BERKELEY LAB

LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

Detecting Credential Spear-phishing Attacks at LBNL

Aashish Sharma
(Grant Ho, Mobin Javed, Vern Paxson, David Wagner)

September 2017

#BroCon2017

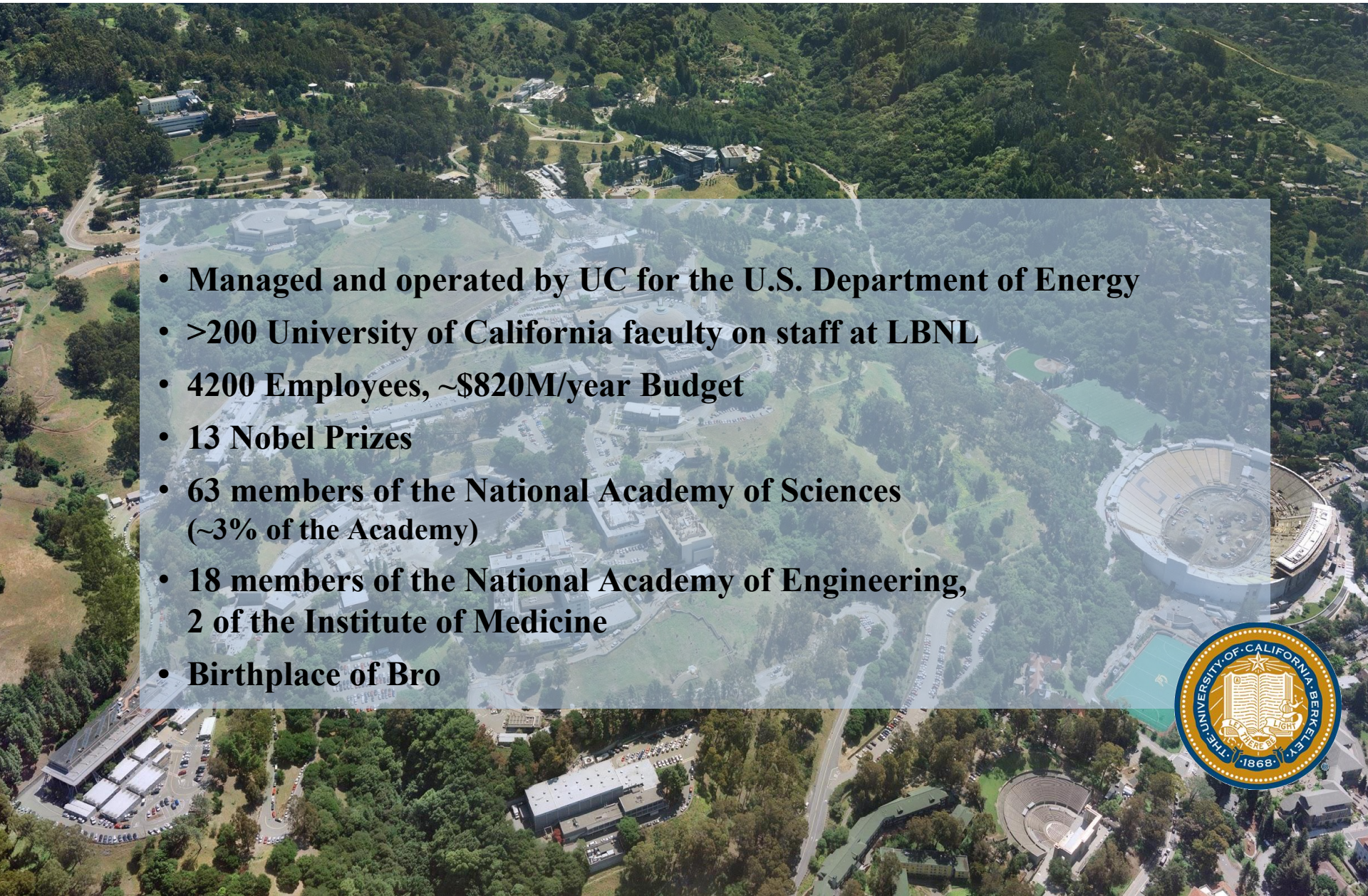
BERKELEY LAB

Lawrence Berkeley National Laboratory

U.S. DEPARTMENT OF ENERGY

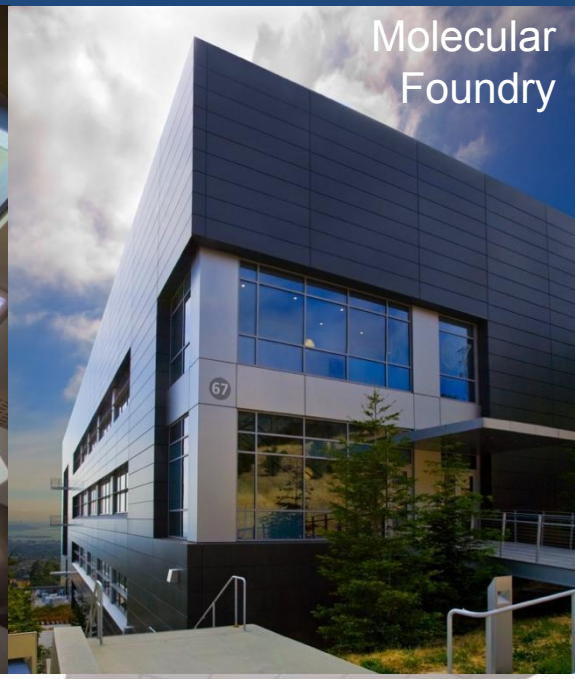
UNIVERSITY OF CALIFORNIA

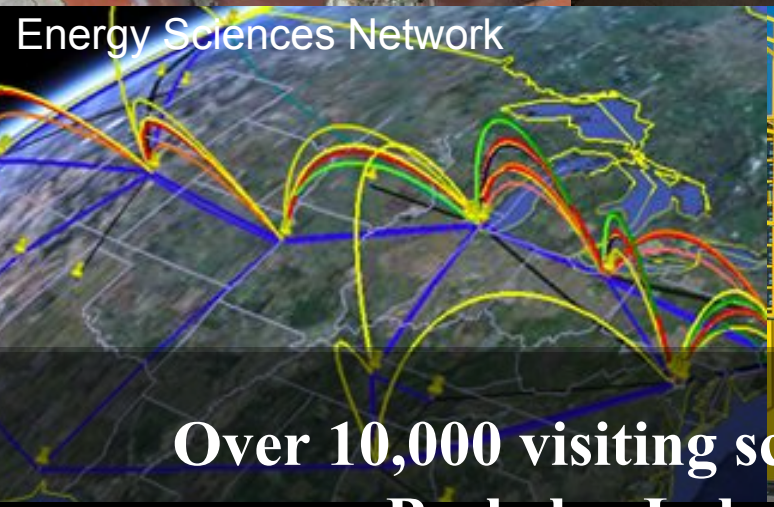
80 Years of World-Leading Team Science at Lawrence Berkeley National Laboratory

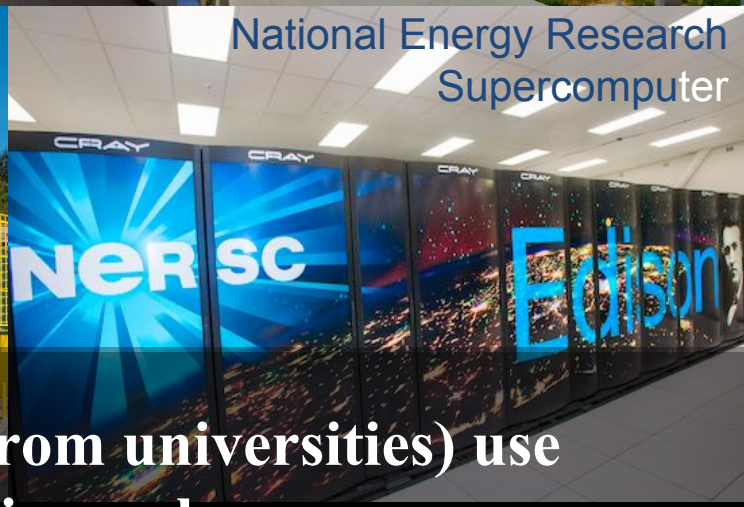


- **Managed and operated by UC for the U.S. Department of Energy**
- **>200 University of California faculty on staff at LBNL**
- **4200 Employees, ~$820M/year Budget**
- **13 Nobel Prizes**
- **63 members of the National Academy of Sciences (~3% of the Academy)**
- **18 members of the National Academy of Engineering, 2 of the Institute of Medicine**
- **Birthplace of Bro**

World-Class User Facilities
Serving the Nation and the World







Over 10,000 visiting scientists (~2/3 from universities) use Berkeley Lab research facilities each year

Overview

- Current state of SMTP
- Gaining visibility into SMTP
- New scripts
  - alerts and False positives
- Realtime detector design for detecting credential stealing spearphish
  - Persistence and reputation databases
  - Scalability
- Implementation and deployment challenges
- Whats next …..

MailFlow

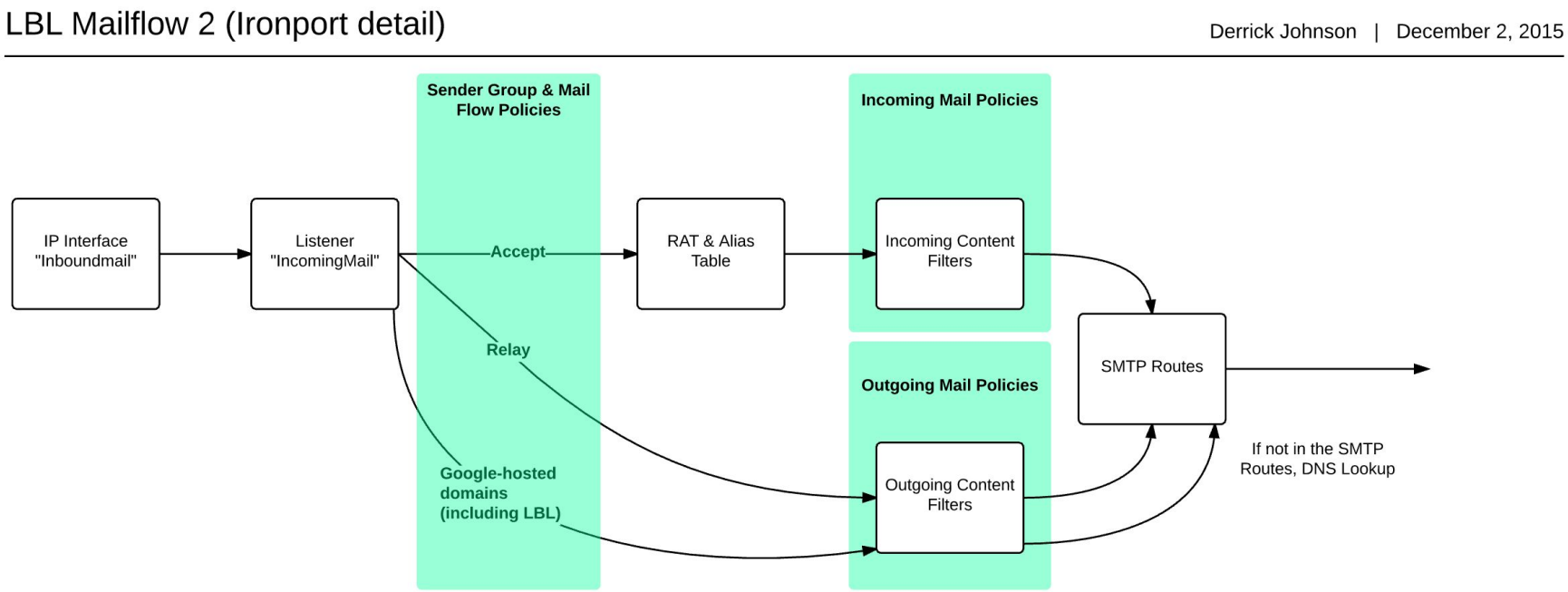


Image Credit: Derrick "The Great" Johnson

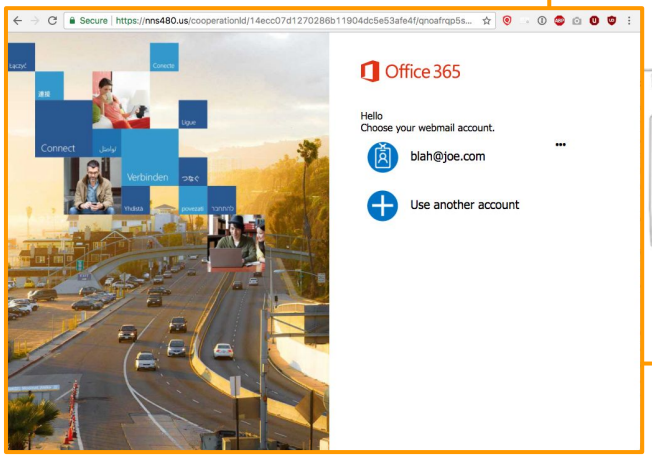
Yes, we do all conventional things

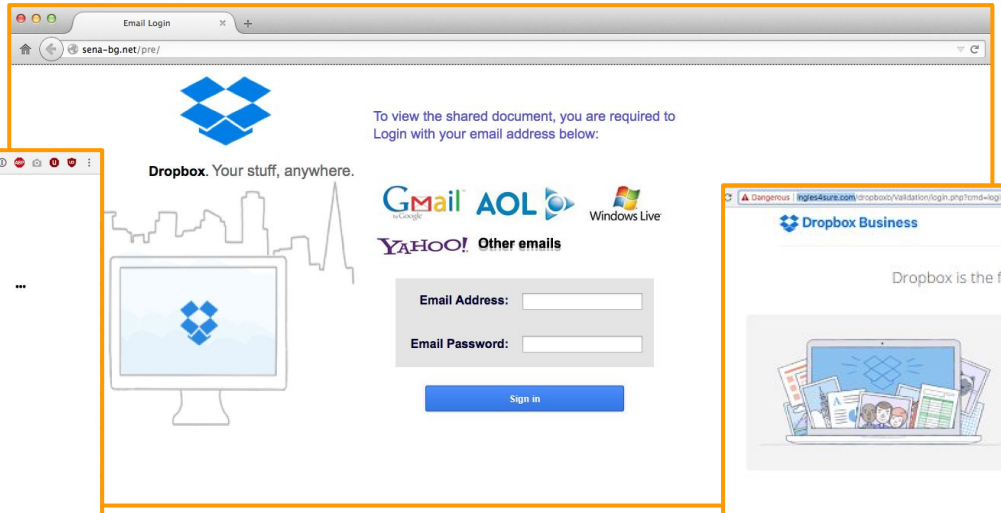
- Gmail
- Ironports
- Phishing specific security training
- Simulated Phishing Exercise
- RPZ
- Other Vendors*

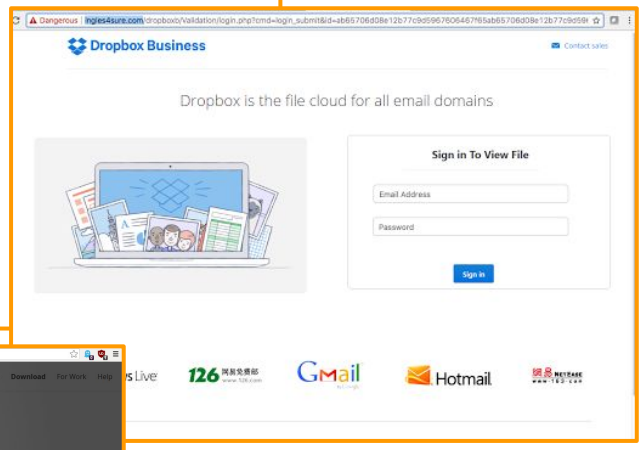
*We learnt that there is not a lot of work in URL analysis as opposed to heavy concentration on attachment analysis.

