

Pandora

(a case study in reverse engineering)

Mark Meiss

(mmeiss@indiana.edu)

March 7, 2006

Web2.0 and AJAX

➤ What is Web2.0?

- Nobody has a concrete notion.
- Something to do with *participatory* and *flexible* sites that let people *customize* them.
- Examples: *Flickr*, *Google Maps*, *Pandora*

➤ What is AJAX?

- *Asynchronous Javascript And XML*
- Pages that *contact the server again* without firing up Java or playing weird `<iframe>` games

Pandora



- Tell it an *artist* or *song* you like
- You get an endless “streaming” *Internet radio station* in return
- You can’t *rewind* or *save* the songs

User interface is in Flash

➤ Why would they use Flash instead of...

- ...*Java?*
 - Web users hate Java.
- ...*Javascript?*
 - No flashy interface.
 - Exposure of source code.
 - Obligation to prevent downloading.
- ...*[your plugin here]?*
 - Users don't have it.
 - Users don't want it.

Taking a look at the traffic

The screenshot displays the Wireshark interface with a packet capture named 'barracuda.dump'. The main pane shows a list of 23 captured packets. The selected packet (No. 1) is a TCP SYN packet from 156.56.103.1 to 209.10.40.106. The packet details pane shows the Ethernet II header, Internet Protocol header, and Transmission Control Protocol header. The packet bytes pane shows the raw data in hexadecimal and ASCII.

No.	Time	Source	Destination	Protocol	Info
1	0.000000	156.56.103.1	209.10.40.106	TCP	37052 > http [SYN] Seq=0 Ack=0 Win=5840 Len=0 MSS=1460 TSV=17
2	0.079907	209.10.40.106	156.56.103.1	TCP	http > 37052 [SYN, ACK] Seq=0 Ack=1 Win=8000 Len=0 MSS=1460
3	0.079953	156.56.103.1	209.10.40.106	TCP	37052 > http [ACK] Seq=1 Ack=1 Win=5840 Len=0
4	0.080060	156.56.103.1	209.10.40.106	HTTP	POST /radio/xmlrpc/v4?rid=2777855P&lid=5200019&method=search&
5	0.159985	209.10.40.106	156.56.103.1	TCP	http > 37052 [ACK] Seq=1 Ack=488 Win=8000 Len=0 MSS=1460
6	0.160007	156.56.103.1	209.10.40.106	HTTP	Continuation or non-HTTP traffic
7	0.239974	209.10.40.106	156.56.103.1	TCP	http > 37052 [ACK] Seq=1 Ack=1111 Win=8000 Len=0 MSS=1460
8	0.240350	209.10.40.106	156.56.103.1	TCP	http > 37052 [ACK] Seq=1 Ack=488 Win=6432 Len=0
9	0.240475	209.10.40.106	156.56.103.1	TCP	http > 37052 [ACK] Seq=1 Ack=1111 Win=8099 Len=0
10	0.345507	209.10.40.106	156.56.103.1	HTTP	HTTP/1.0 200 OK
11	0.345535	156.56.103.1	209.10.40.106	TCP	37052 > http [ACK] Seq=1111 Ack=1461 Win=8760 Len=0
12	0.345548	209.10.40.106	156.56.103.1	HTTP	Continuation or non-HTTP traffic
13	0.345565	156.56.103.1	209.10.40.106	TCP	[TCP Dup ACK 11#1] 37052 > http [ACK] Seq=1111 Ack=1461 Win=8
14	0.345569	209.10.40.106	156.56.103.1	HTTP	Continuation or non-HTTP traffic
15	0.345578	156.56.103.1	209.10.40.106	TCP	37052 > http [ACK] Seq=1111 Ack=3082 Win=11680 Len=0
16	0.345784	156.56.103.1	209.10.40.106	TCP	37052 > http [FIN, ACK] Seq=1111 Ack=3082 Win=11680 Len=0
17	0.425751	209.10.40.106	156.56.103.1	TCP	http > 37052 [ACK] Seq=3082 Ack=1112 Win=8099 Len=0
18	4.838796	156.56.103.1	209.10.40.106	TCP	37053 > http [SYN] Seq=0 Ack=0 Win=5840 Len=0 MSS=1460 TSV=17
19	4.918591	209.10.40.106	156.56.103.1	TCP	http > 37053 [SYN, ACK] Seq=0 Ack=1 Win=8000 Len=0 MSS=1460
20	4.918627	156.56.103.1	209.10.40.106	TCP	37053 > http [ACK] Seq=1 Ack=1 Win=5840 Len=0
21	4.918710	156.56.103.1	209.10.40.106	HTTP	POST /radio/xmlrpc/v4?rid=2777855P&lid=5200019&method=createS
22	4.998961	209.10.40.106	156.56.103.1	TCP	http > 37053 [ACK] Seq=1 Ack=495 Win=8000 Len=0 MSS=1460
23	4.998978	156.56.103.1	209.10.40.106	HTTP	Continuation or non-HTTP traffic

