Outline

• Introduction
• Some motivating stories: real-life attacks
• Efficient network mapping with “Lazy Scan” mode
• Layer 7 extensions
• Scripting Extensions
• Data Mining and Experimental Data sharing network
Introducing presenter

- Fyodor.Y

Interests:
- Intelligence collection/analysis
- Network discovery and network protocols
- AI
Attack Trends
China vs. Taiwan
briefs of cyber “wars”
Mystic redirects
(2009/03/05)
• Large number of users were redirected to malware-infected servers, while trying to visit legitimate web sites hosted outside of Taiwan island (i.e. zdnet, msn.com, etc)
Traces
Guess..

- A node was compromised somewhere en-route. TCP connections were non-blindly hijacked...
Tracing “ghost” node(s)

• some “spaghetti” to quickly discover the node
Discovered attack scenario
Lesson learnt

- Large number of target nodes are to be probed in order to identify potential ‘en-route’ attacks.
- We need a high-performance network discovery tool, capable of operating at Layer7
- we need automated tracing capability
more stuff @L7...
Motivation

• we need more application-level probes
And..

• we could actually correlate L7 data with network probing results
but ..

• we need to minimize network load, because L7 might mean “lots of traffic”
Also..

• Time is another player. We want to be able to monitor network fluctuations in time
So, the Xprobe

now “NG”
Xprobe

- The historical note:
- Xprobe project started as remote fingerprinting tool to probe remote systems using **ICMP** protocol queries.
- Other protocols support was added later. **Fuzzy fingerprinting** mechanism was introduced to improve precision.
Further motivation

- Exploring other protocols running on the top of IP
- Bulk scanning
- Probing “en-route” systems
- Migrating to IPv6
- Honeypots/Nets
- Improving precision by cross-correlation over time
On the top of IP

- SCTP/Sigtrans gateways
- IPv4 to IPv6 gateways
- ...
“en route” findings

- Caching systems, transparent proxies etc.
- L7 switches
- Reactive IDS/IPS
- Application Firewalls
- Active spoofing attacks
Honeypots

- Virtual Machines
- Virtual Networks
- Incomplete Services
Bulk Scanning

- Probing “en-route” devices by large-range scans
- IPv6
Data cross-correlation

- Currently correlating data between L7 and network layers.
Current Improvements
Minimizing Network Load

- Information Gain metrics
- "Lazy-Mode" execution
- "Target" driven execution
- New Scan engine (in progress)
Improving Precision

- Cross correlation between L7 and below
Improving Usability

• Language Extensions: Python (xprobepy)
Information Gain
Information gain

- A “score” calculated for a probe, characterizing how much “information” the probe is going to bring
Benefits

• Highest information gain probes are executed first
• “0” information gain probes are not executed (unless are part of dependency)
• Possible to optimally minimize network overhead by executing “top X”/target
Algorithm

- success
  - exec. dependency
  - not satisfied
    - top module data req?
      - satisfied
        - execute
      - failed
        - disable
  - satisfied
  - reorder by gain
    - false
      - no more modules?
        - true
          - done
        - false
          - all modules gain = 0?
            - true
              - done
            - false
              - not satisfied
                - top module data req?
                  - satisfied
                    - execute
Lazy scan and target-driven execution

discovery process optimizations
Architecture, briefly..
Data dependency chains

• Each module is characterized with type of data it “requires” and “provides”
Data Dependency based execution

- Find dep. provider
- Enough data?
- Exec
- Mark
- Return exec. dependency
- Return fail
- Return success
No “portscan” per se

• This technically makes port scanning “AS IS” unnecessary
• Significantly reduces tool “noise” on the wire
Wire “noise” rough comparison
Benefits of Data Chaining

• Probe focused execution (by specifying “intended” probe)

• Restrictions can be set:
  • no more than X queries/target
  • use only “normalized” packets
Negative impact

• You still may not know about certain ports and applications running on the target system.
Application level
Application level

- Improving fingerprinting precision
- “en-route” interaction
- Honeypots
L7 fingerprinting

- Underlying OS can be probed via L7 tests and correlated with other data

<table>
<thead>
<tr>
<th>Test type</th>
<th>Usable Protocol</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directory Separator</td>
<td>HTTP</td>
<td>Win/Unx</td>
</tr>
<tr>
<td>New line characters</td>
<td>HTTP</td>
<td>Win/Unx</td>
</tr>
<tr>
<td>Special/reserved filenames</td>
<td>HTTP</td>
<td>Win/Unx</td>
</tr>
<tr>
<td>Root directory</td>
<td>FTP</td>
<td>Win/Unx.</td>
</tr>
<tr>
<td>Special characters (EOF,EOL)</td>
<td>HTTP, FTP</td>
<td>..</td>
</tr>
<tr>
<td>Filesystem limitations</td>
<td>HTTP, FTP</td>
<td>..</td>
</tr>
<tr>
<td>Filesystem illegal characters</td>
<td>HTTP, FTP</td>
<td></td>
</tr>
<tr>
<td>Case sensitivity</td>
<td>HTTP, FTP</td>
<td>Win/Unx</td>
</tr>
<tr>
<td>Special filenames handling</td>
<td>HTTP, FTP</td>
<td>Win/Unx</td>
</tr>
<tr>
<td>Special files in directory</td>
<td>HTTP, FTP</td>
<td>Win/Unx</td>
</tr>
<tr>
<td>Binary file fingerprinting</td>
<td>FTP</td>
<td>Win/Unx</td>
</tr>
</tbody>
</table>
Honeypots
VM tricks

- Possible to identify VMs (not all) by TCP stream analysis
Network level tricks

- Analyzing MAC addresses, when available
Application Level Tricks

- We can probe for incomplete implementations of L7 protocols
Current Developments
Work in progress

- Language bindings
- L7 modules
- new engine
- en-route modules
Future Plans

• By designing distributed data sharing network it’d be possible to collect Multi-dimensional data
IPv6 Action plan

- Local node discovery: straightforward (multicast)
- Remote segments: DNS, text file parsing, “educated” guessing, search engine, beforementioned networking capability
Availability

http://xprobe.sourceforge.net
(git push in a couple of days)
http://github.com/fygrave/xprobepy
(due Mid of July)
Questions

if you have no questions, feel free to throw your shoe ;-)  
*jk*

fygrave@o0o.nu
(o-zero-o)