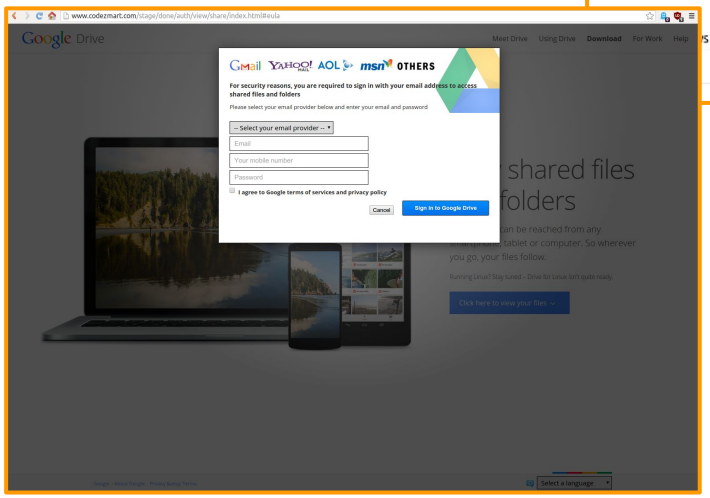
And yet phish makes it way in ….

Source: https://gifrific.com/fish-jumps-out-of-water-and-hits-man/









BERKELEY LAB

Berkeley Lab Login

Not Secure | lbl.gov.freeshare.link/secure/b3c6f3/15d6bb06-f78f-4329-9943-c088f93ebf33/?owner=1

BERKELEY LAB

Log in with your Berkeley Lab Identity

Username

Password

Log in

A service named Google Apps has requested that you login.

Lost or forgot password?

A U.S. Department of Energy National Laboratory
Managed by the University of California

Questions & Comments
Privacy & Security Notice

---

word00000000wcltrh96UQb2W9rsJ9NowN42-LXrnNlj3ES46u4E9pfLOW8p8Vi842Hjwd0ogp-cx3bMFtLXIbrQq28ngWghiKK2vvC

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

A-Z INDEX | PHONE BOOK | JOBS | SEARCH

CENTRAL LOGIN FACILITY

Please login below with your LBNL LDAP username and password:

USERNAME:

PASSWORD:

Login

Berkeley
UNIVERSITY OF CALIFORNIA

Hello,

The University is having a salary increment program again this year with an average of 2.5%.

The Human Resources department evaluated you for a raise on your next paycheck.

Click below to confirm and access your salary revision documents:

Click Here to access the documents

Sincerely,

Human Resources

University of California, Berkeley

---

emporioshop.com.br/ess/lbl-gov/AuthnUserPassword.html

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

A-Z INDEX | PHONE BOOK | JOBS | SEARCH

CENTRAL LOGIN FACILITY

Please login below with your LBNL LDAP username and password:

USERNAME:

PASSWORD:

Login

Are you wondering how you ended up here?

A service offered by Lawrence Berkeley National Laboratory, but without more descriptive details, has requested that we authenticate you. Because your username and password will not be transmitted to this service, you should feel relatively safe logging in here.

If you have concerns about the legitimacy of this site, please contact the Help Desk at x4357 or help@lbl.gov.

A U.S. Department of Energy National Laboratory Operated by the University of California

Office of Science

QUESTIONS & COMMENTS - PRIVACY & SECURITY NOTICE

---

lbl.uni.me/https.login.lbl.gov.html

Reported Web Forgery!

BERKELEY LAB
LAWRENCE BERKELEY NATIONAL LABORATORY

U.S. DEPARTMENT OF ENERGY

A-Z INDEX | PHONE BOOK | JOBS | SEARCH

CENTRAL LOGIN FACILITY

Please login below with your LBNL LDAP username and password:

USERNAME:

PASSWORD:

Login

Are you wondering how you ended up here?

This site serves as the secure login provider for the LBNL's **Google Apps** services.

If you have concerns about the legitimacy of this site, please contact the Help Desk at x4357 or help@lbl.gov.

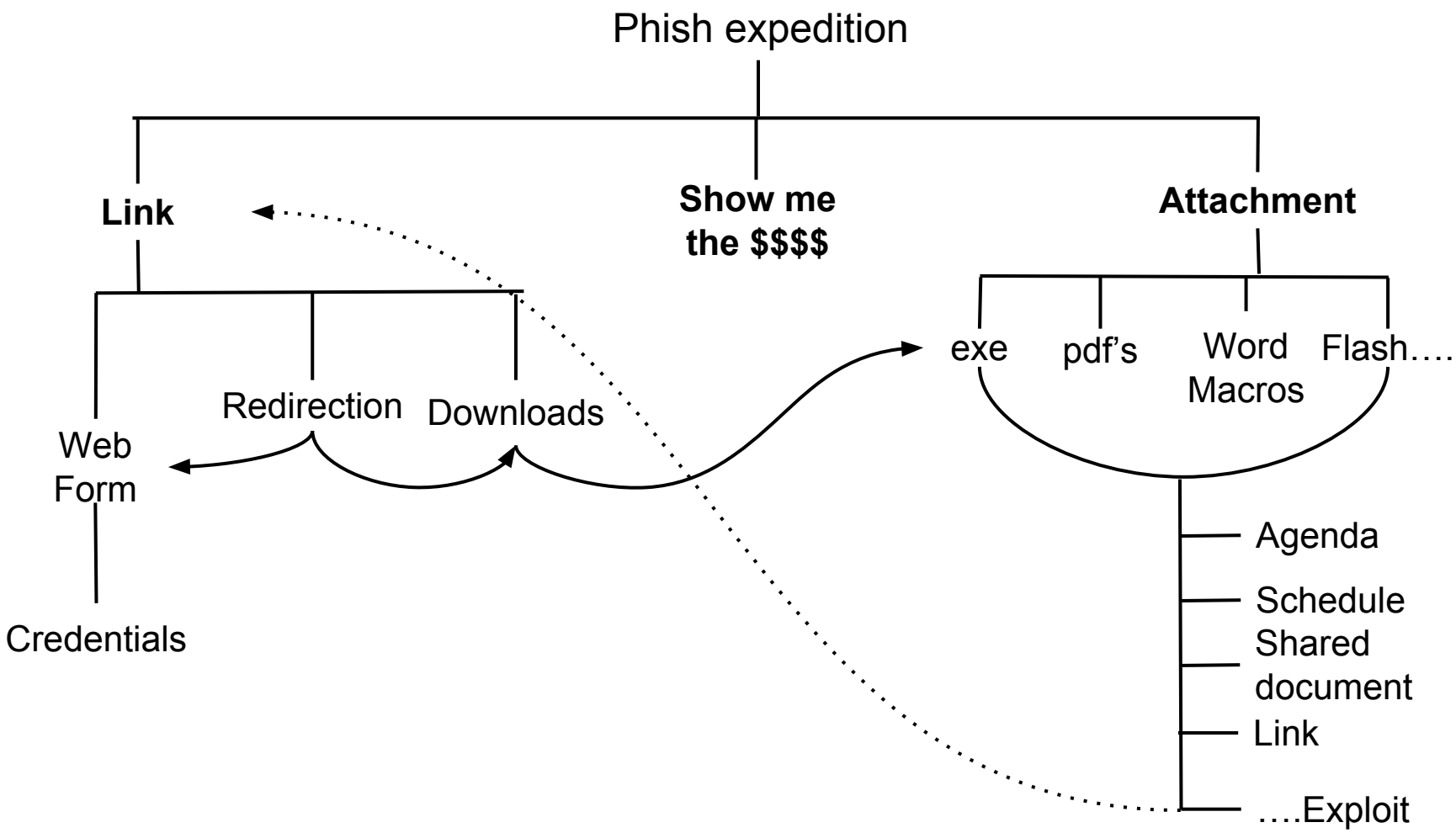
A U.S. Department of Energy National Laboratory Operated by the University of California

Office of Science

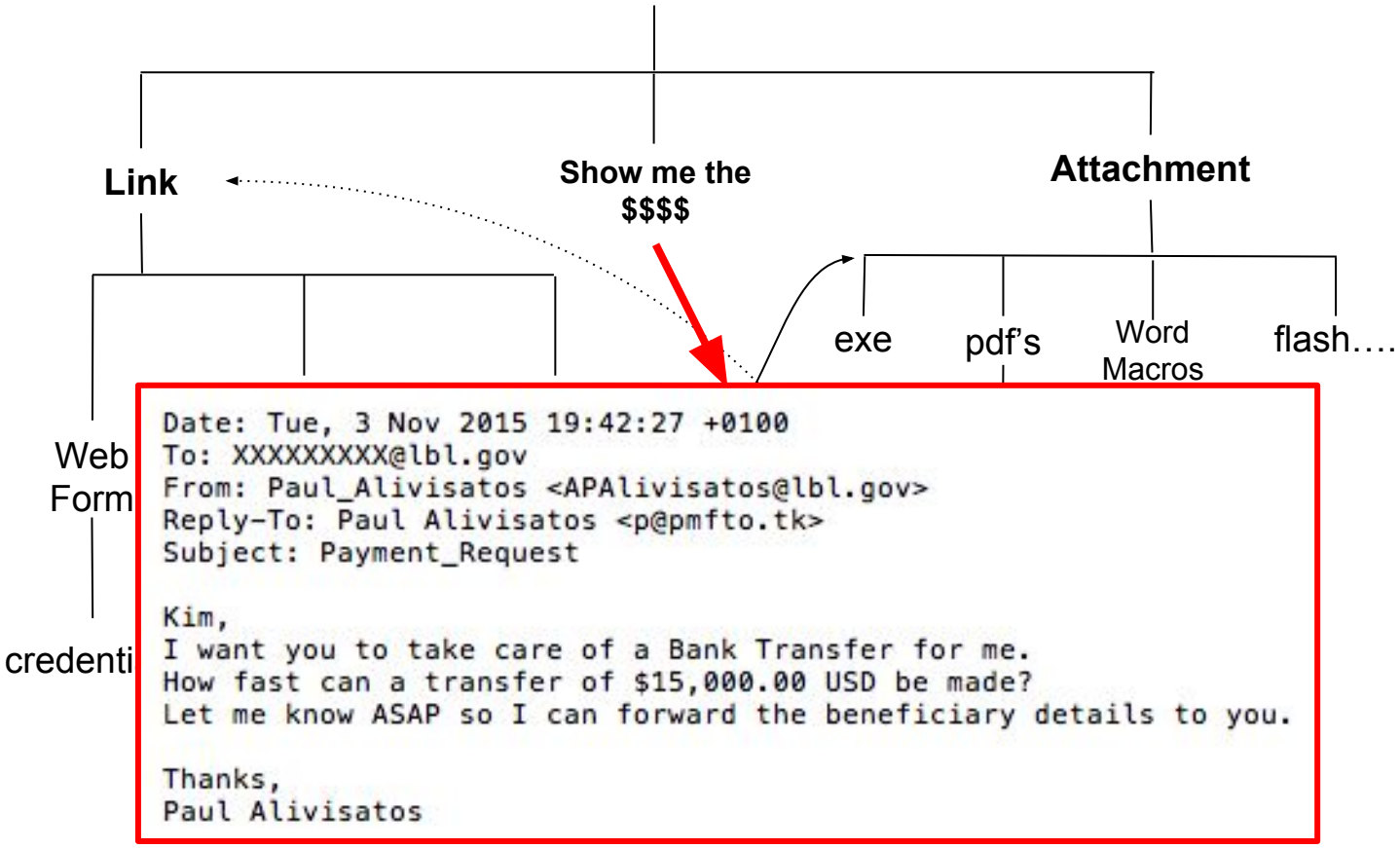
QUESTIONS & COMMENTS - PRIVACY & SECURITY NOTICE

BERKELEY LAB

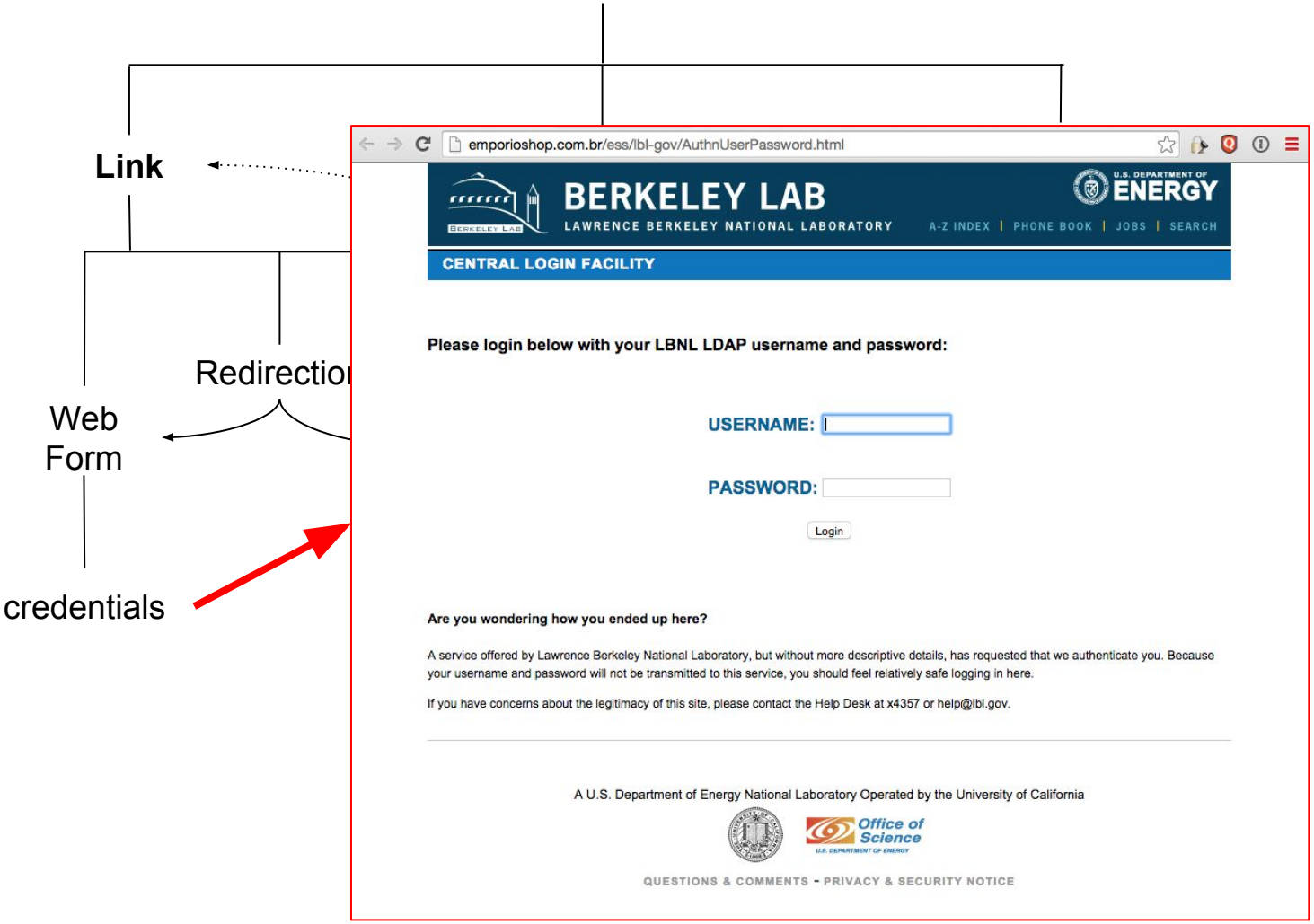
This work is a supplement to the existing technologies we've put in production