Frame 1 (74 bytes on wire, 74 bytes captured)
Ethernet II, Src: 00:30:48:20:1d:9e, Dst: 00:0f:35:b2:b9:00
Internet Protocol, Src Addr: 156.56.103.1 (156.56.103.1), Dst Addr: 209.10.40.106 (209.10.40.106)
Transmission Control Protocol, Src Port: 37052 (37052), Dst Port: http (80), Seq: 0, Ack: 0, Len: 0

```
0000  00 0f 35 b2 b9 00 00 30 48 20 1d 9e 08 00 45 00  ..5....0 H ....E.
0010  00 3c bb e4 40 00 06 82 29 9c 38 67 01 d1 0a    <..@. .).8g...
0020  28 6a 90 bc 00 50 c0 3c 0e b0 00 00 00 00 a0 02  (j...P.< .....J
0030  16 d0 14 f1 00 00 02 04 05 b4 04 02 08 0a 0a 4a  .....T.....
0040  b5 54 00 00 00 01 03 03 00  ..T.....
```

File: barracuda.dump 233 MB 0 P: 267870 D: 267870 M: 0

Taking a look at the traffic

- Packet capture and analysis with *Ethereal*
 - Turn on capture
 - Fire up Pandora
 - Turn off capture
- Ethereal does *TCP stream reassembly*
- In a few minutes, we find:
 - The *IP address blocks* used by Pandora
 - All the traffic consists of *HTTP requests*
 - There is *no streaming* going on!

Client request

```
GET /access?version=4&token=615bWUj2NEB0bZuM%2Be9L%2F5nshW
cc9mHlgMSvGExw5kMs4kfv7KI%2BvGnFkopY97MYovAuNrrLJqoW%2BaP1
D1%2FfR90hEL42Wg4nPxrwc7InyeHD%2B9pEfJHJgISl4Jc j3NA%2FTv3h
m9WK21AgHJueG570aYpIzSEocNVj721Q%2BRZPxj9PNKd9urdGnZd4UzMi
Hb4c2IRlsaH4D%2Fs%3D&lid=5200019 HTTP/1.1
Host: audio-eqx-sjl04.pandora.com
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US;
rv:1.8.0.1) Gecko/20060124 Firefox/1.5.0.1
Accept: text/xml,application/xml,application/xhtml+xml,
text/html;q=0.9,text/plain;q=0.8,image/png,*/*;q=0.5
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
```

Server response

HTTP/1.1 200 OK

Date: Tue, 28 Feb 2006 20:26:39 GMT

Server: Apache

Cache-Control: no-cache, no-store, must-revalidate, max-age=-1

Pragma: no-cache, no-store

Expires: -1

Content-Length: 3736136

Connection: close

Content-Type: application/octet-stream

Extracting the file

- *Save stream* to file
- Use *vi* to *strip* off the HTTP headers
- *file* command shows
 - *MP3 data!*
 - Encoded with LAME, 128kbps Joint Stereo
- Given an *.mp3* extension, plays fine
- Lack of *ID3* information...

