

Performance Improvements in Samba's Pathname Handling

SDC 2022
Fremont

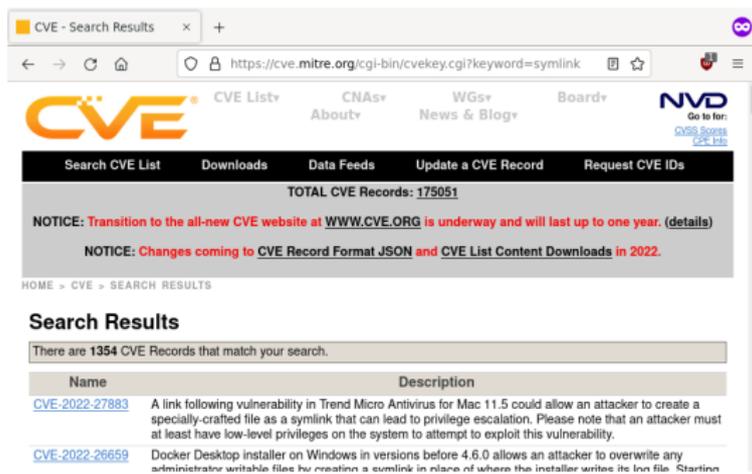
Volker Lendecke

Samba Team / SerNet

2022-09-14

Symlink race – Why care?

- ▶ Search for “symlink” in Common Vulnerabilities and Exposures:



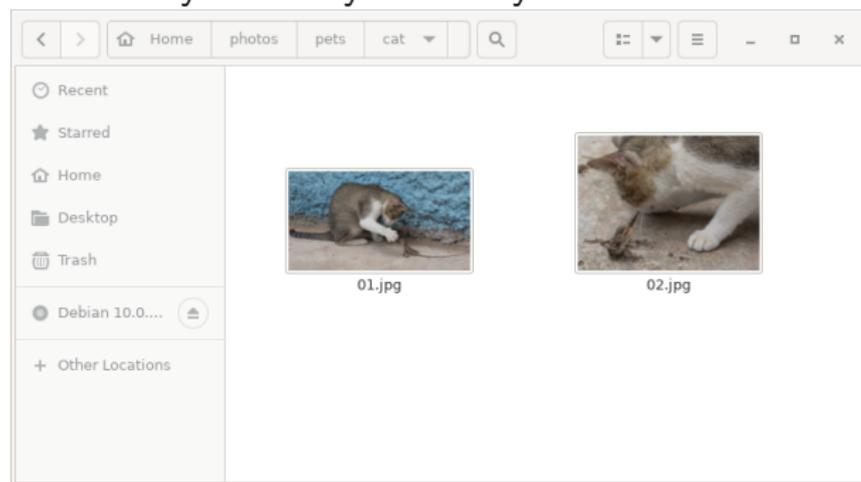
The screenshot shows a web browser window with the URL <https://cve.mitre.org/cgi-bin/cvekey.cgi?keyword=symlink>. The page header includes the CVE logo and navigation links for CVE List, CNAs, WG's, Board, and NVD. Below the header, there are navigation tabs for Search CVE List, Downloads, Data Feeds, Update a CVE Record, and Request CVE IDs. A black banner displays "TOTAL CVE Records: 175051". Two notices are present: one about the transition to the new CVE website (WWW.CVE.ORG) and another about changes to the CVE Record Format JSON and CVE List Content Downloads in 2022. The search results section shows "There are 1354 CVE Records that match your search." and a table with two entries:

Name	Description
CVE-2022-27883	A link following vulnerability in Trend Micro Antivirus for Mac 11.5 could allow an attacker to create a specially-crafted file as a symlink that can lead to privilege escalation. Please note that an attacker must at least have low-level privileges on the system to attempt to exploit this vulnerability.
CVE-2022-26659	Docker Desktop installer on Windows in versions before 4.6.0 allows an attacker to overwrite any administrator writable files by creating a symlink in place of where the installer writes its log file. Starting

- ▶ Ok, there’s something going on – but what??

