Wikimedia Foundation v. NSA
No. 15-cv-0062-TSE (D. Md.)

Plaintiff’s Exhibit 30
Using the XKS CNE dataset and a DISGRUNTLEDDUCK fingerprint, we now see at least 21 TAO boxes with evidence of this intrusion set, most of which are associated with projects aimed at Iran WMD targets." -- MHS, July 2010

March, 2011

xks-cne@r1.r.nsa
The overall classification of this presentation is:

TOP SECRET//COMINT//REL TO USA, FVEY
What is XKEYSCORE?

- A suite of software running on a Linux host
- Classically, used for DNI processing, selection and survey
- A distributed hierarchy of servers at field sites and headquarters
  - Extract and tag metadata & content from traffic
  - Servicing analyst queries and workflows
- Web and programmatic front-ends
What is XKEYSCORE?

- A suite of software running on a Linux host
- Classically, used for DNI processing, selection and survey
- A distributed hierarchy of servers at field sites and headquarters
  - Extract and tag metadata & content from traffic
  - Servicing analyst queries and workflows
- Web and programmatic front-end
Example Search

• Let’s try a search for suspicious stuff...

  http_activity search, 5-eyes defeat, look for fingerprints:
  ndist/discovery/heuristic/BHAM/get_with_content or http/get/with_content

• While the search runs, some gotchas:
  • You choose where your query is run
  • Content and metadata age-off
  • Burden is on user/auditor to comply with USSID-18 or other rules
  • Geolocation based on IP
Search Results

Notes:
- Strange User-Agent
- Probably NOT CNE but definitely something non-standard
- Content: maybe a HTTP tunnel for some weird protocol?
- Should we write a Fingerprint?

Reset from local...
Fingerprints and Appids

- Useful for identifying classes of traffic or particular targets (for SIGDEV or collection):
  - mail/webmail/yahoo
  - browser/cellphone/blackberry
  - topic/s2B/chinese_missile

- appid – a contest, highest scoring appid wins
- fingerprint – many fingerprints per session
- microplugin – a fingerprint or appid that is relatively complex (e.g. extracts and databases metadata)
• Written in language called “GENESIS” (go genesis-language):

```python
appid('encyclopedia/wikipedia', 2.0) = http_host('wikipedia' or 'wikimedia');
fingerprint('dns/malware/MalwareDomains') = dns_host('erofreex.info' or 'datayakoz.info' or 'erogirlx.info' or 'pornero.info' or ...)
```

• If a fingerprint contains a schema definition, a search form automatically appears in the XKEYSCORE GUI

• Power users can drop in to C++ to express themselves
More about searches

- Many different searches
  - Base search is **Full Log DNI**
  - Depending on traffic type, will generate searchable results for (example):

<table>
<thead>
<tr>
<th>HTTP Activity</th>
<th>Network Information</th>
<th>GEO Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>Extracted Files</td>
<td>Email Addresses</td>
<td>Registry</td>
</tr>
<tr>
<td>Logins and Passwords</td>
<td>Document Metadata</td>
<td>Machine Info</td>
</tr>
</tbody>
</table>

- workflow – a user query that is run automatically usually every 24 hours
XKEYSCORE Gotchas

- Not all sites run latest XKEYSCORE software or fingerprints
- Fingerprint submission:
  - XKEYSCORE team weighs mission-worthiness of user fingerprints vs computational cost
- Content and metadata ageoff
• Lots of endpoint data flows into XKS
  TAO (no ECIs), GCHQ (almost all)
• Other limited flows include SIGINT
  Forensics Center, TAO STAT
• XKEYSCORE works well for endpoint data
• Sometimes the paradigm breaks (e.g.
  collected browser history file)
XKEYSCORE CNE (more)

- Payload types:
  - dirwalk, extracted file, system survey, network config, captured credentials, registry query, key logger, etc.
- Labeled dnt_payload in appid/fingerprint ontology
- Let’s look at some DANDERSPRITZ data...
XKEYSCORE CNE (more)
Recent Developments
- Upgrade of XKEYSCORE CNE
- Keyloggers: keylogger/perfect/extension
- PCAP Reingestion
- Router Redirection
Counter CNE Methodology

(refer to Counter CNE Resources slide...)

- **Hypothesis/research-driven**
  - “Could South Korean CNE be using similar selectors to FVEY CNE?”
  - “What keywords could be used to find keyloggers ("example: keylog OR keystroke")

- **Bogus or Unusual Traffic**
  - HTTP GET with content (example in this presentation)
  - HTTP POST at odd hours (from Russia 0200-0359Z)
  - Funky user agents

- **Known-Host or User driven (e.g. drop sites)**

- **XKEYSCORE is GOOD at these kinds of things**
CNE-Specific

- Registry searches (e.g. SIMBAR)
- Fused Active/Passive search
  - common selectors
  - document hashes
- Known Processes (malicious executables or code)
  ... Let’s enhance the process list appid
- map-reduce within CNE cluster using GENESIS calls
XKEYSCORE Doesn’t Do...

- ... at all (well, automatically, anyways)
  - Paired traffic heuristic-based approach
    - HTTP[S] imbalance (e.g. GET without response)
    - IP/DNS mismatch*
- ... on an automatic basis
  - Network or host characterization
  - Changes in IP/DNS mapping over time
  - Changes over time in malware comms
Counter CNE Resources

- How to Discover Intrusions [using XKEYSCORE] by [name and (paper)]
- MHS INDEX – Foreign CNE Discovery Page
  https://wiki.itd.nsa/wiki/Foreign_CNE_Discovery
- CSEC and GCHQ – DONUT (unknown protocols):
  https://tiso.sigint.cse/snipehunt/index.php/DONUT
- GCHQ Discovery Posted some Research of Detecting Man-on-the-Side Attacks:
  https://tiso.sigint.cse/snipehunt/index.php/MOTS
- GCQH Disco Team posts POC’s for different Intrusions and some Details:
  https://wiki.gchq/index.php/Discovery
- The GCHQ DISCO team also posts Discovery Theories they run once a week:
  https://wiki.gchq/index.php/Discovery_Afternoons
- XKEYSCORE Fingerprints
Using TAO-obtained Iranian implant encryption keys, inline decrypt using XKS microplugin – IRGC-QF keylogger data!
Points of Contact

- MHS Index Team
  - [redacted]: [redacted]@nsa.ic.gov

- CES/TRANGRESSION
  - [redacted]: [redacted]@nsa.ic.gov
  - [redacted]: [redacted]@nsa.ic.gov

- NSA/Countering Foreign Intelligence
  - [redacted]: [redacted]@nsa.ic.gov

- NTOC ??

- XKEYSCORE
  - [redacted], [redacted]: xks-cne@r1.r.nsa