Phish expedition
Link
Show me the $$$$
Attachment
Web Form
Redirection
Downloads
Credentials
exe
pdf's
Word Macros
Flash….
Agenda
Schedule
Shared document
Link
….Exploit

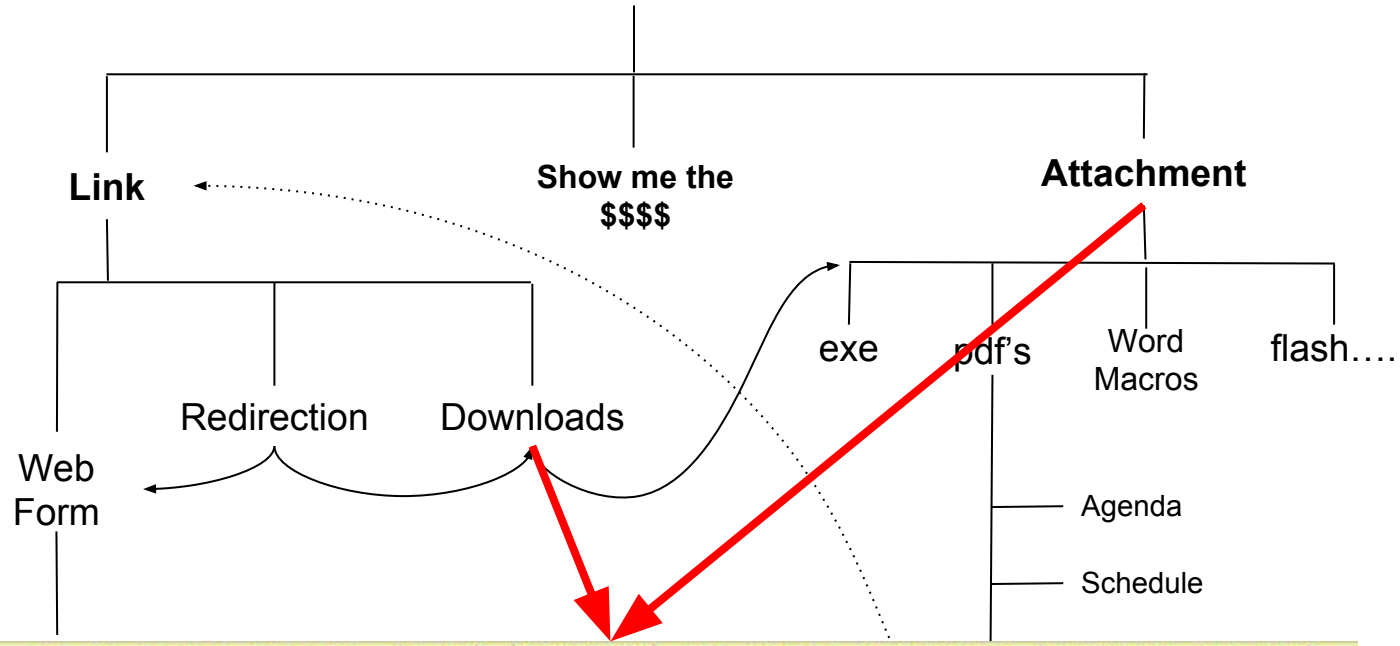
Phish: Exploit Payload



Phish: Exploit Payload



Phish: Exploit Payload




```
42124     Jul 25 Sophos Support  (8.1K) [#7450294] RE: Sample submission (IDENTITY CREATED)
42357     Jul 26 Sophos Support  (8.2K) [#7451388] RE: Sample submission (IDENTITY CREATED)
51186     Aug 16 Sophos Support  (8.2K) [#7503110] RE: Sample submission (IDENTITY CREATED)
51536     Aug 17 Sophos Support  (8.6K) [#7503126] RE: Sample submission (IDENTITY CREATED)
51745     Aug 17 Sophos Support  (8.5K) [#7505760] RE: Sample submission (IDENTITY CREATED)
56996     Aug 24 Sophos Support  ( 10K) [#7519887] RE: Sample submission (IDENTITY CREATED)
56997     Aug 24 Sophos Support  (8.4K) [#7519874] RE: Sample submission (IDENTITY CREATED)
65023     Sep 06 Sophos Support  (  71) [#7548652] RE: Sample submission (IDENTITY CREATED)
```

smtp.log | SMTP transactions

FIELD	TYPE	DESCRIPTION
ts	time	Timestamp when message was first seen
uid & id		Underlying connection info > See conn.log
trans_depth	count	Transaction depth if there are multiple msgs
helo	string	Contents of the HELO header
mailfrom	string	Contents of the MAIL FROM header
rcptto	set	Contents of the RCPT TO header
date	string	Contents of the DATE header
from	string	Contents of the FROM header
to	set	Contents of the TO header
cc	set	Contents of the CC header
reply_to	string	Contents of the ReplyTo header
msg_id	string	Contents of the MsgID header
in_reply_to	string	Contents of the In-Reply-To header
subject	string	Contents of the Subject header
x_originating_ip	addr	Contents of the X-Originating-IP header
first_received	string	Contents of the first Received header
second_received	string	Contents of the second Received header
last_reply	string	Last server to client message
path	vector	Message transmission path, from headers
user_agent	string	Value of the client User-Agent header
tls	bool	Indicates the connection switched to TLS
fuids	vector	File unique IDs seen attached to message
is_webmail[1]	bool	If the message was sent via webmail

[1] *If policy/protocols/smtp/software.bro is loaded*

But data from this SMTP::Info record isn't sufficient anymore

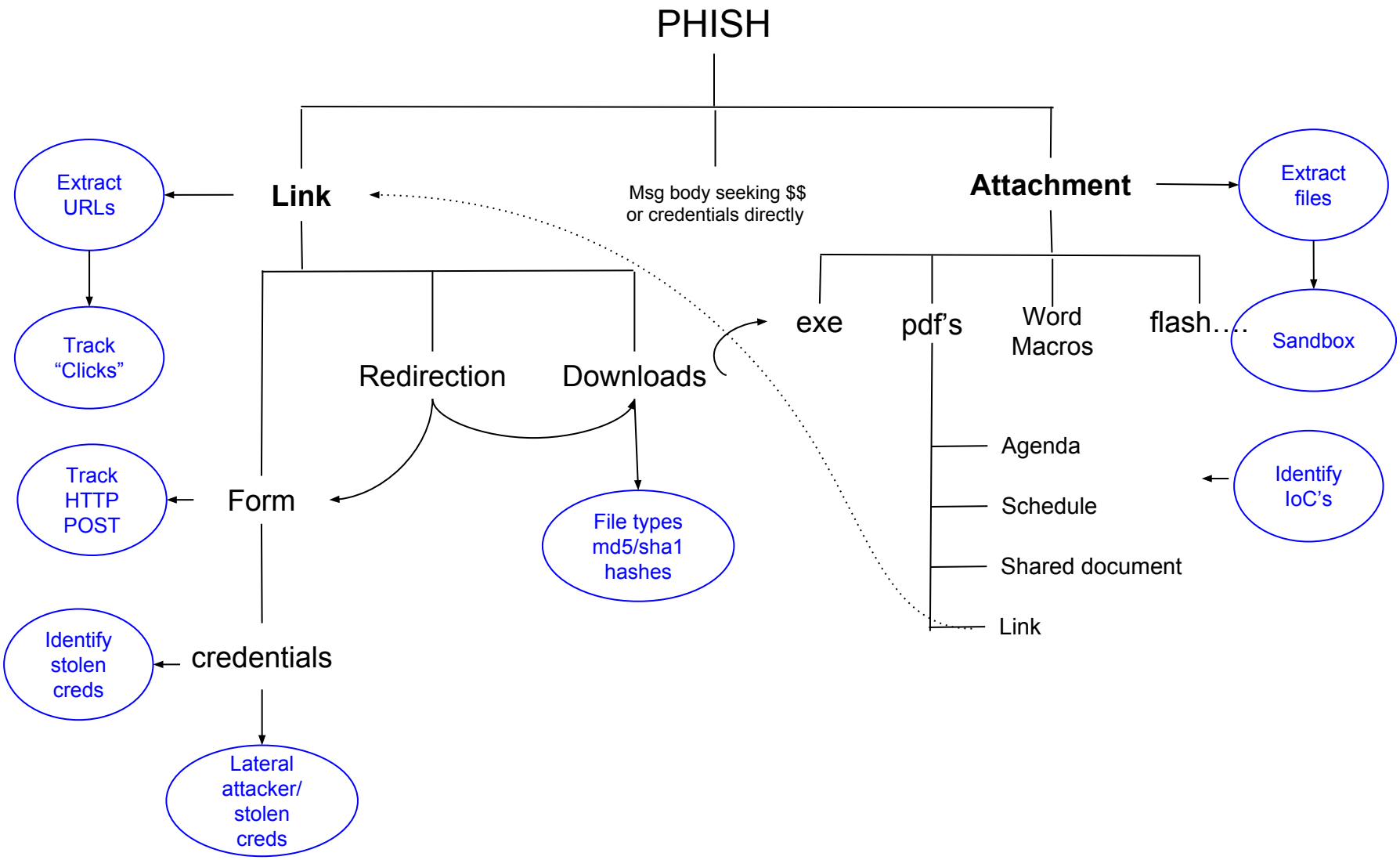
We need more visibility than just timestamp, sender, recipients, subject, dates, path, reply, originating_ip, user_agent etc etc

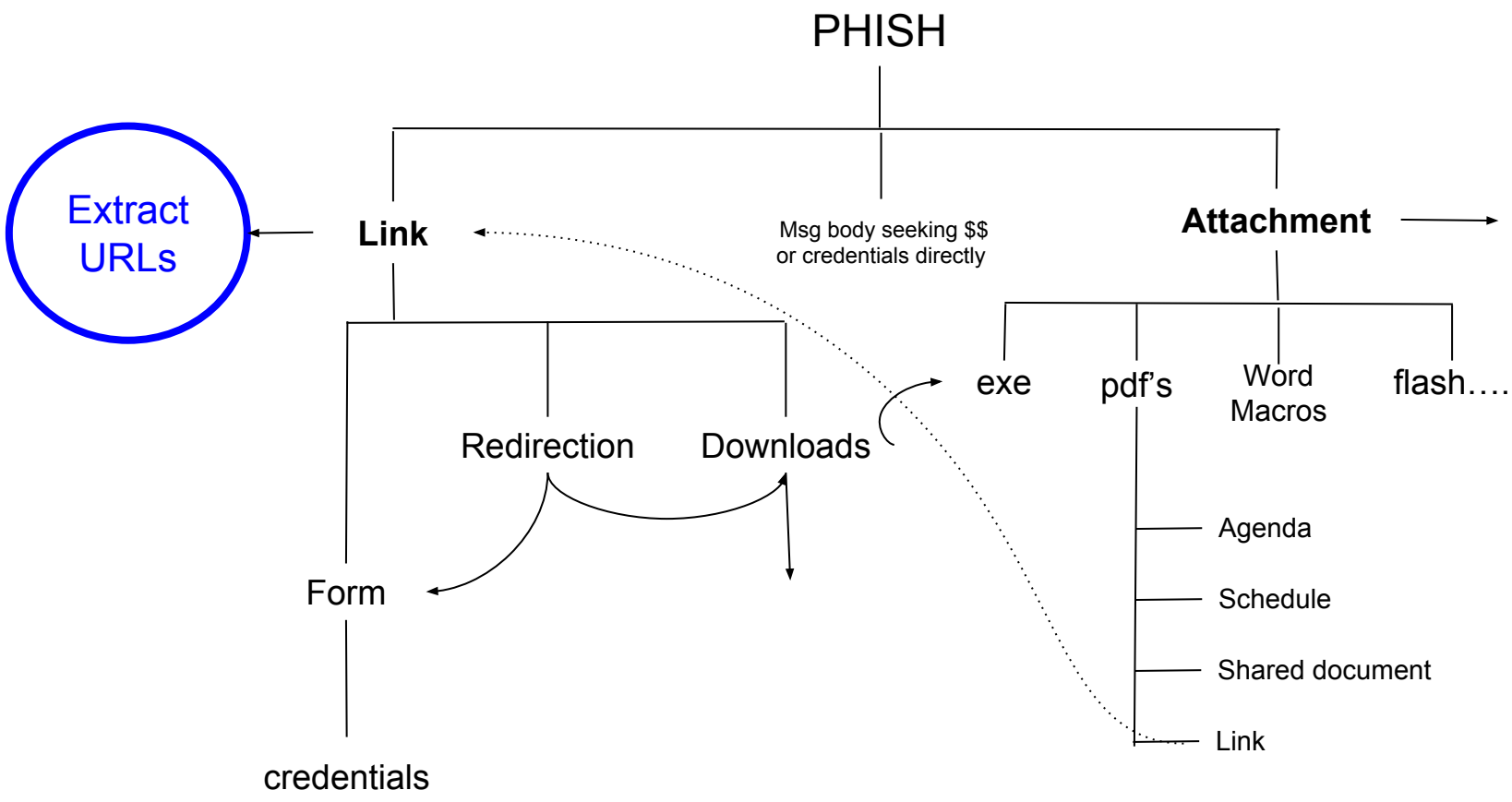
Source: https://github.com/corelight/bro-cheatsheets/blob/master/Corelight-Bro-Cheatsheets-2.5.pdf

BERKELEY LAB

We need more visibility into SMTP

- New log which gives us all the URLs seen in email
- New ability to track
  - which URLs get clicked on
  - signature match on URLs
  - generate alerts based our knowledge from past
- New Alerts to identify if a clicked URL transmitted a
  - File (exe, rar etc)
  - Credentials


PHISH
Link
Msg body seeking $$
or credentials directly
Attachment
Extract
URLs
Track
"Clicks"
Track
HTTP
POST
Identify
stolen
creds
Lateral
attacker/
stolen
creds
Form
credentials
Redirection
Downloads
File types
md5/sha1
hashes
exe
pdf's
Word
Macros
flash….
Agenda
Schedule
Shared document
Link
Extract
files
Sandbox
Identify
IoC's


PHISH
Extract URLs
Link
Msg body seeking $$ or credentials directly
Attachment
Redirection
Downloads
exe
pdf's
Word Macros
flash….
Agenda
Schedule
Shared document
Link
Form
credentials

What does it take to Extract URLs

```
event mime_all_data(c: connection, length: count, data: string)
 &priority=-5
{
    if (! c?$smtp)
                return ;

      local urls = find_all_urls(data) ;

        for (link in urls)
        {
                local url = split_string(link,/ /)[0];
                url = gsub(url,/\]$|\)$/,"");
                event Phish::process_smtp_urls(c, url);
        }

   }
```

```
event Phish::process_smtp_urls(c: connection, url: string)
{
      log_smtp_urls(c, url);
}
```

```
function log_smtp_urls(c:connection, url:string)
{
      local info: Info;

      info$ts = c$smtp$ts;
      info$uid = c$smtp$uid ;
      info$id = c$id ;
      info$url = url;
      info$host = extract_host(url) ;

      Log::write(Phish::Links_LOG, info);
  }
```

