

UNIX mbox (メールボックス) 形式とは何ですか。

目次

[質問 :](#)

質問 :

UNIX mbox (メールボックス) 形式とは何ですか。

UNIX mbox 形式は、メッセージがアーカイブされ (アンチスパムおよびアンチウイルス設定)、ログに記録される (メッセージ フィルタの log() アクション) 場合に AsyncOS で使用されます。

mbox 形式とは、0 個以上のメール メッセージを含むことができる ASCII 形式 (バイナリではない) のファイル形式です。メッセージは mbox ファイル内で連結され、ファイル内の特定の文字列に基づいて分けられます。この形式は、RFC 2821 準拠のメール ゲートウェイ間で転送されるメッセージと同じです。

mbox 形式の各メッセージは、先頭に文字列「From」 (ASCII 文字の F、r、o、m、スペース) が付いた行で始まります。「From」行の後に、エンベロープ送信者、日付、その他のデータ (任意) など、複数のフィールドが続きます。

「From」文字列の後の最初のフィールドは、メッセージのエンベロープ送信者です。mbox ファイルを作成するアプリケーションに応じて、エンベロープ送信者は実際のメールボックスのままに表示されるか、別の文字または文字列になります。通常、実際のエンベロープ送信者が使用できないか不明である場合は、エンベロープ送信者が「-」 (半角ダッシュ) に置き換えられます。ESA によって挿入されるデータ フィールドは、UNIX 標準の asctime() 形式で、常に 24 文字の長さです。AsyncOS 実装以外によって書き込まれる mbox ファイルでは、日付スタンプの後に詳細が続きます。これらの 3 つのフィールドは単一のスペースで区切られます。

単一のメッセージを含む mbox ファイルの例を次に示します。

```
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
  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:
```

```
X50!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-
FILE!$H+H*">X50!P%@AP[4\PZX54(P^)7CC)7}$EICAR-STANDARD-ANTIVIRUS-TEST-FILE!$H+H*
```

```
--IronPort--
```

mbox 形式のファイルを解析する場合、メッセージを分ける「From」行にセマンティクスを読み込みすぎないことをお勧めします。さまざまなユーティリティが mbox ファイルを書き込むため、これらの行にはかなりのバリエーションがあります。ただし、常にメッセージの区切り行として使用される「From」行は、mbox ファイルで新しいメッセージが開始されたことを明確に示します。メッセージ区切りの「From」の後の文字列には、既知の形式が全部で約 20 あります。たいていの場合、これにより解析は非常に困難になります。

「From」行の後には、一連のメッセージ本文のヘッダー、空白行、さらにメッセージ本文の内容と続く RFC 2822 形式の電子メールメッセージが続きます。

メッセージが適切に区切られるように、文字列「From」で始まる行の先頭には常に「>」が 1 つ付加されます。さまざまなバリエーションの mbox ファイルが「>From」で始まる行をそれぞれ異なる方法で処理します。mbox ファイルを書き込むアプリケーションの初期の実装では、これらの行自体は記載されませんでした。AsyncOS のログ ファイルでは、1 つ以上の「>」文字の後に「From」が続く文字列で始まる行の先頭に、常に「>」が付加されるようになります。

次の例では、メッセージ内の行に「From」、「>From」、「>>>>From」が開始文字列として含まれている mbox ファイルを示します。

```
From jtrumbo@example1.com Sun Dec 12 12:27:33 2004
X-IronPort-RCPT-TO: trumbo@example1.com
From: jtrumbo@example1.com
      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
```

従来、mbox 形式ファイルのメッセージの最後は空白行で示されます。ただし、空白行が必ず使用されるわけではありません (AsyncOS では配置されます)。mbox 形式ファイルを解析する場合は、新しいメッセージの開始 (空白行がある場合は削除する) またはファイルの終了によって

、メッセージの最後を示す必要があります。

mbox 形式の別のバリエーションでは、メッセージヘッダー内の「Content-Length」フィールドにメッセージの長さを示す必要がありました。この形式では、引用符付きの「From」行は使用されませんでした。AsyncOS ではこの形式を使用せず、「Content-Length」フィールドも挿入しません。