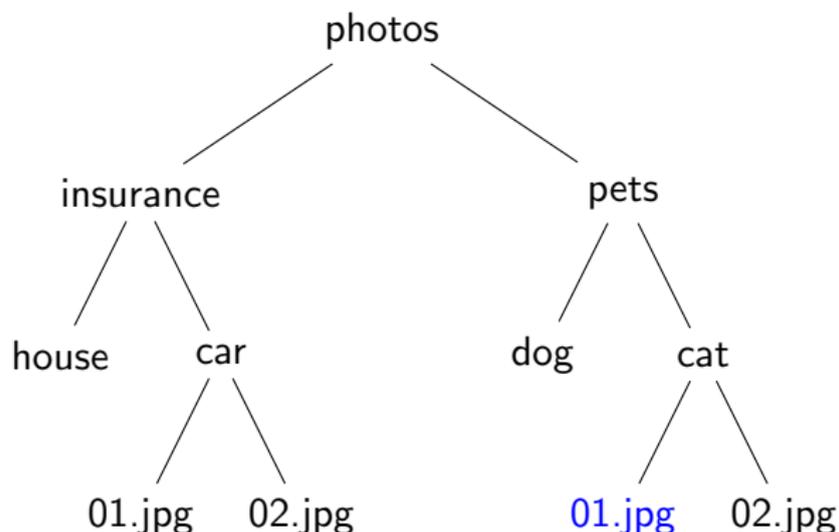
Publish cat photos to my blog

► Browse my directory hierarchy



- Copy file path `/photos/pets/cat/01.jpg` into Browser
- Press upload

My directory hierarchy

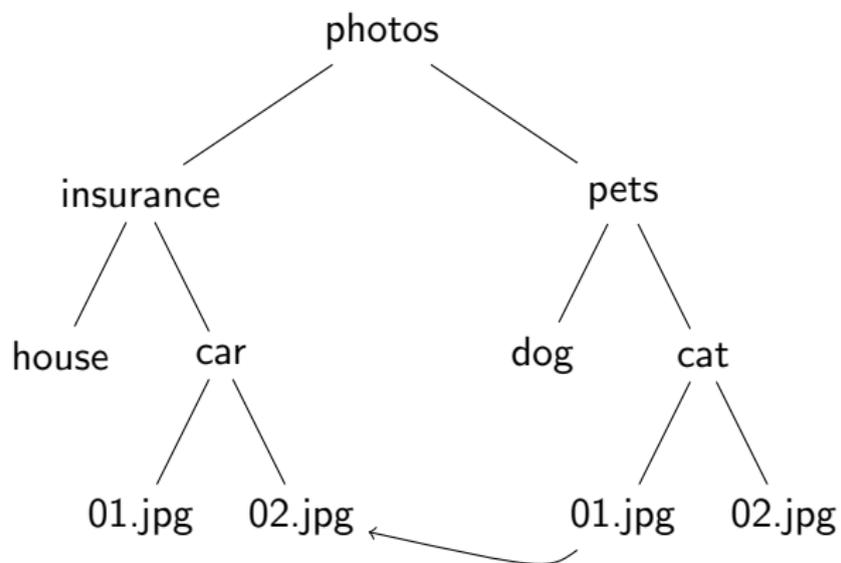


- ▶ Cute cat under `/photos/pets/cat/01.jpg`

Symlink Race

- ▶ Publish cat photo
 - ▶ Browse my directory hierarchy
 - ▶ Copy file path `/photos/pets/cat/01.jpg` into Browser
 - ▶ A few seconds for an **ATTACK**
 - ▶ Press upload
- ▶ The attacker replaces “01.jpg” with a symbolic link to the latest communication with my insurance
- ▶ In `-sf /photos/insurance/car/02.jpg /photos/pets/cat/01.jpg`

My new directory hierarchy



What gets published?

- ▶ /photos/pets/cat/01.jpg gets uploaded, but... oops



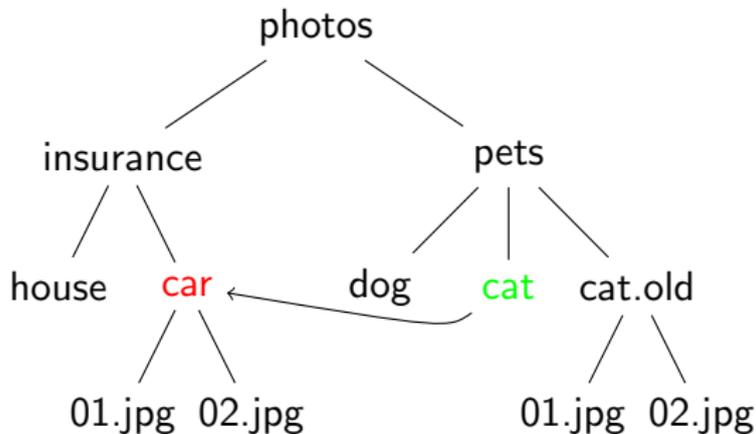
- ▶ This could have been your Passwords.DOCX