New Log: smtpurl_links.log

```
#fields ts      uid     id.orig_h       id.orig_p       id.resp_h       id.resp_p       host    url
#types  time    string  addr    port    addr    port    string  string
Aug 31 12:02:57 CD59BMs7lSWO9XL5e       119.76.101.162  52594   128.3.30.30     25      bayimpex.be     http://bayimpex.be/dropbox.html
Aug 31 12:02:59 CFxh8P37ZmEoWxWjt4      113.160.129.28  29056   128.3.30.30     25      aegelle.com     http://aegelle.com/dropbox.html
Aug 31 12:03:08 CJReqg2Cgof2XUlnm7      109.98.108.83   49885   128.3.30.30     25      eifel-netz.de   http://eifel-netz.de/dropbox.html
Aug 31 12:03:08 CCqP6GmCx1DaC0Fol       180.148.210.162 5957    128.3.30.30     25      busad.com       http://busad.com/dropbox.html
Aug 31 12:03:16 CyRHVE3L86qGKXaqZg      187.126.98.37   52553   128.3.30.30     25      fachwerkhaus.ws http://fachwerkhaus.ws/dropbox.html
Aug 31 12:03:33 C6lnBp4j7OgO16HwD5      122.166.114.144 63353   128.3.30.30     25      avtokhim.ru     http://avtokhim.ru/dropbox.html
Aug 31 12:03:42 CLUmwx4AEmoqzzdhcd      187.37.84.194   62255   128.3.30.30     25      potamitis.gr    http://potamitis.gr/dropbox.html
Aug 31 12:03:43 CyVZGk2OeciHWuVgAk      113.167.126.130 58407   128.3.30.30     25      busad.com       http://busad.com/dropbox.html
Aug 31 12:04:12 CsDFfI2vNOXwzISJU2      182.187.89.116  5685    128.3.30.30     25      avtokhim.ru     http://avtokhim.ru/dropbox.html
Aug 31 12:04:15 Cef4Wp3dJcyvvhmfB       150.107.8.186   35495   128.3.30.30     25      albion-cx22.co.uk       http://albion-cx22.co.uk/dropbox.html
Aug 31 12:04:30 CSY2S13EFmSvGcTJAa      178.149.36.9    54414   128.3.30.30     25      patrickreeves.com       http://patrickreeves.com/dropbox.html
Aug 31 12:05:15 CT7T5eDKEzQUkDsmd       138.204.89.129  57988   128.3.30.30     25      melting-potes.com       http://melting-potes.com/dropbox.html
Aug 31 12:05:22 CDO9uih7h60LoN6y1       5.152.239.178   51524   128.3.30.30     25      tasgetiren.com  http://tasgetiren.com/dropbox.html
Aug 31 12:05:31 CdL7ZY9pwW7Gs3p24       190.97.254.210  63941   128.3.30.30     25      conlin-boats.com        http://conlin-boats.com/dropbox.html
```

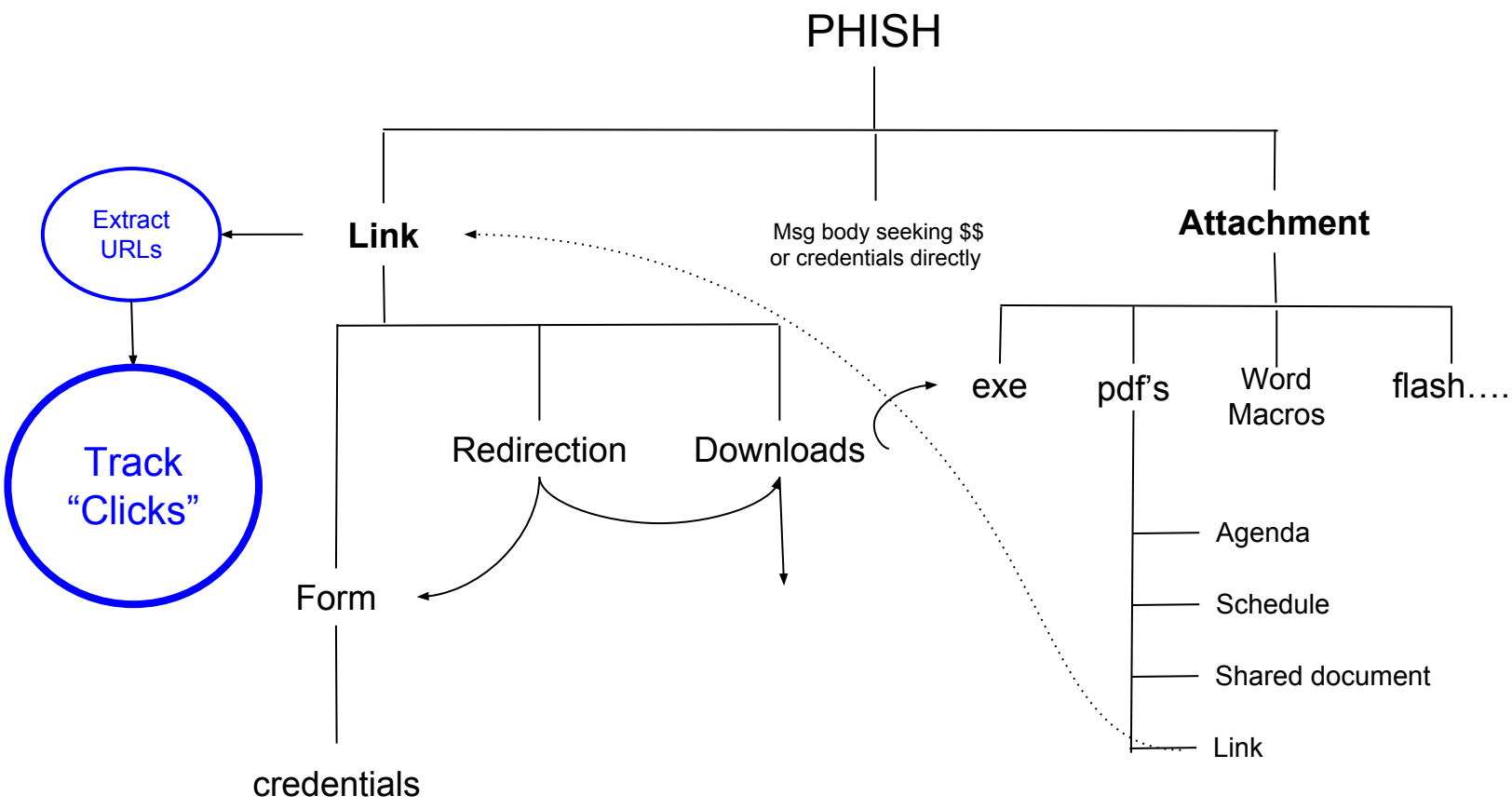
URL Extraction Internals

- Fairly simple to extract URLs from msg body
- Occasional parsing issues
- Correct Regex for URL to be extracted is the only tricky part here

```
const url_regex = /^https?:\/\/([a-z0-9A-Z]+(:[a-zA-Z0-9]+)?@)?[-a-z0-9A-Z\-]+(\.[-a-z0-9A-Z\-]+)*((:[0-9]+)?)(\/[a-zA-Z0-9;:\/\.\-_+%~?&amp;@=#\(\)]*)?/
```

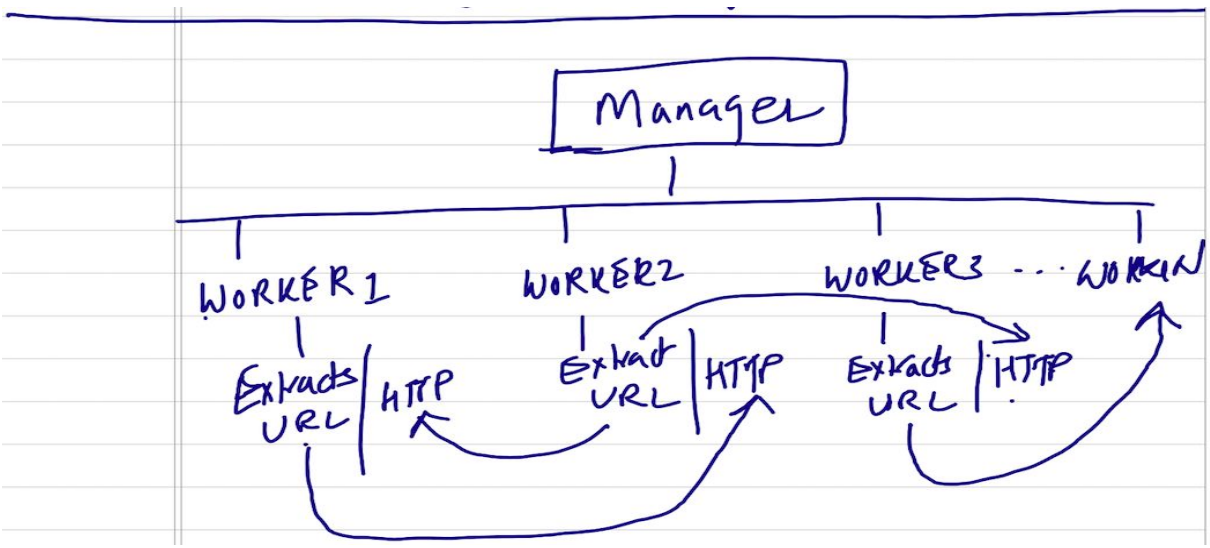
- Bro takes care of logging etc
- Logging framework makes cluster/standalone transparent

PS - Ironports don't do *ALL* URL extractions from Emails



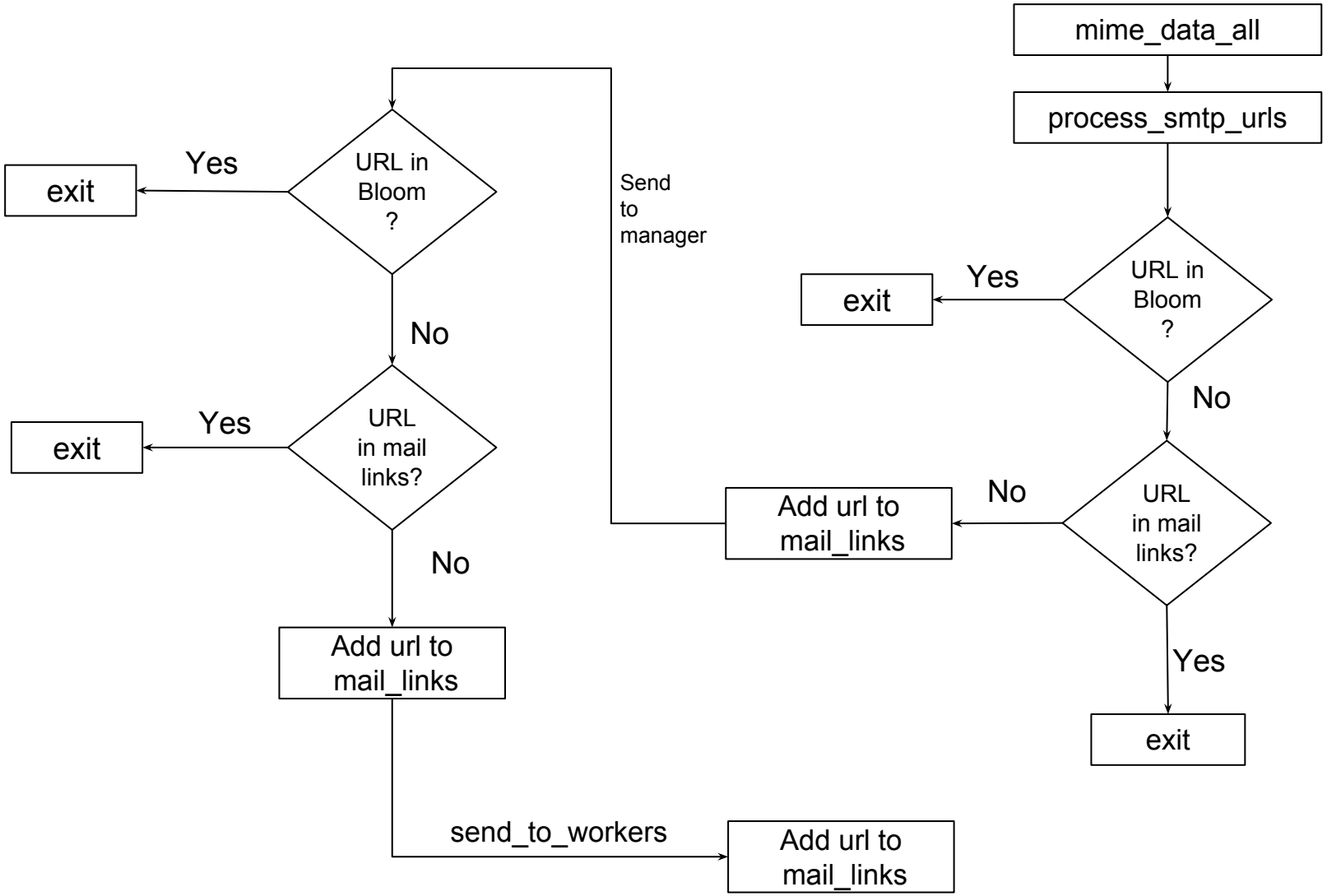
Since we've got SMTP URL's logged won't it be nice to know which ones got clicked-on ? and by who ?

Problem in tracking URLs: Clusterization



- Worker-X processes a SMTP session and extracts a URL
- Worker-Y processes the HTTP GET request for that specific URL
- In short: on a cluster it is mostly unpredictable which worker will process what traffic
- So to track *every* click for *every* extracted URL we need to have
  - All Extracted URLs go to all workers, **or**
  - All HTTP traffic go to all workers, **or**
  - URLs and HTTP traffic go to Manager, **or**
  - Imagine a nice data node which see's all logs