Another client request

```
POST /radio/xmlrpc/v4?rid=2777855P&lid=5200019&method=getFirstFragment&arg1=15483129
HTTP/1.1
Host: www.pandora.com
User-Agent: Mozilla/5.0 (X11; U; Linux i686; en-US; rv:1.8.0.1) Gecko/20060124
Firefox/1.5.0.1
Accept: text/xml,application/xml,application/xhtml+xml,text/html;q=0.9,text/plain;
q=0.8,image/png,*/*;q=0.5
Accept-Language: en-us,en;q=0.5
Accept-Encoding: gzip,deflate
Accept-Charset: ISO-8859-1,utf-8;q=0.7,*;q=0.7
Keep-Alive: 300
Connection: keep-alive
Cookie: vlad=1:5:0
Content-Type: text/xml
Content-Length: 592
```

```
e70ab41fe73bdfdb1ed6ce227c3e2ca43ad523f8963075faa0404a49a979eacf8a46bd5382d57be10c32530
775b454f2ca9ccd87f8a2bae991ee111a7616c7787f4f5a25aa0e0ff9e64f572a15eae934bdf5a29e9f8fc2
cbd4552a498ddf21791c0cc66c23b2e5d650314d2bb70f7893538e1b16f4279899d7ea387b2f2a0b4af4d1d
736ac336adbf3ec1a1eb753eb413ad5632cb6ee7f01f1f30f322c2a49b4bcd720b16b01a48adfdcdf0ebd4b
2547aae94960f738d790f97e677123bfa313d278a43ff9f78730ba3448773cafe283f4d1d736ac336adbf3e
c1a1eb753eb413ad5632cb6ee7f01792cc8c378c230093535bf05fc9303f495d14bf7dc8f66d53f8f99aa61
937aef075a766f66ae4199f7618628380f7247e956ca41636a1ff90d6c6fa565d9933f
```

Another server response

```
HTTP/1.0 200 OK
Date: Tue, 28 Feb 2006 20:26:38 GMT
Server: Jetty/5.1.3 (Linux/2.6.13.1-20050914 i386 java/1.4.1_07
Content-Type: text/xml
Content-Length: 10203
Connection: close
```

```
<?xml version="1.0" encoding="UTF-8"?><methodResponse><params><param><value><array><data><value><struct><member>
<name>amazonUrl</name><value>&path=search-handle-url%2Findex%3Dmusic%26field-artist%3DHeart%26field-title%3D
Jupiter%2527s%2BDarling</value></member><member><name>type</name><value>SEED</value></member><member><name>audio
URL</name><value>http://audio-eqx-sjl02.pandora.com/access?version=4&token=YGNWZzwutda%2F%2BECHwORbegXtT9QBS
ur5SBDR%2F5e8yZ0Be%2F3wXtYhxvhBHqGkhtGUG50wCBdpARjy74tL1Mbd6HPvrYh5rsq1Y0b7la4VaxZsSWxp%2BxCux%2Baw%2B%2BDz61y%2
FvtTUGBL4g7x3WMqVgPjzP%2B0aR9jO3uk%2BobN%2FUDf2E8XyExMB%2BUsNot4NginGqXUqSJOOmSZ1M8%3D</value></member><member>
<name>itunesUrl</name><value>&RD_PARM1=http%253A%252F%252Fphobos.apple.com%252FWebObjects%252FMZSearch.woa%2
52Fwa%252FadvancedSearchResults%253FartistTerm%253DHeart%252526%2526songTerm%253DOldest%252Bstory%252BIn%252BThe
%252BWorld%2526originStoreFront%253D143441%2526partnerId%3D30</value></member><member><name>focusTrait</name><va
lue></value></member><member><name>musicComUrl</name><value>http://search.music.com/?b=470286336%7C1073741824%7C
1551368192&pl=10%7C3%7C1&t=1551368192&queryType=batchPage&title=&name=&keywords=&q=H
eart+Jupiter%27s+Darling&querySrc=pandora</value></member><member><name>musicId</name><value>S170121</value>
</member><member><name>artRadio</name><value>http://images.pandora.com/images/amazon/8/2/3/5/616948195328_160W_1
60H.jpg</value></member><member><name>focusTraits</name><value><array><data><value><struct><member><name>focusTr
aitName</name><value>a subtle use of vocal harmony</value></member><member><name>focusTraitID</name><value>F5236
</value></member></struct></value><value><struct><member><name>focusTraitName</name><value>a vocal-centric aesth
etic</value></member><member><name>focusTraitID</name><value>F4788</value></member></struct></value><value><stru
ct><member><name>focusTraitName</name><value>minor key tonality</value></member><member><name>focusTraitID</name
><value>F4782</value></member></struct></value><value><struct><member><name>focusTraitName</name><value>electric
guitar riffs</value></member><member><name>focusTraitID</name><value>F1426</value></member></struct></value><va
lue><struct><member><name>focusTraitName</name><value>an emotional female lead vocal performance</value></member
><member><name>focusTraitID</name><value>F5188</value></member></struct></value><value><struct><member><name>foc
name><value>http://www.pandora.com/music/song/860ec159ca18c30d</value></member><member><name>artistSummary</name
```