O_NOFOLLOW

- ▶ Posix/Linux can protect against this kind of attack
- ▶ Opening a file in Posix via C function `open()`
- ▶ Adding the flag `O_NOFOLLOW` to `open()` ensures “01.jpg” is not a symlink:
 - ▶ If the trailing component (i.e., basename) of pathname is a symbolic link, then the open fails, with the error `ELOOP`
- ▶ When using `O_NOFOLLOW`, the upload function will get an error
- ▶ However:
 - ▶ Symbolic links in earlier components of the pathname will still be followed.

My latest directory hierarchy

- ▶ Another **attack**



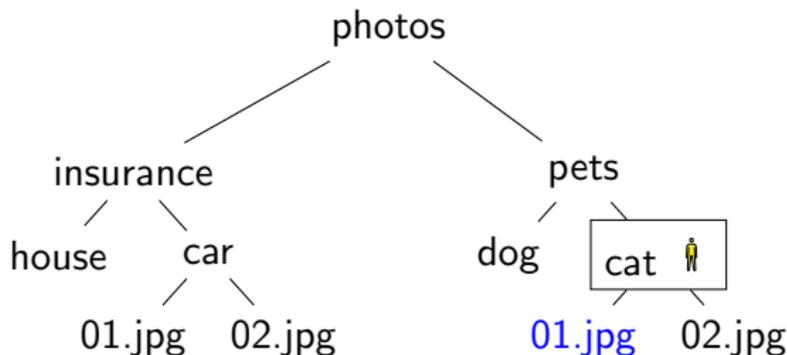
- ▶ `/photos/pets/cat` → `/photos/insurance/car`
- ▶ `/photos/pets/cat/01.jpg` → `/photos/insurance/car/01.jpg`

What gets published?

- ▶ /photos/pets/cat/01.jpg gets uploaded, but... ouch!



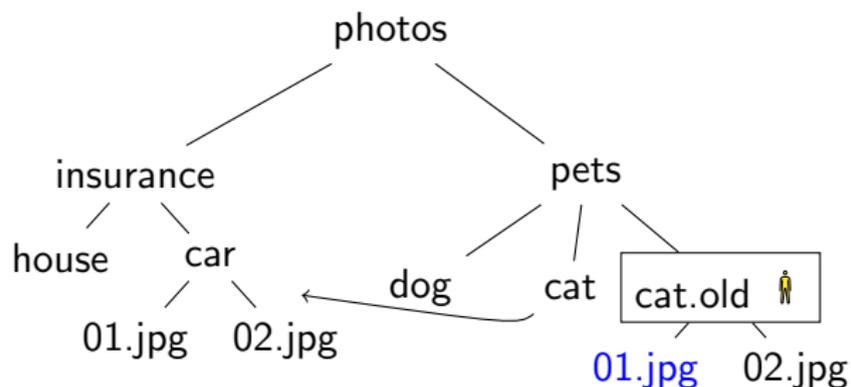
Change into /photos/pets/cat



► Four expensive steps

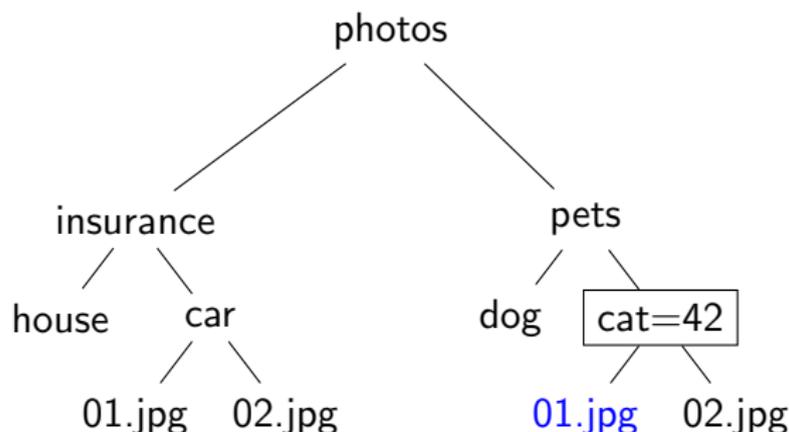
- `chdir /photos/pets/cat`
- check I am really in `/photos/pets/cat`
- `open("01.jpg", O_NOFOLLOW)`
- `chdir /`