High Level architecture of
distribution of URLs in a Cluster



New log: smtp_clicked_urls.log

Connection Record

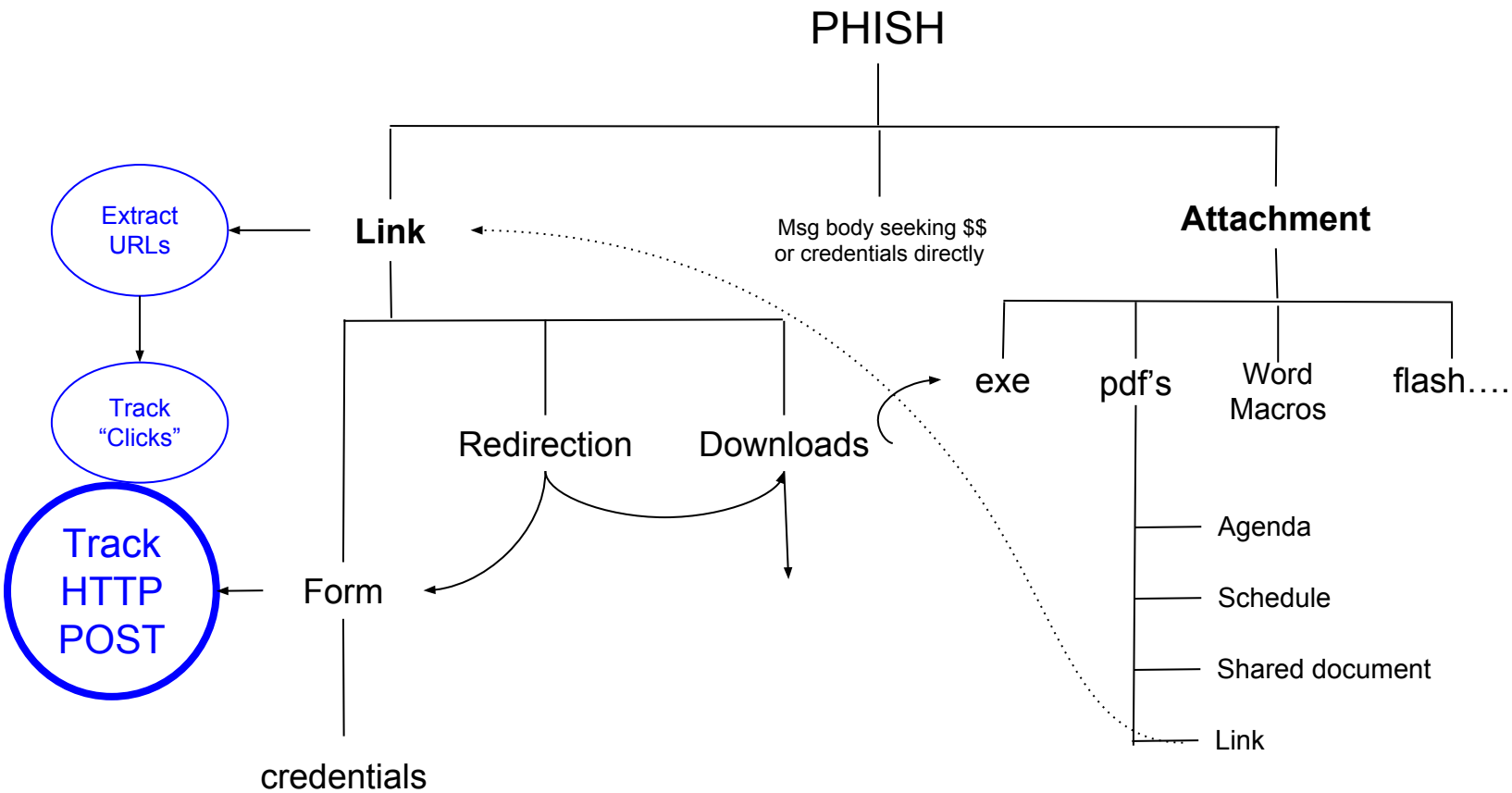
```
1481062180.295358        C3W4S51MSDKicZfirj      128.3.x.y
39017   107.21.6.90      80      lbl.gov.invoicenotices.com

http://lbl.gov.invoicenotices.com/0cb548/?login_id=c25acd74-aed4
-43f3-89a5-563a03a0d9cc
```

URL

```
1481050626.364467        CgP4Rc3LGXkLOhkjWc      Frank Zuidema
<fzuidema@lbl.gov>     XXXXX@lbl.gov  Document review -
Invitation to edit    (empty)
```

First email in which this URL was seen



- Identify passwords transmitted in HTTP POSTs
- Identify FileDownload
- Alert on "SensitiveURIs"
  - Simply signature matching parts of URL string

Tracking HTTP Posts

- Since we can track link clicks, we can identify if any passwords are transmitted over HTTP:

1467998894.642754 CiGsfc4XOymomXJTH8 128.3.X.Y 64310 104.16.58.61 80 - - - tcp
HTTP::HTTPSensitivePOST Request: /electacta/login_action.asp - Data:
username=XXXXXXX@lbl.gov&password=Lopzecz$19&rememberMe=on&role=editor&bypass=&rememberUser=1&ignoreWarning=0 - 128.3.X.Y 104.16.58.61 80 - bro Notice::ACTION_LOG 3600.000000 F - - - - -

If password matches certain complexity Criteria

1467998894.642754 CiGsfc4XOymomXJTH8 128.3.X.Y 64310 104.16.58.61 80 - - - tcp
HTTP::HTTP_Sensitive_Passwd Request: /electacta/login_action.asp - Data:
username=XXXXXXX@lbl.gov&**password=Lopzecz$19**&rememberMe=on&role=editor&bypass=&rememberUser=1&ignoreWarning=0 - 128.3.X.Y 104.16.58.61 80 - bro Notice::ACTION_LOG 3600.000000 F

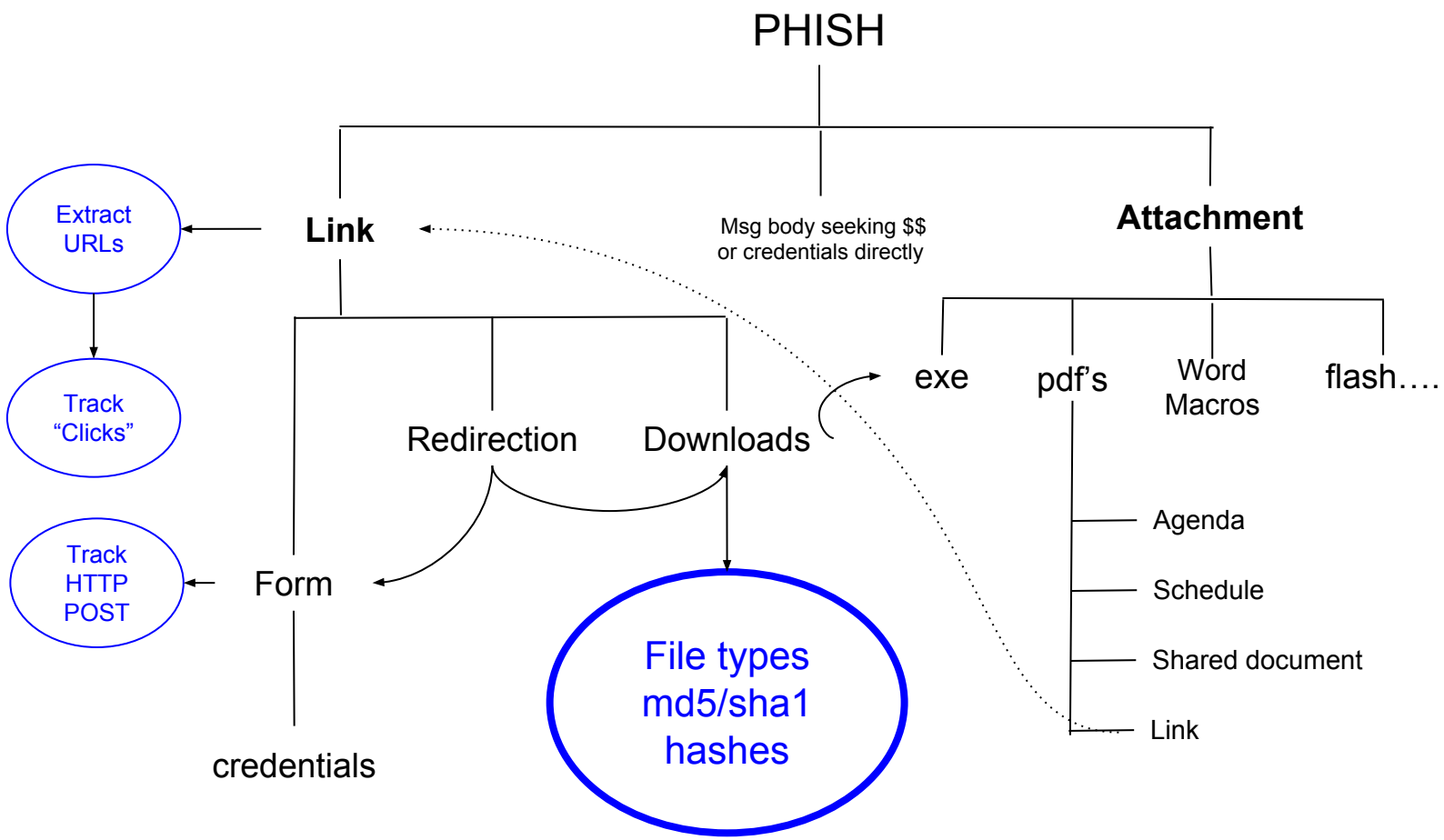
Credit: Jim Mellander

New Alert: SensitiveURI

1351714828.429308 Cu8Nlk1PAJLiEM4Kd9 128.3.41.133 1277 209.139.197.113 25 - - - tcp **Phish::SensitiveURI** Suspicious text embedded in URL

http://avtokhim.ru/dropbox.html from Cu8Nlk1PAJLiEM4Kd9 - 128.3.41.133209.139.197.113 25 - bro Notice::ACTION_LOG 3600.000000 F

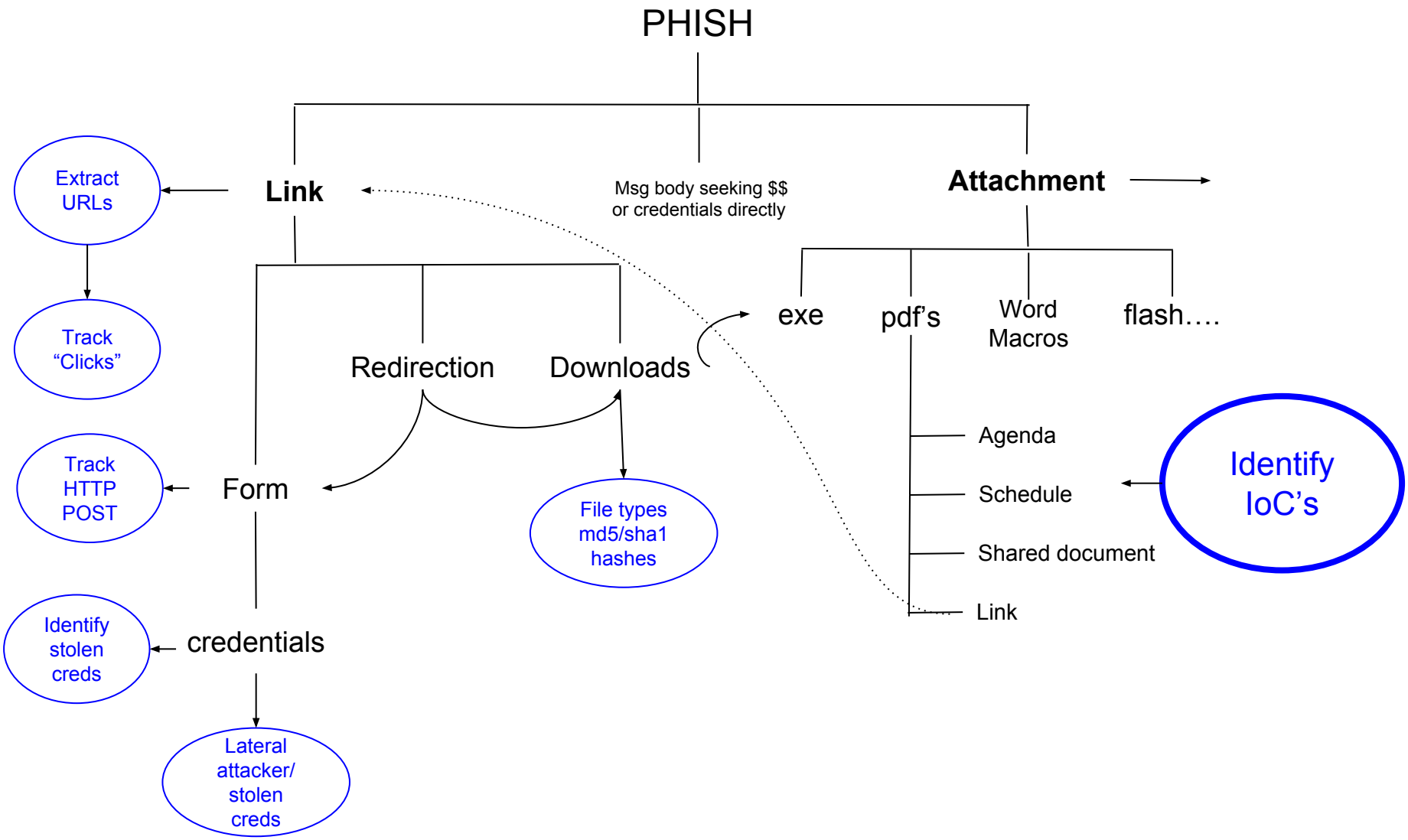
- Signature Match on specific strings within the URLs.
- Mostly useful to flag phishing campaigns built with phishing toolkit
  - /dropbox/dropbox.html


PHISH
Link
Msg body seeking $$
or credentials directly
Attachment
Extract
URLs
Track
"Clicks"
Track
HTTP
POST
Form
credentials
Redirection
Downloads
File types
md5/sha1
hashes
exe
pdf's
Word
Macros
flash….
Agenda
Schedule
Shared document
Link

New Alert: FileDownload

1481499234.568566 C59XU64EvEHN5fr1Th 128.3.x.y 49067 46.43.34.31 80 FxrREO3dgcnSlAQZO8 application/x-dosexec http://the.earth.li/~sgtatham/putty/0.67/x86/putty.exe tcp **Phish::FileDownload** [ts=1481431889.562629, uid=CCCqPL3ZaXmxqdMFJ1, from=cmdline <cmdline@gmail.com>, to=GUI <gui_person@lbl.gov> , subject=putty.exe, referrer=[]] http://the.earth.li/~sgtatham/putty/0.67/x86/putty.exe 128.3.x.y 46.43.34.31 80 - bro Notice::ACTION_LOG 3600.000000 F

We can further Notice on "**Phish::WatchedFileType**", ex: URLs containing .pdf or .rar or .tar.gz or …


PHISH
Link
Msg body seeking $$
or credentials directly
Attachment
Extract
URLs
Track
"Clicks"
Track
HTTP
POST
Identify
stolen
creds
Lateral
attacker/
stolen
creds
Form
credentials
Redirection
Downloads
File types
md5/sha1
hashes
exe
pdf's
Word
Macros
flash….
Agenda
Schedule
Shared document
Link
Identify
IoC's

Identifying Known Known's: Intel feeds

- Malicious Sender
- Malicious Subject
- Malicious Attachment
  - MD5/SHA1
  - Name
  - Mime-type
- Targeted Recipient (ex. Honeypot addresses)
- Malicious reply_to, rcptto,
- Malicious IP origin or in path

New Policy: smtp-malicious-indicators.bro

- Periodic cron to dump all smtp indicators into one flat file
- Bro reads these smtp indicators using input-framework
- Matches against various event attributes
- Generate a notice or an alert

New Alert: Known Malicious Actors

```
1504682044.991930       C4nWFy2vtwAcvOQt8l      107.161.187.234 45086   128.3.41.120    25  tcp
Phish::Malicious_Mailfrom  [indicator=german.mendoza@gpm.com.ve, description=bad-sender],
german.mendoza@gpm.com.ve  german.mendoza@gpm.com.ve     Notice::ACTION_EMAIL,Notice::ACTION_LOG 60.000000       F
```

```
1504682044.991930       C4nWFy2vtwAcvOQt8l      107.161.187.234 45086   128.3.41.120    25  tcp
Phish::Malicious_from   Malicious Sender :: [indicator="Mr. Seigfrid Hernandez"
<german.mendoza@gpm.com.ve>, description=full-bad-sender], "Mr. Seigfrid Hernandez" <german.mendoza@gpm.com.ve>
"Mr. Seigfrid Hernandez" <german.mendoza@gpm.com.ve>    Notice::ACTION_EMAIL,Notice::ACTION_LOG 60.000000       F
```

```
1504682044.991930       C4nWFy2vtwAcvOQt8l      107.161.187.234 45086   128.3.41.120    25  tcp
Phish::Malicious_reply_to      Malicious reply_to:: [indicator=german.mendoza@gpm.com.ve,
description=bad-sender], german.mendoza@gpm.com.ve   german.mendoza@gpm.com.ve
Notice::ACTION_EMAIL,Notice::ACTION_LOG 60.000000       F
```

```
1504682044.991930       C4nWFy2vtwAcvOQt8l      107.161.187.234 45086   128.3.41.120    25  tcp
Phish::Malicious_subject       Malicious Subject:: [indicator=RFQ # 170227 - Atlas Copco Spare Parts -
Jerwia, description=bad-subject], Notice::ACTION_EMAIL,Notice::ACTION_LOG 60.000000       F
```

```
1504682044.991930       C4nWFy2vtwAcvOQt8l      107.161.187.234 45086   128.3.41.120    25  tcp
Phish::Malicious_Path   Blacklisted IP in smtp relay Path: [indicator=185.29.10.121, description=bad-ip ]
185.29.10.121 185.29.10.121   107.161.187.234 128.3.41.120 25   bro     Notice::ACTION_LOG      60.000000       F
```

