Long Email Headers (an overview)

Long Email Headers (an overview)

What is a long email header?

A long email header (or full email header) is the log of an email's journey from one email server to another. When a piece of email travels from a sender to a recipient, mail isn't simply put into the person's mailbox. Rather, the email usually goes through a list of servers which hand off the email to other servers before reaching the recipient. Because a long email header is a log of this journey, they are usually helpful in detecting the sender of spam mail.

Before going into the specifics, let's first find an email header...

Where can I find long email headers?

Directions on how to find the long email header can be found here.

Now let's take a look at some email headers...

What do long email headers tell me?

The most relevant parts of email headers when looking through the path an email traveled is the **Received:** sections.
In this test email sent from my Yahoo account to my Gmail account, you can see the path an email travels by looking at the **Received:** lines.

```
Delivered-To: lightningdragons@gmail.com
Received: by 10.100.153.5 with SMTP id a5cs174126ane;
Mon, 6 Jul 2009 07:38:22 -0700 (PDT)
Received: by 10.142.199.15 with SMTP id w15mr1581803wff.78.1246891101583;
Mon, 06 Jul 2009 07:38:21 -0700 (PDT)
Return-Path: <calvin_x_hobbes@yahoo.com>
Received: from web57701.mail.re3.yahoo.com (web57701.mail.re3.yahoo.com (68.142.236.53))
by mx.google.com with SMTP id 12si3311531yxe.126.2009.07.06.07.38.20;
Mon, 06 Jul 2009 07:38:20 -0700 (PDT)
Received-SPF: pass (google.com: domain of calvin_x_hobbes@yahoo.com designates 68.142.236.53 as permitted sender)
client-ip=68.142.236.53;
Authentication-Results: mx.google.com; spf=pass (google.com: domain of calvin_x_hobbes@yahoo.com designates 68.142.236.53
as permitted sender) smtp.mail=calvin_x_hobbes@yahoo.com; dkim=pass (test mode) header.i=@yahoo.com
Received: (qmail 55301 invoked by uid 60001); 6 Jul 2009 14:38:19 -0000
... (continued)
```

This area of the email header tells us that the email has gone through 4 steps from sender to recipient:

1. **Starting from the sender, the email is sent out via qmail.**
   Starting from the bottom, the first received message indicates:

```
(qmail 55301 invoked by uid 60001); 6 Jul 2009 14:38:19 \-0000
```

Yahoo's mail system uses the qmail program, which is yahoo's program of mail delivery. More information on the qmail system can be found here.
The time and date following this message indicates the time this process was carried out.

1. **Google verifies the authenticity of this message.**
   Next, Gmail verifies the authenticity of the sender through **Receive-SPF** (*Sender Policy Framework*):

```
Received-SPF: pass ...; Authentication-Results: mx.google.com;
spf=pass (google.com: domain of calvin_x_hobbes@yahoo.com designates
68.142.236.53 as permitted sender...
```

In this case, this email has passed since yahoo's IP address was recognized as a permitted sender to google.com emails. Here the email has noted for us the IP address as the sender, which can be used to track the true sender, were it not from the sender it was said to be from, `calvin_x_hobbes@yahoo.com`.

2. **Google receives this message from yahoo.**
   After having verified the authenticity, Google now accepts this message:

   ```
   Received: from web57701.mail.re3.yahoo.com (web57701.mail.re3.yahoo.com (68.142.236.53))
   by mx.google.com with SMTP id 12si3311531yxe.126.2009.07.06.07.38.20;
   Mon, 06 Jul 2009 07:38:20 \-0700 (PDT)
   ```

   Here, the mail server that sends the email is `web57701.mail.re3.yahoo.com` and the server that receives this mail is `mx.google.com`. The server which sent the email calls itself `web57701.mail.re3.yahoo.com`. The server receiving this email notes the IP address of the server that sends this mail (68.142.236.53) and does a reverse DNS lookup on it, showing that the true server name is `web57701.mail.re3.yahoo.com`, which is what the original server identified itself as. This message is received via SMTP, (Simple Mail Transfer Protocol), and is assigned a temporary ID of `12si3311531yxe.126.2009.07.06.07.38.20` for this transfer. The transfer occurs at the specified time and date.

3. **Google sends the message to the individual's mailbox server.**
   Having accepted this mail, Google transfers this email through a few more servers before reaching the destination mailbox.

   ```
   Received: by 10.100.153.5 with SMTP id a5cs174126ane;
   Mon, 6 Jul 2009 07:38:22 \-0700 (PDT)
   Received: by 10.142.199.15 with SMTP id w15mr1581803wff.78.1246891101583;
   Mon, 06 Jul 2009 07:38:21 \-0700 (PDT)
   ```

