Tachyon

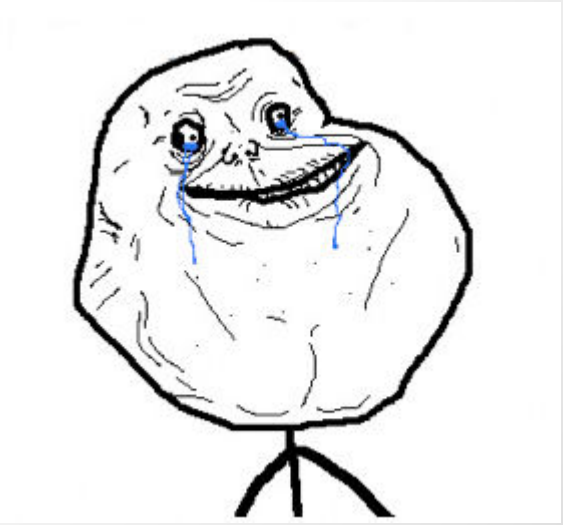
Sloppiness is bliss

µdrink

- This talk implements udrink
- When I do a mistake and you call it, iDrink.
- At the end of the talk if you ask a good question, uDrink.
- The drink of the day is....

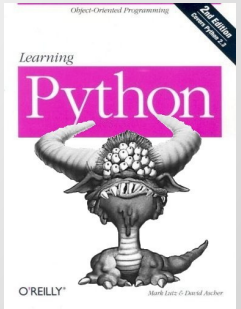
Seriously, who is this guy?

- Security Hobbyist for more than 10 years
- CTF Monkey with the CISSP Groupies
- Hackus 2012 Python Track sadist-in-chief
- Homebrewer
- Amateur Photographer
- Retard



What is Tachyon

- Offensive tool for penetration testers
- Weapon against sysadmin sloppiness
- Intelligent Web discovery tool
  - _hidden_ files and folders
  - Backups
  - Temporary copies
  - Test and Dev artifacts
  - Dumps and more!



What it is not (use a british accent)

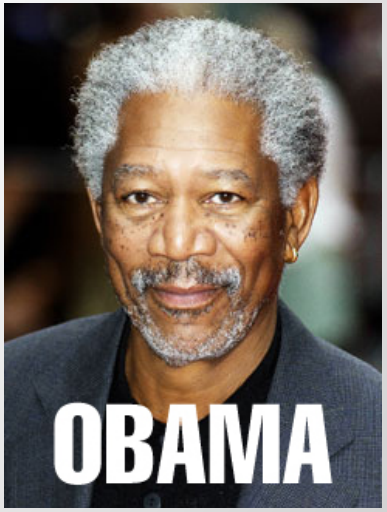
- Vulnerability scanner
- All-purpose framework
- Reporting tool
- Web Crawler
- Dumb bruteforcer

Nothing else does it correctly

Why not <insert tool name> ?

- Dirbuster
  - Grotesque bruteforcer (/Queer, /richard_macmanus)
  - Unmaintained (2009)
  - No plugin architecture
  - Dumb

- Skipfish
  - False positive factory
  - In C! c'mon man we're in 2011!
  - No plugin architecture



Why not <insert tool name> ?

- Nessus, Vega, etc...
  - Vulnerabilities scanners
  - Fancy and heavy UI
  - Reporting tools
  - Not made for this task

- Nikto
  - Slower than water at 0K
  - Vulnerability scanner
  - Database clusterfuck

Features

- Tor support
- Plugin engine
- Fast multi-threading
- HTTPS
- DNS caching
- Automatic variable rate limiter
- Recursive scanning (credits: EiNSTeiN@CISSP Groupies)
- Cutting edge 2.0 command line interface


How It's Made

Path database

- Cheap (The root of evil)
- Each scan start with a path lookup
- Python dict (Plugins++)

```
# Interresting path
{"url" : "/_assets", "description" : "_assets directory"}
{"url" : "/adm", "description" : "Admin section"}
{"url" : "/admin", "description" : "Admin section"}
{"url" : "/administration", "description" : "Admin section"}
{"url" : "/administrator", "description" : "Admin section"}
{"url" : "/auth", "description" : "Auth section"}
{"url" : "/ajax", "description" : "Ajax section"}
```

File database

- Costs a lot more
- Provides more control
  - "No suffix" : Ignore all extension matching
  - "match-string" : String must match to be a valid hit

```
{"url" : "rpc", "description" : "Website API"}
{"url" : "json", "description" : "Website API"}
{"url" : "json-api", "description" : "Website API"}

# PHP common developer files
{"url" : "info.php", "description" : "Common phpinfo() wrapper script", "match_string" : "PHPE9568F34-D428-11d2-A769-00AA001ACF42", "no_suffix": True}
{"url" : "infos.php", "description" : "Common phpinfo() wrapper script", "match_string" : "PHPE9568F34-D428-11d2-A769-00AA001ACF42", "no_suffix": True}
```

Extensions database

- Cost more than a PONY!
- _Hardcoded_ (Python, yeah right...)

```
# Values used to generate file list (should be configurable?)
file_suffixes = ['.sql', '.bak', '-bak', '.old', '-old', '.dmp', '.dump', '.zip', '.rar', '.7z',
                 '.tar.gz', '.tar.bz2', '.tar', '.tgz', '-bak', '~', '.conf.old','.conf',
                 '.conf.orig', '.conf.bak', '.cnf', '.ini', '.inc', '.inc.old', '.inc.orig', '.log', '.txt', '_log',
                 '.php.old', '.php.inc', '.php.orig', '.sql.old', '.sql.bak', '0', '1', '2', '.xml',
                 '.csv']
```