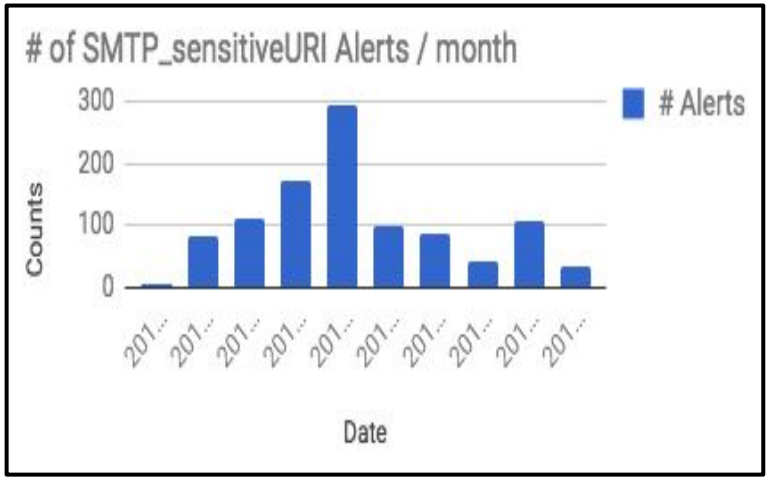
```
1504682044.991930       C4nWFy2vtwAcvOQt8l      107.161.187.234 45086   128.3.41.120    25  tcp
Phish::Malicious_rcptto :: [indicator=XXXXXXX@lbl.gov, description=recipient],  bro
Notice::ACTION_EMAIL,Notice::ACTION_LOG 60.000000       F
```

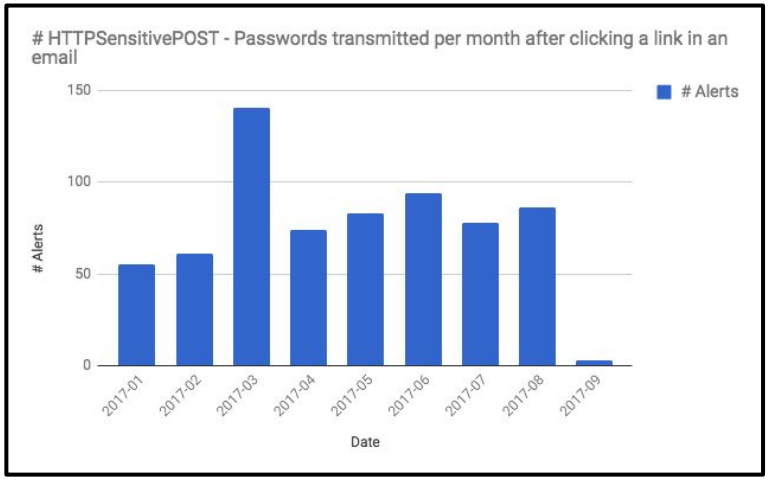
BERKELEY LAB

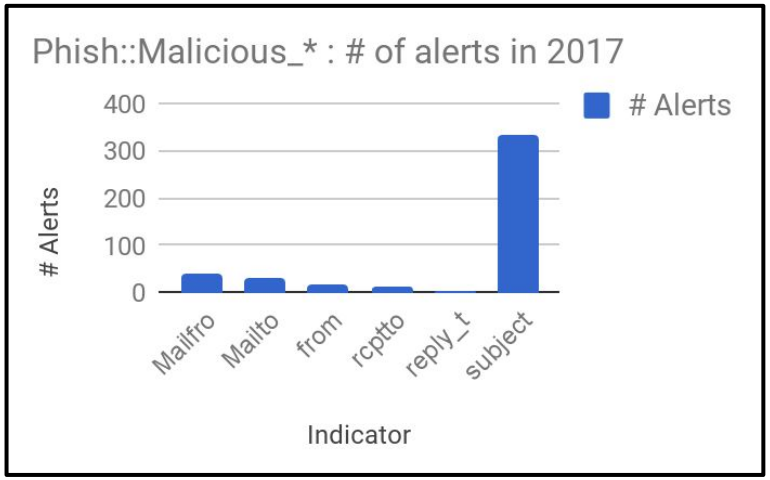
Gaining Visibility: Summary

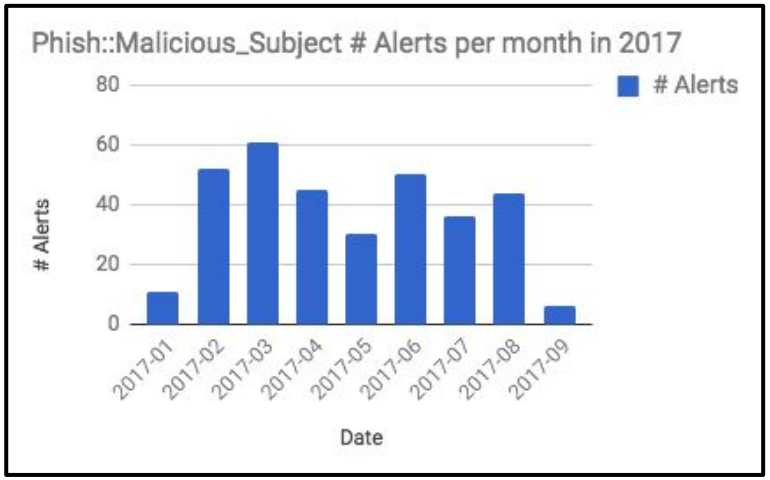
- We've got capability of
  - Identifying URLs from email
  - Signature matching on those URLs
  - Signature matching on smtp record based on intel-feeds
  - Identifying actions as consequence of the URL giving us a solid forensic trail
    - Clicks
    - HTTP POSTS
    - Downloads
    - Estimate on the file types pointed to by URL
- Now let's look at the performance ……

Performance: Number of Alerts Generated









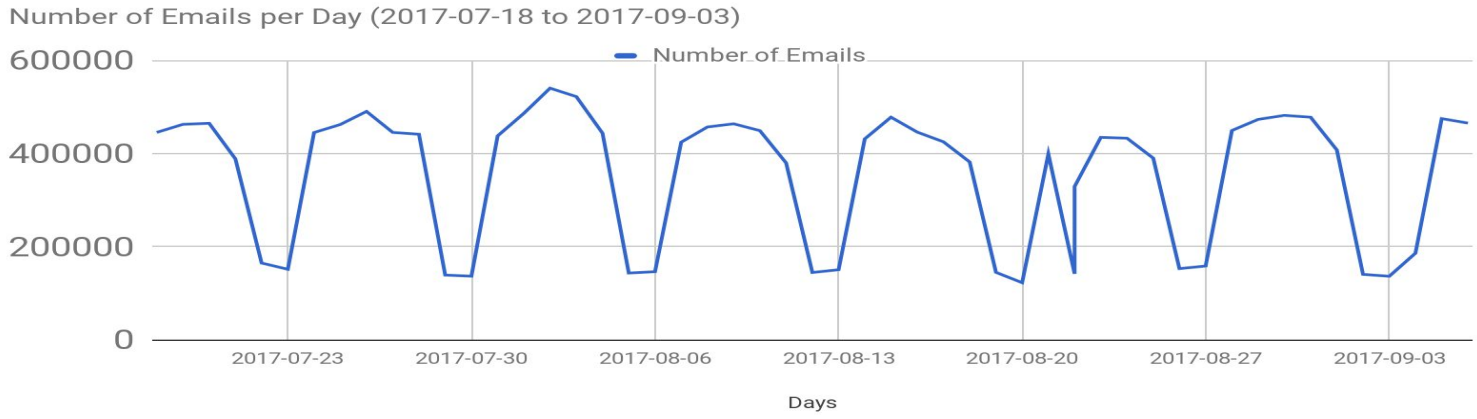
BERKELEY LAB

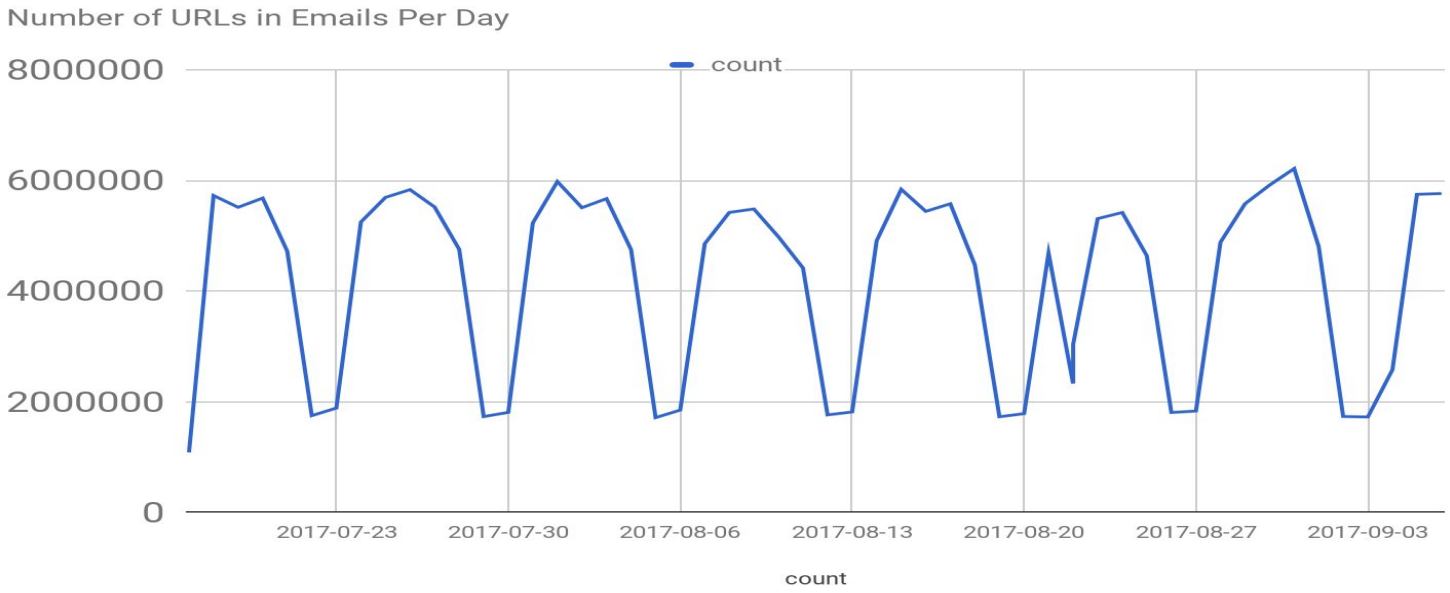
Yet, these are not enough…

Alerts just too damn high (272 on a random day)

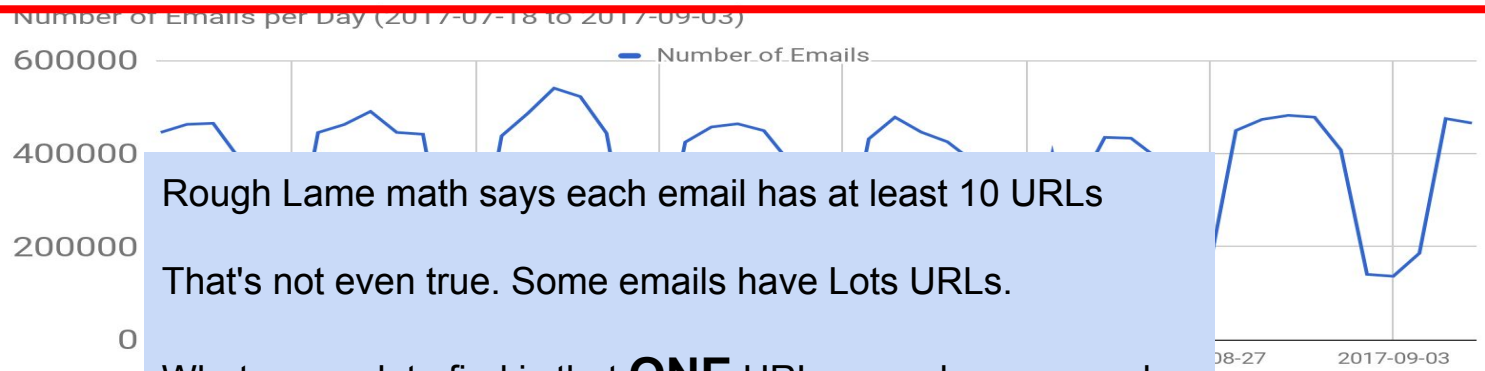
And We still need to catch the phish we gotta catch

Observation: #URLs in Email = 10 x # Emails





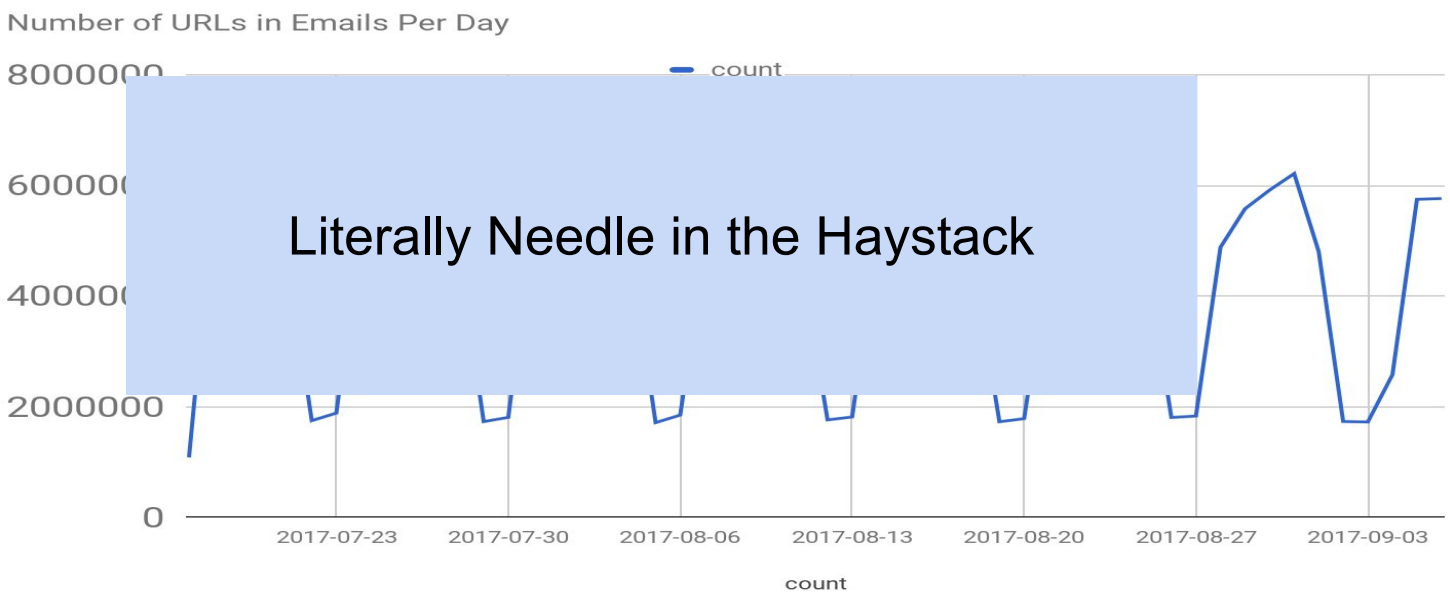
Observation: # URLs in Email = 10x # Emails



Rough Lame math says each email has at least 10 URLs

That's not even true. Some emails have Lots URLs.

