings na het bericht van "From", wat het over het algemeen heel moeilijk maakt om ze te parseren.

Nadat de regel "Van" een e-mailbericht in RFC 2822-indeling is, met een reeks berichtselkop gevolgd door een lege regel gevolgd door extra berichttekst.

Om er zeker van te zijn dat de berichten goed van elkaar worden gescheiden, worden de lijnen die beginnen met de string "From" altijd met één enkele "**>**" toegevoegd. Verschillende varianten van mbox dossiers behandelen lijnen die met "**>**Van" anders beginnen. In vroege implementaties van toepassingen die mbox bestanden schreven, werden deze lijnen zelf niet geciteerd. AsyncOS-logbestanden maken altijd een "**>**" voor lijnen die beginnen met een of meer "**>**" tekens, gevolgd door "Van".

Hier is een voorbeeld van een veldbestand dat een bericht bevat met regels die de beginkoorden "Van", "**>**Van" en "**>>>**Van" erin bevatten:

```
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
```

Het einde van een bericht in een veldformaat wordt traditioneel gemarkeerd door een lege regel.

Dit is echter niet altijd aanwezig (hoewel AsyncOS het daar wel plaatst). Wanneer een veld-formaat bestand wordt geparseerd, dient u het einde van een bericht te signaleren aan het begin van een nieuw bericht (verwijder de lege regel indien deze aanwezig is) of aan het einde van het bestand.

Een andere variant in de notatievorm vroeg om de lengte van het bericht te laten signaleren in een veld "Content-Lengte" in de berichtkop. In die indeling werd geen regel "Van" geciteerd. AsyncOS gebruikt deze indeling niet en voegt geen veld van de contentlengte toe.