

# Forum:[Hack Forums](#)

Thread: [Python Proxy Leecher](#)

v0id\_v0mit - December 28, 2010, 1:15 pm

Proxy leecher, leeches proxies of a list of urls.

It extracts the ip:port from the page and stores it, and the sorts it.

It reads the urls from a file and stores the proxies in another, it also makes a list of the sites it found proxies on.

<http://pastebin.com/G12CHaEB>

<http://darkpy.net/python/pypb.py>

Code ""

Created on 28. des. 2010

@author: v0mit

""

```
import urllib, threading, re, sys, Queue, time
```

```
if len(sys.argv) < 2:
```

```
    sys.exit("Usage python pypb.py <proxylist.txt> [-t threads] [-l leechlist] [-s sort]")
```

```
version = 0001
```

```
def getorders():
```

```
    name = sys.argv[1]
```

```
    try:
```

```
        maxThreads = int(sys.argv[sys.argv.index("-t")+1])
```

```
    except ValueError: maxThreads = 40
```

```
    try:
```

```
        leechlist = sys.argv[sys.argv.index("-l")+1]
```

```
    except: leechlist = "sites.txt"
```

```
    try:
```

```
        if sys.argv.index("-s"): bSort = True
```

```
    except ValueError: bSort = False
```

```
    return [name, maxThreads, leechlist, bSort]
```

```
def printlogo():
```

```
    print("\t      _ _ _ _ _")
```

```
    print("\t    / _ \   ( _ |")
```

```
    print("\t _ _ _ | | | _ _ _ _ | _")
```

```
    print("\t \ \ / / | | | ' _ \ | _ |")
```

```
    print("\t \ V / | | | | | | | _")
```

```
    print("\t \_ / \_ / | | | | | _")
```

```
    print("\tversion:%s    v0mit@darkpy.net\n" % version)
```

```

def sort():
    proxylist = open(orders[0], "r")
    setList = set()
    nCounter = 0

    for aaa in proxylist:
        aaa = aaa.strip()
        setList.add(aaa)
        nCounter += 1

    proxylist.close()
    proxylist = open(orders[0], "wb")
    for bbb in setList:
        proxylist.write(bbb+"\n")

    proxylist.close()

    print("%s proxies in list, %s unique proxies in list." % (nCounter, len(setList)))

class scan_thread(threading.Thread):
    def run (self):
        while True:
            url = threadPool.get()
            try:
                get = urllib.urlopen(url)
            except IOError:
                threadPool.task_done()
                continue
            data = get.read()
            validProxies = re.findall('(?:[d]{1,3})\.(?:[d]{1,3})\.(?:[d]{1,3})\.(?:[d]{1,3})\.(?:[d]{2,5})',
data)
            print("[*] %s proxies found at %s" % (len(validProxies), url))
            for proxy in validProxies:
                proxylist.write(proxy+"\n")

            if len(validProxies) > 0:
                good_sites.write(url+"\n")
                threadPool.task_done()

orders = getorders()
printlogo()

try:
    proxylist = open(orders[0], "a")
except IOError:
    print("[!] Could not create/open %s" % orders[0])

try:
    sites = open(orders[2], "rb")
except IOError:

```

```

print("[!] Could not open %s" % orders[2])
sys.exit()

try:
    good_sites = open("good_sites.txt", "wb")
except IOError:
    print("[!] Could not create good_sites.txt")
    sys.exit()

threadPool = Queue.Queue(0)

for x in xrange(orders[1]):
    scan_thread().start()

while True:
    for url in sites:
        url = url.strip()
        if url.startswith("http://"):
            threadPool.put(url)
        else:
            threadPool.put("http://" + url)
    while True:
        if threadPool.empty():
            break
        time.sleep(1)

    threadPool.join()
    break

print("[*] Done leeching.")
proxylist.close()
good_sites.close()
if orders[3]: sort()
A list with some good working urls:
http://darkpy.net/python/good\_sites.txt
http://pastebin.com/WSfUZyhk

```

**Arata Goru** - January 28, 2015, 6:49 pm

How do I use this python tool on Ubuntu?

Pls, guide me .. I am new to Ubuntu Linux.

Thank you