What we seek to find is that **ONE** URL ……a day ...every day



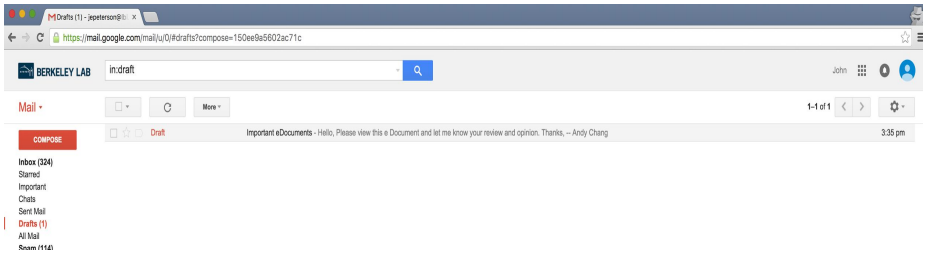
Literally Needle in the Haystack

Time to bring in the Big Guns: Adding smartness into the system

Identifying credentials spearphish

- Current challenges due to base rate issues
  - 500K emails/day @0.1% FP = 500 alerts a day
  - Reality we see is about avg 50-100 smtp-related alerts a day
- Primary Reason is that all the alerting so far is either
  - Lure centric, or
  - Exploit centric

Impersonation Attacks : Spearphish attackers send an email under the identity of a trusted or authoritative entity and include some compelling content in the email to take an action on.

Type of Impersonation	Forge Name	Forge Email	Real Life Example
Address Spoofer	May or may not	YES	Date: Fri, 14 Aug 2015 12:04:00 -0500 (CDT) From: "A Alivisatos" <aalivisatos@lbl.gov> To: XXXXXX@lbl.gov Subject: Good Morning ← Note Reply-To: aalivisatoslbl@mail.com Send me the balance on all our accounts as of today's date. Thanks
HistoricallyNewAttacker	Unseen Name	Unseen Email (@lbl.gov = forged)	Date: Tue, 08 Nov 2016 17:38:28 +0000 From: Computer Maintenance <compmaint@lbl.gov> To: afXXXXXXXh@lbl.gov Subject: Urgent: Email reactivation
NameSpoofer	Yes	Yes	From: Steven Chu <david@huismanauction.com> Date: January 9, 2017 at 11:22:42 PM PST To: undisclosed-recipients:; Subject: Steven Chu shared a File with you
Lateral Attacker (Stolen Credentials)	No need to spoof - given account ownership	No need to spoof - given account ownership	

Lateral Attacker

Important eDocuments

lovelysunshine0923@hotmail.com, Bcc: Lin@aol.com, ptjliu@hotmail.com 398 Bcc

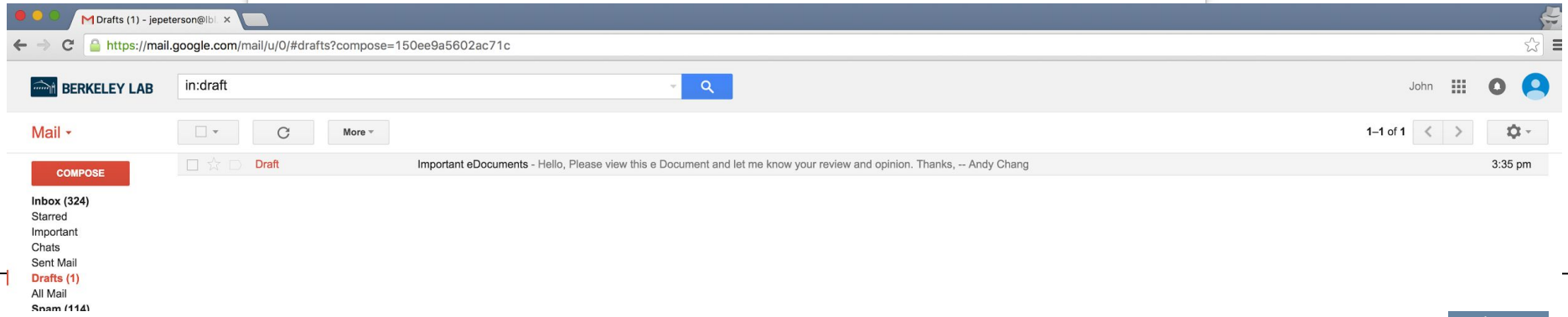
Important eDocuments

Hello,

Please view this **e Document** and let me know your review and opinion.

Thanks,

--
Andy Chang
President/CEO
Continental Assets Management
130 W.Huntington Dr
Arcadia, Ca.91007
626-358-9888 Fax 626-358-9800

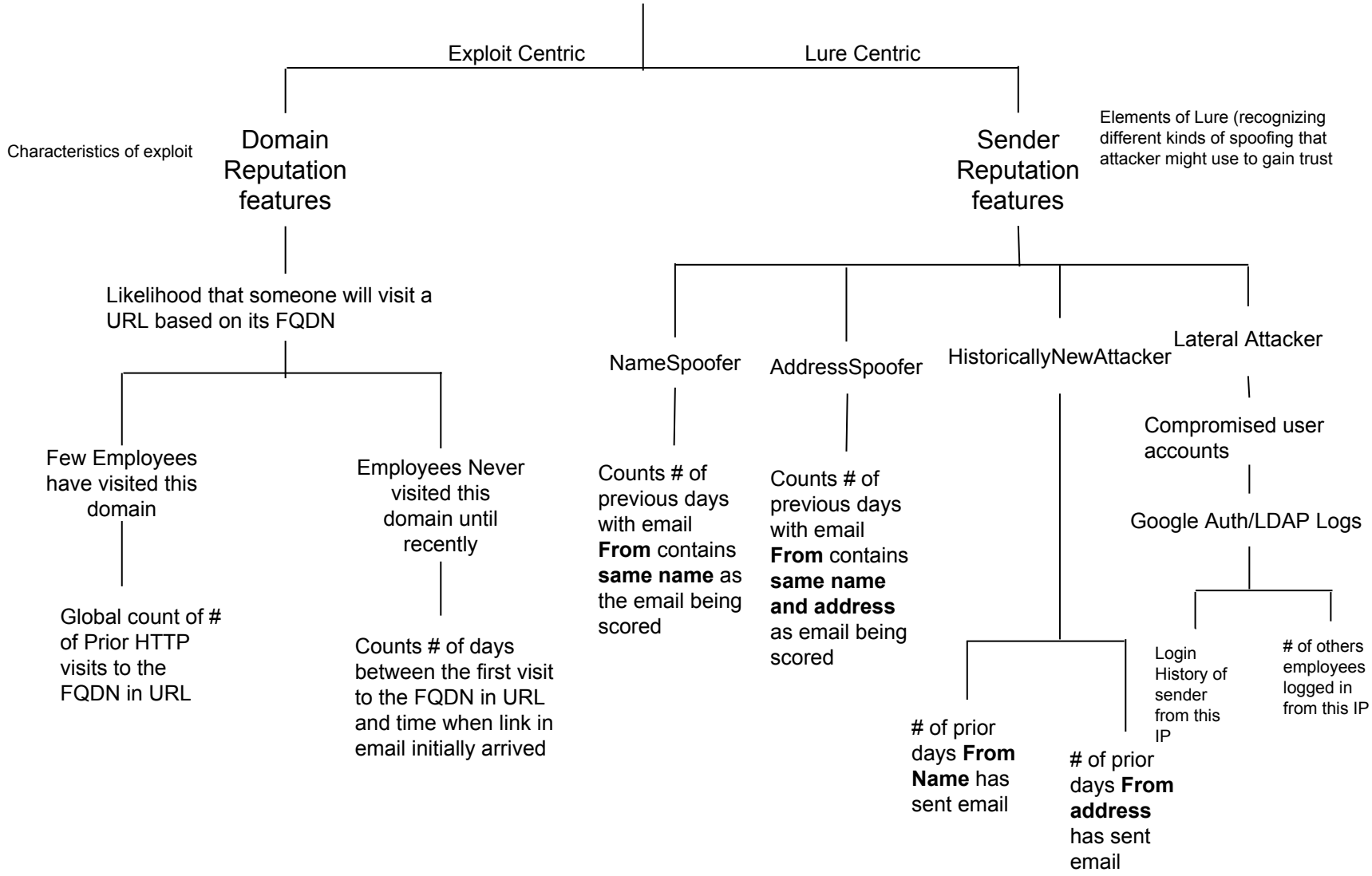


Important eDocuments

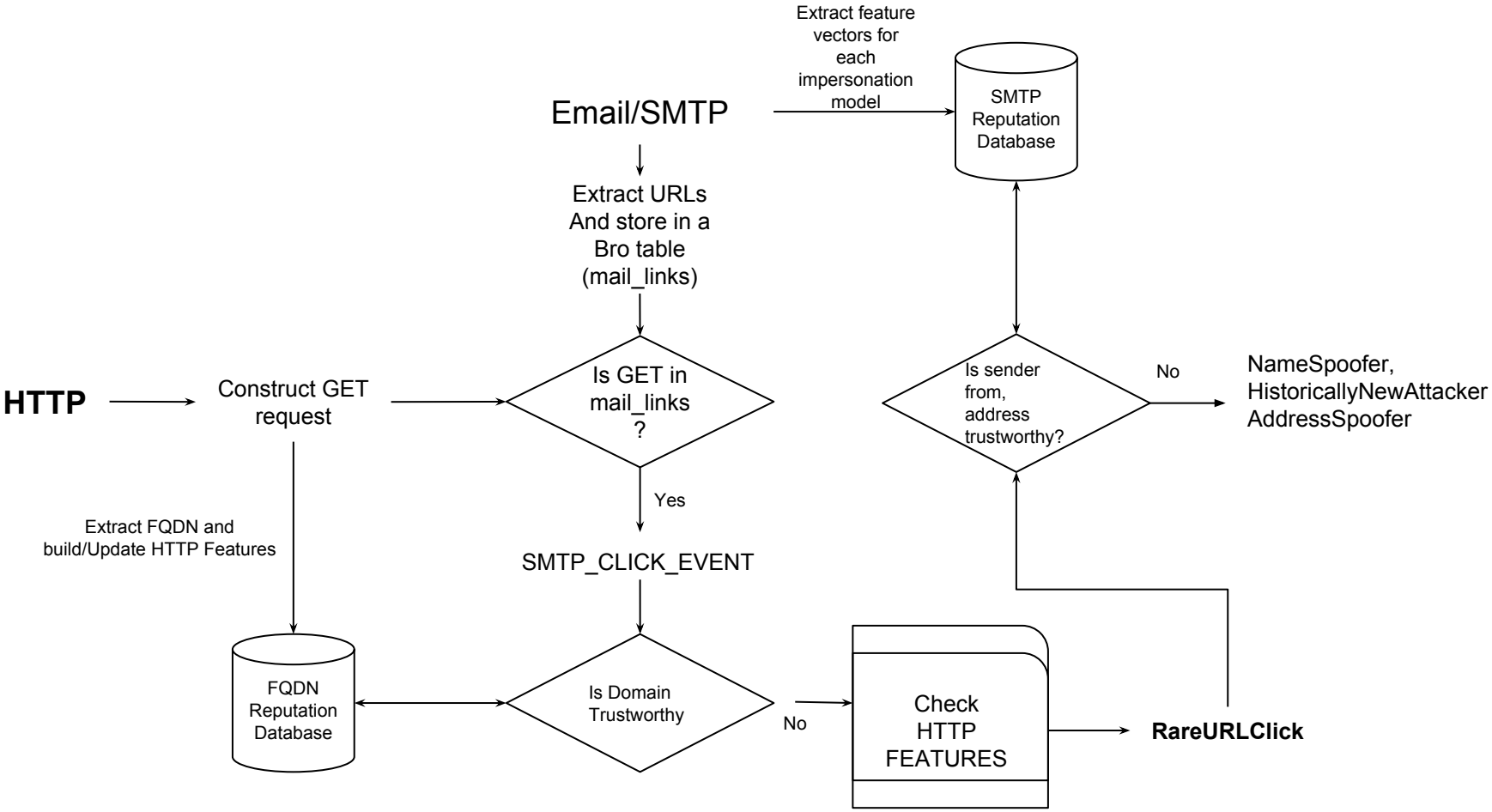
To lovelysunshine0923@hotmail.com

Bcc Lin@aol.com ptjliu@hotmail.com ptrendacosta@frandzel.com ptseng@tpg.com.au puakailima77@yahoo.com puikihung@gmail.com punditcz@yahoo.com puwiduuxihlo@yahoo.com Pvspinosa@aol.com pzia@yahoo.com qmartyn@maui.net q@bristolyachtinc.com q@bristolyachtsinc.com q5jgf-2988665077@hous.craigslist.org qppkm-2992226341@hous.craigslist.org supnova@hotmail.com queenie@cadailyfunding.com queenie@cafunding.com queeniecjw@hotmail.com queeniejw@hotmail.com socalqm@yahoo.com quickpinnacle@yahoo.com.cn quidilee@gmail.com qunyu@yahoo.com smucker007@gmail.com rkim1078@hotmail.com r.cadiz@ymail.com r.lin@takisun.com rachelk@jadeescrow.com Rachel.Cubas@cbre.com Rachel.Lin@situs.com rmoore@stantonarchitecture.com rachelngo1967@yahoo.com Rachel.Lin@situscompanies.com radiantchase@yahoo.com rerobles@msn.com RAFIKP@glendalenisinf.com rafsip@aol.com rahirst@gmail.com raiji.hanashima@hsbc.co.jp raj_raman99@yahoo.com rajiv.trivedi@laquinta.com chrimprovements@gmail.com ralphdivino@aol.com rkannan@doheny.org RamasarA@dwaf.gov.za ramiaht21@gmail.com mortpri@cox.net rappeldorn@opusbank.com RKirby@ci.arcadia.ca.us randym@1stvalley.com randy@indianridgecc.com rim@warrenmarcus.com ranty.h.liang@jpl.nasa.gov rantyliang@aol.com Ranty.H.Liang@jpl.nasa.gov saroya12@gmail.com Ranya.Ku@eastwestinsurance.com rapittet@usfca.edu sobalvarro_r@med.usc.edu belliappa@gmail.com rwilcox@aranewmark.com rkhama04@yahoo.com rkhamaguchi@gmail.com raymcwong@yahoo.com ray@socalunits.com raychao@pacbell.net ray@brei.com Rbaez@crescentheights.com rcamire@jbaia.com rcanalez@hpapts.com rcarrera@lasc.co.la.ca.us RCH9876@aol.com rchaikin@retailadvisorygroup.com rdantas7@hotmail.com re@ssinter.com realesta@islandprodesign.com realestate@debbiehanna.com realestate@jasonhoopai.com realestate@pilottravelcenters.com rebecca16331@cox.net Rebecca_lee@cathaybank.com reed_smileycpa@bookkeepinghelp.com reegsimpex@gmail.com reese@loanmarket.net Regchua07@hotmail.com reginac329@gmail.com reichenbaum@sciproperties.com reinaldo@bennpacific.com ringmotion@yahoo.com dhuang@renderholic.com renedalaten@yahoo.com thirtylove19@gmail.com Renee.Williams@ffslaw.com rrettally@opusbank.com Renirose@aol.com RENOAPT4SALE@aol.com gacx@earthlink.net rentals@discounthawaiicarrental.com reo_database@bankofamerica.com reoportfolios@yahoo.com reowalt@earthlink.net reply_bjfemg_jnwtqhi@cp20.com -402483_HTML-82165558-44478@email.sciprope -402483_HTML-82165558-44478@email.scipropertie… reservation@thrifty.com reservation2@998.com reservations@twobunchpalms.com resnicka@usc.edu ReUnion.Alert@max.fairopen.net toothgap9@yahoo.com rtacsuan@yahoo.com reyscar@gmail.com Reza_Etedali@mail.vresp.com Reza.Ghaffari@marcusmillichap.com rfaust@firstam.com rflebbe@hawaii.rr.com RFlurry@MARCUSMILLICHAP.com RFochtman@bernards.com rgsummers@keyconstruction.com rhastings@hcbahawaii.com rhc@lava.net Rbanuelos@olivermcmillan.com

