

Newton Linux User Group – Bash Scripting Basics

Summary

Someone suggested the topic "Writing Shell Scripts" or more specifically, how to use command line tools such as xargs, awk, sed, and grep. We will cover a real world example where we have a directory filled with quarantined email messages (spam) stored as text files. We want to extract specific info from each file (ie: Subject, From, Date, Spam-Score headers) in some automated way if the data is available.

Bash Scripting References:

Basic Bash Scripting:

<http://mirror.datility.net/pub/linux/ldp/LDP/Bash-Beginners-Guide/html/index.html>

Advanced Bash Scripting:

<http://mirror.datility.net/pub/linux/ldp/LDP/abs/html/index.html>

Our Script

```
#!/bin/bash

echo "Extracting data from quarantined items..."
echo ""

for message in message.* ; do

    echo "Processing file: $message"
    echo ""

    FROM=`sed -n '/From:/p' $message`
    DEST=`sed -n '/To:/p' $message`
    SUBJ=`sed -n '/Subject:/p' $message`
    SPAM=`sed -n '/X-MX-Spam:/p' $message`

    # Removing 'Reply-To:' and 'Deliver-To:' headers
    # Using 'printf' to preserve line feeds
    DEST=`printf "$DEST" | sed /\-To:/d`

    FROM=`echo $FROM | awk 'BEGIN { FS=":" } { print $2 }'`
    DEST=`echo $DEST | awk 'BEGIN { FS=":" } { print $2 }'`
    SUBJ=`echo $SUBJ | awk 'BEGIN { FS=":" } { print $2 }'`
    SPAM=`echo $SPAM | awk 'BEGIN { FS=":" } { print $2 }'`

    echo "FROM:$FROM"
    echo "DEST:$DEST"
    echo "SUBJ:$SUBJ"
    echo "SPAM:$SPAM"

    echo ""

done

echo ""
echo "Finished extracting data!"
echo ""
```

Our Output

```
[jklein@web-d02 spamquar]$ ./extract.sh
Extracting data from quarantined items...
```

```
Processing file: message.1
```

```
FROM: "Shannon Ivey" <vmrbt@ccdvd.com>
DEST: sales@example.net
SUBJ: Get nominated for MBA
SPAM: exempt
```

```
Processing file: message.2
```

```
FROM: "save@familysavings.biz" <save@familysavings.biz>
DEST: ceo@example.net
SUBJ: Free vacation + Airline, Hotel, Theme Park Discounts
SPAM: exempt
```

```
Processing file: message.3
```

```
FROM: "Lourdes Sachdeva" <imbalancestalent@dpcbell.net>
DEST: ceo@example.net
SUBJ: anti-i^mpotence drug to win a*pproval from the u.s. f d a
SPAM: exempt
```

```
Finished extracting data!
```

Our Data Files

```
[jklein@web-d02 spamquar]$ cat message.1
Return-Path: <vmrbt@ccdvd.com>
Delivered-To: catchall@example.net
Received: (qmail 21295 invoked by uid 107); 8 Jun 2004 19:58:10 -0000
Received: from c1m6.emaildefenseservice.com (216.40.36.38)
    by mx3.datility.net with SMTP; 8 Jun 2004 19:58:10 -0000
Received: from dl37-186-172-214.abhsia.telus.net [137.186.172.214] (HELO dl37-186-172-
214.abhsia.telus.n)
    by c1m6.emaildefenseservice.com (mx1_mta-1.3.7-29) with SMTP id
9aa16c04.6128.030.c1m6;
    Tue, 08 Jun 2004 15:59:37 -0400 (EDT)
Received: from 24.241.224.252 by 137.186.172.214; Tue, 08 Jun 2004 16:59:40 -0400
Message-ID: <TTVYYVOMAETTDYBPLKUE@gdf.com>
From: "Shannon Ivey" <vmrbt@ccdvd.com>
Reply-To: "Shannon Ivey" <vmrbt@ccdvd.com>
To: sales@example.net
Subject: Get nominated for MBA
Date: Tue, 08 Jun 2004 21:53:40 +0100
MIME-Version: 1.0
Content-Type: multipart/alternative;
    boundary="--29317920244797388"
X-MX-Spam: exempt
X-MX-MAIL-FROM: <vmrbt@ccdvd.com>
X-MX-SOURCE-IP: [137.186.172.214]
```

```
I am a sample message #1. I rock.
```

```
[jklein@web-d02 spamquar]$ cat message.2
Return-Path: <loumredhead@wmconnect.com>
Delivered-To: ceo@example.net
Received: (qmail 11267 invoked by uid 107); 5 Jun 2004 09:41:26 -0000
Received: from c1m6.emaildefenseservice.com (216.40.36.38)
    by mx3.datility.net with SMTP; 5 Jun 2004 09:41:26 -0000
Received: from unknown [66.205.218.24] (EHLO localhost)
    by c1m6.emaildefenseservice.com (mx1_mta-1.3.7-29) with ESMTTP id
47591c04.4123.047.c1m6;
    Sat, 05 Jun 2004 05:42:12 -0400 (EDT)
From: "save@familysavings.biz" <save@familysavings.biz>
To: ceo@example.net
Subject: Free vacation + Airline, Hotel, Theme Park Discounts
MIME-Version: 1.0
Content-type: text/html
Message-Id: <20040605134419.817D128F351@localhost>
Date: Sat, 5 Jun 2004 09:44:19 -0400 (EDT)
X-MX-Spam: exempt
X-MX-MAIL-FROM: <loumredhead@wmconnect.com>
X-MX-SOURCE-IP: [66.205.218.24]
```

This is sample message #2.

```
[jklein@web-d02 spamquar]$ cat message.3
Return-Path: <imbalancestalent@pacbell.net>
Delivered-To: ceo@example.net
Received: (qmail 15989 invoked by uid 107); 3 Jun 2004 19:14:44 -0000
Received: from c1m6.emaildefenseservice.com (216.40.36.38)
    by mx3.datility.net with SMTP; 3 Jun 2004 19:14:44 -0000
Received: from csd.k12.ar.us [165.29.2.62] (EHLO csdweb.k12.ar.us)
    by c1m6.emaildefenseservice.com (mx1_mta-1.3.7-29) with ESMTTP id
db87fb04.21425.210.c1m6;
    Thu, 03 Jun 2004 15:15:09 -0400 (EDT)
Received: from courtroom (r200-40-238-206-dialup.adsl.anteldata.net.uy
[200.40.238.206]) by csdweb.k12.ar.us with SMTP (Microsoft Exchange Internet Mail
Service Version 5.5.2653.13)
    id MA221VTY; Thu, 3 Jun 2004 14:11:09 -0500
From: "Lourdes Sachdeva" <imbalancestalent@pacbell.net>
To: ceo@example.net
Subject: anti-i^mpotence drug to win a*pproval from the u.s. f d a
Mime-Version: 1.0
Content-Type: text/html; charset=us-ascii
Content-Transfer-Encoding: 7bit
X-MX-Spam: exempt
X-MX-MAIL-FROM: <imbalancestalent@pacbell.net>
X-MX-SOURCE-IP: [165.29.2.62]
```

ceo:

C1AL`IS & LEV1`TRA works in as little as 30 minutes and lasts for up to 36 HOURS.

- 1: 80+% savings O_rdering !
- 2: No p`rescription required .
- 3: DOCTOR & F_D_A a.pproved !
- 4: Overnight s|hipping !

P~1~a~c~e YOur 0~r~d~e~r H~e~r~e Today