Plugins

- Access to almost everything in Tachyon
  - Data structures
  - Classes
  - Python does _not_ support protection :)
- Dedicated execution levels:
  - Before path test
  - Before file-path combination
  - Link validation level (still in dev)
- No documentation for now, but some examples


Instructions:
Nursing Baby
YES
NO

Semi automatic scanning

- Most efficient
- You need to know what you are doing
- Usefull for manually discovered paths
- Two modes
  - Search only for subpath in specified path
  - Search only for files in specified path

Automatic scanning

- Lazy _lulzsec_ mode
- Longer (a lot in some cases)
- More false positives
- Better to find irregularities (/img/pass.txt)

Recursive scanning

- Powerful but hard to control
- Lot of assumptions on recursive directory existence
- Easier to use in semi-automatic mode
- Depth limitable

Tor support

- Trough Privoxy in http proxy mode
- A lot slower
- More prone to timeouts
- Useful for WAF evasion
  - Tor has pseudo-random latency

A false positive story

- Incoherent Error codes
- Redirections
- Batshit insane webserver
- Variable 404 output

False positive detection

- Benchmark target
  - CRC32 evaluation
  - Generate 404's with uuid
  - Probe common file handling
    - .html, .php, .asp, .txt, no-ext
  - Store each 404 CRC in lookup table

False positive detection

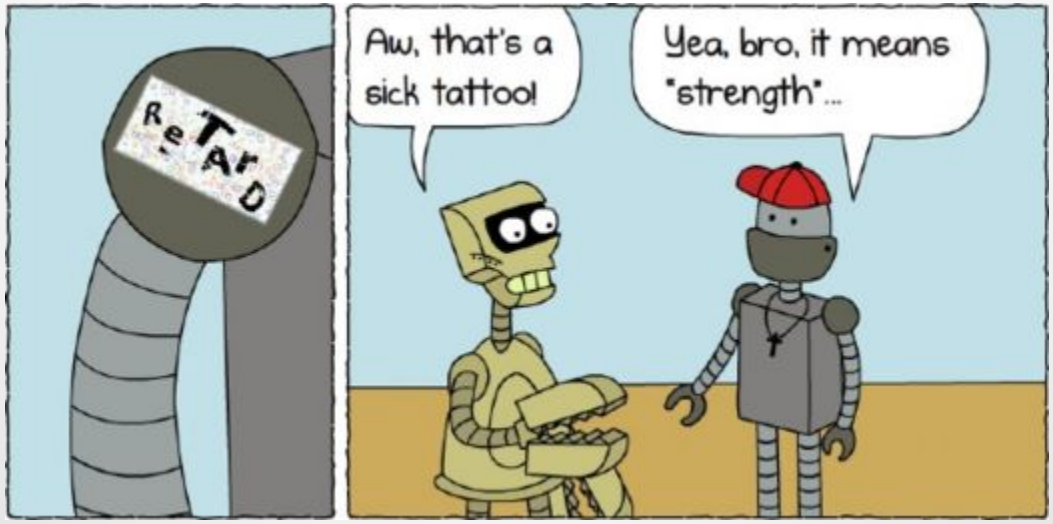
- Redirect evaluation
  - Using Ratcliff-Obershelp (not levenshtein!!!)
  - pony.com -> www.pony.com == Valid
  - pony.com/pony.txt -> www.pony.com/pony.txt == Valid
  - pony.com/test -> pony.com/error/404 -> 200 -> invalid!

False positive detection pitfall

- CRC32 Testing of first 200 bytes
  - False positive if there's a timestamp in those bytes
  - False positive if there's anything variable in those bytes
- Ratcliff-Obershelp
  - Too much granularity in some cases
    - what is the best differential ratio? 0.6 VS 0.65?

Plugins - Robots.txt parser

- Leverage your little "Secrets"
- Dissalow path : new target path
- Dissalow file: new target file



Plugins - Path Generator

- Used to generate trivial paths and files
- Reduce database pollution
  - /0-9
  - /a-z

Plugins - Host Processor

- Try to generate pobable filenames with hostname
- adomain.com -> domain, adomain, adomain.sql etc...

Plugins - /.svn/entries parser

- Parse xml svn entries
- Old non-xml format will be supported
- Generate paths

Limitations

- Recursive scanning
  - Predictability
  - <Forbidden 1>/<Forbidden 2>/ == 403. Does not guarantee "2" exists (nginx)
- Various edge cases
- Tor support trough privoxy only

Todo

- Faster (urllib3: keep-alive pooling)
- Cute and/or parseable output
- Plugin system
  - Callbacks
  - Documentation
- Pattern eclusion
- HTML Crawler (buy me _some_ **good** beer then maybe)

Demo!

You want to contribute?

- Google "github tachyon"
- Contact me first!!! (initnull@gmail.com)
- You **need** a github account
  - no i won't handle your un-mergeable .patch
  - I work on Linux **AND** Windows
  - I don't care
- Priorities:
  - Documentation
  - Clever plugins
- Also looking for website directory listing

The end!