Detector Design: Features per attack stage



Detector Design



Feature Vectors and comparators per sub detectors

Feature	Description
isRareURLClick	# If a domain has been seen fewer than 3 times in previous HTTP clicks, it is rare. # If a domain has been seen at least 3 times in prior HTTP traffic, and the time of the 3rd visit was more than 3 days ago, it is rare. # Otherwise, it is not-rare.
isHistoricallyNewAttacker	if (RareURLClick && *from_name:days_sent <= 2* && *from_email_addr:days_sent <= 2*)
isSpoofworthyFromName	SpoofworthyFromName is a boolean OR-clause where: (from_name:days_sent >= 14 || from_name:num_clicks > 1 || from_name:emails_recv > 1)
isNameSpoofer	- if (RareURLClick && *SpoofworthyFromName* && *full_from_field:days_sent <= 1*)

Final version gets rids of these parameters which is a really nice property

Detector Design: Features per attack stage

Exploit Centric

Lure Centric

Characteristics of exploit

Domain Reputation features

Likelihood that someone will visit a URL based on its FQDN

Few Employees have visited this domain

Global count of # of Prior HTTP visits to the FQDN in URL

Employees Never visited this domain until recently

Counts # of days between the first visit to the FQDN in URL and time when link in email initially arrived

Elements of Lure (recognizing different kinds of spoofing that attacker might use to gain trust

Sender Reputation features

NameSpoofer

Counts # of previous days with email **From** contains **same name** as the email being scored

AddressSpoof

Attacker

Lateral Attacker

promised user

/LDAP Logs

# of others employees logged in from this IP

From address has sent email

Replaced with Directed Anomaly Scoring (DAS)

TO be Replaced by DAS: Directed Anomaly Score

- Requires no training data
- Operates in non-parametric fashion
- Order of magnitude better performance than standard anomaly detection heuristics

~~# Alerts just too high 272/day~~

Alerts <= 10 and make sure those are fast to deal with

(Good read - go.lbl.gov/credphish)

Converting theory into bro scripts

- Domain and sender reputation features demand persistence (~14 days of buildup)
- Optimize URL storage in BRO given we've got ~600,000 URLs per day
- In short this all means bro cannot crash (..... or restart :)

Bro+Postgres

```
# bro -N Johanna::PostgreSQL
Johanna::PostgreSQL - PostgreSQL log writer and input reader (dynamic, version 0.1)
```

Available as Bro-Package and at: https://github.com/0xxon/bro-postgresql

Postgres Plugin Automatically creates tables and scheme

Smart to translate Bro's native data types into Postgres and back (both r/w)

Bro records <-> Postgres tables

Bro Script

```
type fqdn_rec : record {
     domain: string ;
     days_visited: vector of time ;
     num_requests: count &default=0;
      last_visited: time ;
      trustworthy: bool &default=F;
      } &log ;

global http_fqdn: table[string] of
fqdn_rec &write_expire=10 days;
```

Postgres DB

```
=> \d http_fqdn
                                    Table "public.http_fqdn"
    Column     |        Type        |                        Modifiers
---------------+--------------------+----------------------------------
----------------------
 id            | integer            | not null default
 domain        | text               |
 days_visited  | double precision[] |
 num_requests  | integer            |
 last_visited  | double precision   |
 trustworthy   | boolean            |

Indexes:
    "http_fqdn_id_key" UNIQUE CONSTRAINT, btree (id)
    "domain_idx" btree (domain)
```

```
lbl.gov.invoicenotices.com  - [days_visited=[1481051156.986024, 1481062180.295358],
num_requests=48, last_visited=1481062276.631609, interesting=T]
```

Design Decisions

- Size of mail_links table and ability to track URLs over days
  - Convert to bloomfilter
  - Problem: Loose the mail_info relation
  - Solution: Fetch mail_info from postgres store

- Should we track every URL or be selective
  - .gif, .jpeg, .png

Problem: Postgres storage works a bit too good

Having more data in storage tables than you need - 25M URLs in 43 days

- Previously we'd struggle to store URLs for 4hrs or 12 hrs in a table
- Limit it to 30 days

Design decisions for fqdn reputation

Optimization	Problem
If we see fqdn_domain in mail_links update the http_fqdn by reading database, unless it's already in table	- Way too many DB queries - We may not even have a 'click'
Read everything from fqdn database and fill up the table	- Too much data - Not quite useful to keep everything in table
Create a trusted_db bloom and untrusted fqdn's go into a table Expire untrusted fqdn's after N days	- Graduating untrusted to trusted syncs - What is that 'N'

Jumpstarting reputation code

https://github.com/initconf/reputation-db-scripts-for-phish-analysis

- When starting to run code from scratch we need to have a reputation database built
- Bro takes at last 2 weeks to build it
  - Until then HUGE number of false positives which reduce per day
- So, we've got two Python scripts which read historic logs and populate reputation database

Challenges

- Postgres database design
  - Non-normalized data
  - Only INSERT and no UPSERT
    - Using adhoc workarounds, for now
    - Delete all but last inserted record
- Operational Problems
  - Cannot get it running on 50 worker 5 box cluster
    - 400GB process size
  - Works perfectly awesome on 20 worker 1 cluster-in-a-box

How is it working for LBNL

- Code is mostly working stable
- Credential spearphish
  - Implementation in Intermediary stage of research paper.
  - Running semi-production state
  - **Need to incorporate Directed Anomaly Scoring (DAS)**

Promising Results

FY17 Performance Review Problem

Inbox x ASHARMA x

Ingrid Peters <ipeters@lbl.gov>

to me

12:41 PM (18 hours ago)

Image ers@lbl.gov

Reviev

from: **Ingrid Peters** <ipeters@lbl.gov>

to: asharma@lbl.gov

date: Mon, Sep 11, 2017 at 12:41 PM

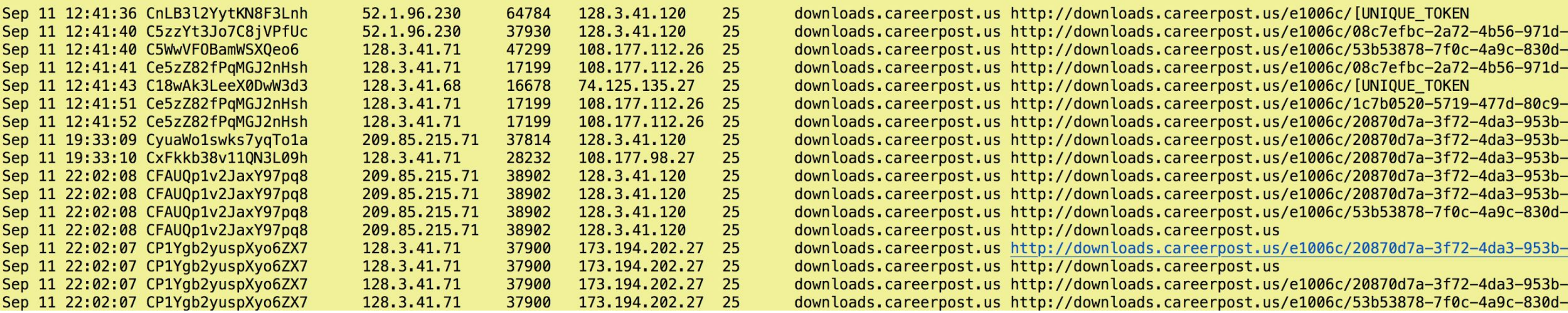
subject: FY17 Performance Review Problem

: Important mainly because of the people in the conversation.

pdf FY17 Performance Review.pdf

attachment

So we’ve got URLs



```
Sep 11 12:41:36 CnLB3l2YytKN8F3Lnh      52.1.96.230    64784   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/[UNIQUE_TOKEN
Sep 11 12:41:40 C5zzYt3Jo7C8jVPfUc      52.1.96.230    37930   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/08c7efbc-2a72-4b56-971d-
Sep 11 12:41:40 C5WwVFOBamWSXQeo6       128.3.41.71    47299   108.177.112.26  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/53b53878-7f0c-4a9c-830d-
Sep 11 12:41:41 Ce5zZ82fPqMGJ2nHsh      128.3.41.71    17199   108.177.112.26  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/08c7efbc-2a72-4b56-971d-
Sep 11 12:41:43 C18wAk3LeeX0DwW3d3      128.3.41.68    16678   74.125.135.27   25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/[UNIQUE_TOKEN
Sep 11 12:41:51 Ce5zZ82fPqMGJ2nHsh      128.3.41.71    17199   108.177.112.26  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/1c7b0520-5719-477d-80c9-
Sep 11 12:41:52 Ce5zZ82fPqMGJ2nHsh      128.3.41.71    17199   108.177.112.26  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 19:33:09 CyuaWo1swks7yqTo1a      209.85.215.71  37814   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 19:33:10 CxFkkb38v11QN3L09h      128.3.41.71    28232   108.177.98.27   25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 22:02:08 CFAUQp1v2JaxY97pq8      209.85.215.71  38902   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 22:02:08 CFAUQp1v2JaxY97pq8      209.85.215.71  38902   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 22:02:08 CFAUQp1v2JaxY97pq8      209.85.215.71  38902   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/53b53878-7f0c-4a9c-830d-
Sep 11 22:02:08 CFAUQp1v2JaxY97pq8      209.85.215.71  38902   128.3.41.120    25      downloads.careerpost.us http://downloads.careerpost.us
Sep 11 22:02:07 CP1Ygb2yuspXyo6ZX7      128.3.41.71    37900   173.194.202.27  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 22:02:07 CP1Ygb2yuspXyo6ZX7      128.3.41.71    37900   173.194.202.27  25      downloads.careerpost.us http://downloads.careerpost.us
Sep 11 22:02:07 CP1Ygb2yuspXyo6ZX7      128.3.41.71    37900   173.194.202.27  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-
Sep 11 22:02:07 CP1Ygb2yuspXyo6ZX7      128.3.41.71    37900   173.194.202.27  25      downloads.careerpost.us http://downloads.careerpost.us/e1006c/53b53878-7f0c-4a9c-830d-
```

Alerts

2017-09-11-19:29:50 CWnRow2SReNVDJOCnd 131.243.223.32 52451 54.236.212.118 80 tcp

Phish::RareURLClick

http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-926ee75f7926/? ####

[ts=1505158912.44534, uid=Ce5zZ82fPqMGJ2nHsh, from=Ingrid Peters <ipeters@lbl.gov>, to=asharma@lbl.gov, subject=FY17 Performance Review Problem, referrer=<uninitialized>] ####

[domain=downloads.careerpost.us, days_visited=[1504909578.01473, 1505156245.335107], num_requests=0, last_visited=1505183380.538836, trustworthy=F] -

131.243.223.32 54.236.212.118 80 - worker-22 Notice::ACTION_LOG 3600.000000 F

2017-09-11-19:29:50 CWnRow2SReNVDJOCnd 131.243.223.32 52451 54.236.212.118 80 tcp

Phish::HistoricallyNewAttacker

http://downloads.careerpost.us/e1006c/20870d7a-3f72-4da3-953b-926ee75f7926/? ####

[ts=1505158912.44534, uid=Ce5zZ82fPqMGJ2nHsh, from=Ingrid Peters <ipeters@lbl.gov>, to=asharma@lbl.gov, subject=FY17 Performance Review Problem, referrer=<uninitialized>] ####

[domain=downloads.careerpost.us, days_visited=[1504909578.01473, 1505156245.335107], num_requests=0, last_visited=1505183380.538836, trustworthy=F]

131.243.223.32 54.236.212.118 80 - worker-22 Notice::ACTION_LOG 3600.000000 F

Sep 11 19:28:32 131.243.x.y/127.0.0.1 %ASA-4-713228: Group = LBL-IPSec, Username = asharma, IP = 64.134.160.79, Assigned private IP address 131.243.223.32 to remote user

More example alerts

Date: Fri, 8 Sep 2017 02:38:04 -0700 (PDT)

From: bro <bro@bro.lbl.gov>

To: test@lbl.gov

Subject: [Bro] Phish::RareURLClick

Connection: CjjBiP3hSjecIpFKIa, [orig_h=128.3.5.17, orig_p=39017/tcp, resp_h=107.21.6.90, resp_p=80/tcp]

SMTP:: [ts=1481050626.364467, uid=Ch12mp1noGiPWMwtne, from=Frank Zuidema <fzuidema@lbl.gov>, to= xxxxx@lbl.gov , subject=Document review - Invitation to edit, referrer=[]]

HTTP:: [domain=lbl.gov.invoicenotices.com, days_visited=[1481051156.986024], num_requests=24, last_visited=1481051156.986024, trustworthy=F]

Clicked URLs:

http://lbl.gov.invoicenotices.com/0cb548/?login_id=c25acd74-aed4-43f3-89a5-563a03a0d9cc

Example Alert-2

Subject: [Bro] Phish::RareURLClick

Connection: [orig_h=128.3.153.65, orig_p=50212/tcp, resp_h=93.88.255.126, resp_p=80/tcp]

SMTP:: [ts=1504523783.99904, uid=CpOtuI2Hz7lEUKzODj, from="Training FSRM" <training@fsrm.ch>, to="XXXXXX YYYY"<XXXXXXYYYY@lbl.gov>, subject=Next FSRM courses (Attn. XXXXXX YYYY), referrer=<uninitialized>]

HTTP:: [domain=www.fsrm.ch, days_visited=[1503420632.37621, 1504864837.885044], num_requests=0, last_visited=1504864837.886409, trustworthy=F]

Clicked URLs:

http://www.fsrm.ch/gfx/social/In-2C-28px-TM.png

http://www.fsrm.ch/gfx/social/YouTube-logo-30.png

http://www.fsrm.ch/gfx/social/FB-f-Logo__blue_29.png

How to get smtp-url-analysis running

```
$ bro-pkg install smtp-url-analysis
The following packages will be INSTALLED:
  bro/initconf/smtp-url-analysis (master)

Proceed? [Y/n] Y
Running unit tests for "bro/initconf/smtp-url-analysis"
all 8 tests successful
Installing "bro/initconf/smtp-url-analysis"....
Installed "bro/initconf/smtp-url-analysis" (master)
Loaded "bro/initconf/smtp-url-analysis"
```

Questions ?

security@lbl.gov

SMTP Detection with Bro