Looking at a Piece of Spam

Now let's take a look at a spam email that has been sent straight to my Spam Folder:

```
Delivered-To: lightningdragons@gmail.com
Received: by 10.100.153.5 with SMTP id a5cs283861ane;
Tue, 7 Jul 2009 13:56:17 -0700 (PDT)
Received: by 10.143.5.20 with SMTP id h20mr2094156wfi.279.1247000174962;
Tue, 07 Jul 2009 13:56:14 -0700 (PDT)
Return-Path: <MauriciowilsonJuarez@energybulletin.net>
Received: from south-station-annex.mit.edu (SOUTH-STATION-ANNEX.MIT.EDU (18.72.1.2))
by mx.google.com with ESMTP id 28si2068780wfd.4.2009.07.07.13.56.13;
Tue, 07 Jul 2009 13:56:14 -0700 (PDT)
Received-SPF: neutral (google.com: 18.72.1.2 is neither permitted nor denied by best guess record for domain of
MauriciowilsonJuarez@energybulletin.net) client-ip=18.72.1.2;
Authentication-Results: mx.google.com; spf=neutral (google.com: 18.72.1.2 is neither permitted nor denied by best guess record for
domain of MauriciowilsonJuarez@energybulletin.net) smtp.mail=MauriciowilsonJuarez@energybulletin.net
Received: from fort-point-station.mit.edu (FORT-POINT-STATION.MIT.EDU (18.7.7.76))
by south-station-annex.mit.edu (8.13.6/8.9.2) with ESMTP id n67KuAj9006970;
Tue, 7 Jul 2009 16:56:10 -0400 (EDT)
Received: from mit.edu (W92-130-BARRACUDA-2.MIT.EDU (18.7.21.223))
by fort-point-station.mit.edu (8.13.6/8.9.2) with ESMTP id n67Ku0X1023450
for <ultimate@mit.edu>; Tue, 7 Jul 2009 16:56:01 -0400 (EDT)
Received: from christine (localhost (127.0.0.1))
by mit.edu (Spam Firewall) with SMTP
id 401FE235BA78; Tue, 7 Jul 2009 16:55:58 -0400 (EDT)
Received: from christine (22.227.9-93.rev.gaoland.net (93.9.227.22)) by mit.edu with SMTP id SNBnBwgPwPvyjCM6; Tue, 07 Jul
2009 16:55:58 -0400 (EDT)
X-Originating-IP: (147.159.192.139)
X-Originating-Email: (FedericoparentageLara@energybulletin.net)
X-Sender: VinceaerogeneShannon@energybulletin.net
Received: (qmail 1906 by uid 239); Tue, 7 Jul 2009 22:55:48 -0100
Message-Id: <20090622743.2356.qmail@energybulletin.net>
To: <apac@mit.edu>, <darons@mit.edu>, <ultimate@mit.edu>, <cwilcox@mit.edu>,
<agarrawa@mit.edu>
Subject: Your order approved and shipped
From: Sebastian Sexton <QuincyparapetWinters@energybulletin.net>
MIME-Version: 1.0
Importance: High
Content-Type: text/html
Date: Tue, 7 Jul 2009 16:55:58 -0400 (EDT)
X-Spam-Score: 9.002
X-Spam-Level: ********* (9.002)
X-Spam-Flag: YES
X-Scanned-By: MIMEDefang 2.42
... more
```

The sender of this email says he/she is from energybulletin.net, which is a legitimate site. However, this email is obviously not from this sender (You can't purchase anything off EnergyBulletin.net and I certainly didn't buy anything from there...). How can we trace who it's from?

If the proper information is available, this isn't so hard to find out. We simply look at the first **Received:** header which offers detailed information about servers:

```
Received: from christine (22.227.9-93.rev.gaoland.net (93.9.227.22)) by mit.edu
with SMTP id SNBnBwgPwPvyjCM6; Tue, 07 Jul 2009 16:55:58 \-0400 (EDT)
```

Here, the sending server identifies itself as "Christine." However, a reverse DNS done by the receiving server (mit.edu) shows that the server is a part of gaoland.net and has an IP address of 93.9.227.22. If we check ourselves using a reverse DNS lookup, we find that this IP address originates from France. But from my DNS lookup, this IP address indicates a possibly forged server.

Conclusion
While we are unable to pinpoint a sender, we did find out the approximate origin of this email. Seeing as how I know no one from France, I can only assume this as a random spam mail. But while the path can tell us who sent it, we can forgo some of this work by looking at MIT's spam filter system.

Spam Filtering

✓ **MIT's Email Architecture**
For a more visual explanation of the servers involved in MIT's mail filtering, reference Jacob's powerpoint.

Now let's take a look at another part of the email header: the Spam Level and Score. When an MIT email is recognized as spam, there are generally a few things on the long email header that will indicate that it is spam and the extent to which that email is "spammy." Specifically, these parts:

For Webmail...
... **X-Spam-Score**: 9.002 **X-Spam-Level**: ******** * (9.002) **X-Spam-Flag**: YES ...

When email passes through MIT's central mail hub, email goes through a spam filter named Spam Assassin. This spam filter scores the email, indicating how much that email resembles spam. This is done by adding a few extra headers, including `X-Spam-Score`, `X-Spam-Level` and `X-Spam-Flag`. If this score passes a certain level, this email is blocked and given a spam flag when it reaches your mailbox. The higher the number of the spam score, the more the email will be flagged as spam.

References and Resources:

About Email Headers
More about Email Headers
Reverse DNS Lookup
Another Reverse DNS Lookup