Attack mitigated



- ▶ Attacker moves me to cat.old
- ▶ `open("01.jpg", O_NOFOLLOW)` references the correct file
- ▶ Opening files becomes expensive due to the four steps

Openat-Call



- ▶ Four expensive steps done just once
- ▶ Hold and cache reference to /photos/pets/cat (=42)
- ▶ Cheap openat(42, "01.jpg", O_NOFOLLOW)

Symlink races from 30,000ft

- ▶ Paths in Posix are prone to input validation problems
- ▶ Symlink races are a Time-Of-Check Time-Of-Use (TOCTOU) problem
- ▶ Sanitizing paths is possible, but tedious
⇒ Nobody does it
- ▶ Current status in Samba with 4.16:
 - ▶ Many places in Samba chose the `chdir/check/open/chdir` way
 - ▶ Work is ongoing to pass directory handles (the “42”) everywhere
- ▶ Work ongoing in Linux to make sanitizing easier:
 - ▶ `opent2(RESOLVE_NO_SYMLINKS)`
 - ▶ Mount option disallowing symlinks?

Samba 4.16

- ▶ OEMs complain to SerNet that Samba 4.15 is slower than 4.12
 - ▶ Benchmarks like specsfs swbuild do a lot of stat-opens
 - ▶ Just ask for file metadata like timestamps or filesize
- ▶ Number of syscalls is a problem:
- ▶ `strace -p <smbd-pid> of smbclient -c "open sub/0/1/2/3.txt"`:
 - ▶ v4-12-strace: 29 lines
 - ▶ v4-16-strace: 45 lines
 - ▶ v4-17-strace: 17 lines
- ▶ v4-17 removes `chdir/check/open/chdir` in initial pathname processing
 - ▶ More work ongoing to remove many more of those
 - ▶ We need to cache directory handles from initial processing

More work to do

- ▶ Remove calls to `non_widelink_open()`
 - ▶ This does the `chdir/check/open/chdir` steps
- ▶ Deleting a directory:
 1. Open file with `DELETE` access permissions
 - ▶ Check parent access for `DELETE_CHILD`
 2. Set `DELETE_ON_CLOSE` disposition
 - ▶ Check for empty directory
 3. Close
 - ▶ `parent_pathref()` to prepare for `unlinkat()`
- ▶ In every step there is one or more `non_widelink_open()`
- ▶ We need to cache the parent fd with every file handle

How to speed up filename handling?

- ▶ test_chdir: Do what we do today
 - ▶ Get current working directory
 - ▶ Verify full pathname with realpath()
 - ▶ chdir() into target directory
 - ▶ Verify “.” inside current working directory
 - ▶ open(“filename”, O_NOFOLLOW);
 - ▶ chdir back
- ▶ test_openat:
 - ▶ next = openat(dirfd, “single component”, O_PATH|O_NOFOLLOW);
 - ▶ close(dirfd);
 - ▶ dirfd = next;
 - ▶ This is the path to proper SMB2 server-side symlink handling
- ▶ test_openat2: Just use RESOLVE_NO_SYMLINKS

Measuring smbd runtime

- ▶ Fetch 10k files in 3-level hierarchy with smbclient tar-mode
- ▶ All numbers are in milliseconds
- ▶ Run “time smbd -d0 -i”, ignore wall-clock time

	1st	2nd	3rd	avg	relative
4.12 user	863	796	863	840	100%
sys	744	851	756	783	100%
4.16 user	1100	1089	983	1057	125%
sys	881	1033	975	963	122%
master user	1011	979	958	982	116%
sys	737	708	800	748	95%

- ▶ We have to work on user space now
 - ▶ Stefan Metzmacher is right now optimizing internal lock handling

Thanks for your attention

```
vl@samba.org / vl@sernet.de  
https://www.sernet.de/  
https://www.samba.org/
```