CVE-2023-0386 - Informe Técnico

Nombre de la vulnerabilidad: CVE-2023-0386

Versión afectada: Kernels Linux entre versiones **5.11 hasta 6.1.8** (excepto la corrección posterior, como en 5.15.91).

Fecha de publicación: Enero 2023

Severidad: Alta (Escalado de privilegios locales)

Herramientas / Componentes involucrados:

- Archivo ejecutable escrito en (C, Python, GO)
- OverlayFS (sistema de ficheros superpuesto)
- FUSE (Filesystem in Userspace) opcionalmente para exploits

Tipo de vulnerabilidad: Escalada de privilegios local mediante error en la validación de metadatos de archivos en OverlayFS.

Descripción de la vulnerabilidad:
OverlayFS es un sistema de ficheros que permite combinar dos directorios, generalmente llamados "lowerdir" (solo lectura) y "upperdir" (escribible).
El fallo ocurre porque el kernel no valida correctamente los metadatos de los archivos que provienen del lowerdir, permitiendo que un archivo con UID/GID falsificados y el bit setuid se copie al upperdir como si tuviera privilegios de root legítimos.

Impacto:
Un usuario local puede crear un archivo especial en un lowerdir manipulado (por ejemplo, usando FUSE) y, al combinarlo con OverlayFS, obtener un archivo con permisos de root en el upperdir. Ejecutar este archivo podría otorgar privilegios de root en el sistema.

Explicación técnica detallada:

1. **OverlayFS:** Permite combinar un lowerdir y un upperdir en una vista única. Las modificaciones se escriben en el upperdir mientras que el lowerdir permanece intacto.
2. **Lowerdir manipulado:** Un atacante puede usar un filesystem falso, como uno montado con FUSE, para presentar archivos con UID 0, GID 0 y setuid activado.
3. **Error de validación:** OverlayFS copiaba estos archivos al upperdir sin comprobar si los metadatos eran legítimos en el namespace del usuario, resultando en archivos efectivamente setuid root.
4. **Payload:** Un binario mínimo que ejecute `/bin/sh` con `setuid(0)` puede ser suficiente para obtener una shell de root.

5. **FUSE:** Sirve como herramienta de usuario para crear lowerdirs falsos que engañen a OverlayFS. No se modifica el binario exploit en sí; FUSE solo genera los metadatos falsos visibles en lowerdir.

¿Cómo se corrigió?
El parche añade una verificación que asegura que el UID y GID del archivo estén mapeados en el namespace del usuario actual, usando kuid_has_mapping y kgid_has_mapping, abortando la operación si no lo están.

Glosario de Términos

---

1. OverlayFS

- Sistema de ficheros que combina dos directorios (lowerdir y upperdir) en una sola vista.
- *Lowerdir*: directorio de solo lectura.
- *Upperdir*: directorio escribible.
- Ejemplo técnico: OverlayFS combina `/usr` (lowerdir) y `/tmp/overlay` (upperdir) para ver un solo árbol de archivos.
- Ejemplo analógico: Imagínate una biblioteca donde los libros no se pueden escribir (lowerdir). Ponés una hoja de calcar encima (upperdir) y escribís en ella: parece que modificaste el libro, pero el libro original sigue intacto.

2. FUSE (Filesystem in Userspace)

- Permite crear un filesystem virtual desde el espacio de usuario.
- Se puede inventar metadatos y archivos sin ser root.
- Ejemplo técnico: Montar un directorio `/mnt/fusemnt` donde `ls` muestra archivos que realmente no existen en disco.
- Ejemplo analógico: Es como tener una vitrina de exhibición de libros falsa: podés poner etiquetas y decir que son de alguien específico, pero no existen físicamente.

3. UID y GID

- UID: Identificador de usuario.
- GID: Identificador de grupo.
- Ejemplo técnico: root tiene UID 0 y GID 0.

4. setuid

- Bit especial de permisos que hace que un programa se ejecute con los privilegios de su propietario.
- Ejemplo técnico: Un binario con setuid root se ejecuta como root aunque lo lance un usuario normal.

5. Namespace

- Entorno aislado donde se pueden mapear UIDs/GIDs sin afectar al sistema global.
- Ejemplo técnico: UID 0 dentro de un namespace puede no ser root real fuera de él.
- Ejemplo analógico: Es como jugar a ser el director de la biblioteca dentro de un salón de juego: tenés poderes ahí, pero no afectan al verdadero director.

6. Lowerdir manipulado

- Lowerdir creado con FUSE o similar para presentar archivos con UID/GID falsificados y bit setuid.
- Ejemplo técnico: Un archivo "`exploit_bin`" con UID 0, GID 0, setuid activo dentro de un FUSE mount.
- Ejemplo analógico: La hoja de calcar tiene una firma falsa del director, aparenta ser oficial, aunque no lo es.

7. Upperdir

- Directorio escribible donde OverlayFS copia los archivos del lowerdir.
- Ejemplo técnico: `/tmp` usado como upperdir.
- Ejemplo analógico: La hoja de calcar está sobre un libro original; cualquier cambio que hagas se ve reflejado allí sin tocar el libro.

8. Payload

- Código mínimo que se ejecuta para conseguir la escalada de privilegios.
- Ejemplo técnico: Binario que ejecuta `/bin/sh` con `setuid(0)` para abrir una shell de root.

9. Parche / corrección

- Se añade verificación de mapeo de UID y GID usando `kuid_has_mapping` y `kgid_has_mapping`.
- `kuid_has_mapping`: función del kernel que comprueba si un UID (usuario) está mapeado en el namespace actual.
- `kgid_has_mapping`: función del kernel que comprueba si un GID (grupo) está mapeado en el namespace actual.
- Ejemplo técnico: Si un archivo tiene UID 0 pero no está mapeado en el namespace del usuario, OverlayFS aborta la copia.
- Ejemplo analógico: El bibliotecario revisa si el carnet y la firma son reales; si no coinciden, no te deja copiar la hoja de calcar sobre el libro.

Referencias:

- CVE-2023-0386: https://cve.mitre.org/cgi-bin/cvename.cgi?name=CVE-2023-0386
- Linux Kernel Security Advisory 2023

Nota: Este informe es de carácter técnico/educativo y resume la información pública sobre la vulnerabilidad CVE-2023-0386. Contiene descripciones de cómo la vulnerabilidad puede ser explotada a nivel conceptual, pero **no incluye instrucciones de explotación funcional ni código dañino**.