...

Extracted XML data

albumTitle : Presence

amazonUrl : &path=search-handle-url%2Findex%3Dmusic%26field-artist%3DLed%2BZeppelin%26field-title%3DPresence

artRadio : http://images.pandora.com/images/amazon/8/2/9/3/075679243928_160W_159H.jpg

artistSummary : Led Zeppelin

audioURL : <http://audio-eqx-sjl01.pandora.com/access?version=4&token=YGNWZzwutda%2F%2BECHwORbegXtT9QBSur5SBDR%2F5e8yZ0Be%2F3wXtYhxhvBHqGkhtGUzSKg8Od5KQRzJq2Ksqbh%2BdmN05CWohDZXIKIn9ToZcUWKTguanWJMcPrLXqoO5nCzm53qgJkug9yj%2BLrvYleXJiZ2od95joxUo5ka656N%2BGLBJPEN0h0csfl1CmnqP2oFEZ9SKdXEJQ%3D>

focusTrait :

focusTraitId :

fragmentS2 : S27517

identity : 8b1bc1ea93cb46374fd2a3833c8f32a7

itunesUrl : &RD_PARM1=http%253A%252F%252Fphobos.apple.com%252FWebObjects%252FMZSearch.woa%252Fwa%252FadvancedSearchResults%253FartistTerm%253DLed%252BZeppelin%252526%2526songTerm%253DAchilles%252BLast%252BStand%2526originStoreFront%253D143441%2526partnerId%3D30

musicComUrl : http://search.music.com/?b=470286336%7C1073741824%7C1551368192&pl=10%7C3%7C1&t=1551368192&queryType=batchPage&title=&name=&keywords=&q=Led+Zeppelin+Presence&querySrc=pandora

musicId : S32945

songDetailURL : http://www.pandora.com/music/song/860ec159ca18c30d

songTitle : Achilles Last Stand

type : HIT

Putting it together

- Data capture: *libpcap*
- Stream reassembly: *libnids*
- XML parsing: *expat* and *XML::Parser*
- ID3 tagging: *id3tools*

Total time to write: about 2 hours
(C and Perl)

Detection

- The extraction is *completely passive*.
- They have *no way* at all of knowing.

Moral of the Story

- Obscurity *does not* provide security.
- Users *will* poke under the hood.
- Compromises take *hours*, not *weeks*.
- *One user* is enough; they can give it to *everyone*.