LINUX KEYLOGGERS: THE STATE OF THE ART

$WHOAMI

Malware Detection Researcher at SentinelOne.

Ex-Security Researcher at Uptycs.

Work primarily on Linux malware detection and security automation.

FOSS contributor and selfhosting nerd.

AGENDA

What is a keylogger

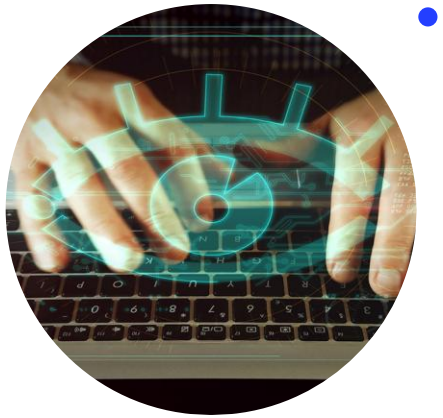
Linux-based keyloggers

X11 Keyloggers

Wayland Keyloggers

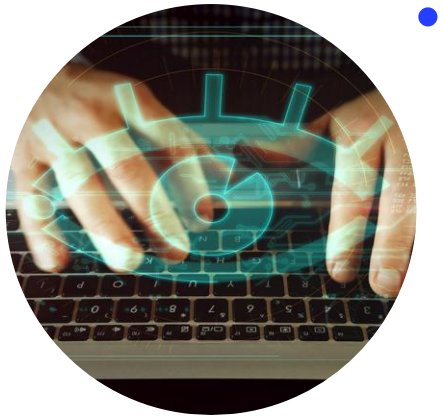
Q&A

WHAT IS A KEYLOGGER?



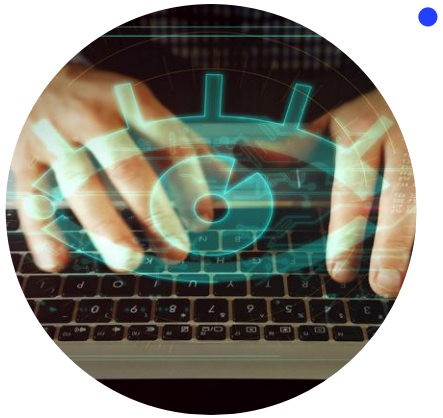
WHAT IS A KEYLOGGER?

- A keylogger is a program that is designed to record your keystrokes.
- Can be made to then exfiltrate this data to a remote server.
- Would be best to do it with the least amount of privileges.
- If root, there are a few ways:
  - Linux keyboard device files
  - Hooking into sshd
  - Hooking into readline(3)



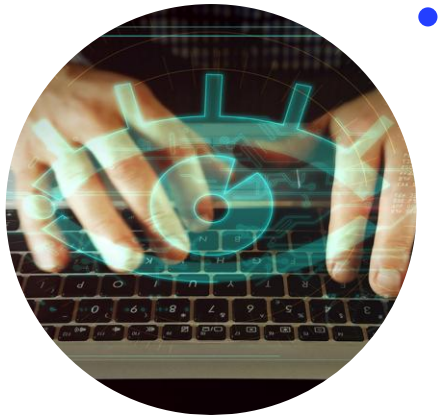
ASIDE: LINUX DESKTOP

- Linux desktop market share is much lower than server-side market share.
- The desktop has some unique problems and security gaps, which are interesting to look at.
- Will be looking at the two major display servers in Linux, X11 and Wayland.



GLOSSARY

- Display server: Server that actually creates the display that you look at. Refers to X11 or X.org Server (or just X). Not to be confused with X (formerly Twitter).
- Desktop Environment: What most Linux users call their “desktop”. Really multiple programs working in tandem to give a cohesive GUI experience. Examples: KDE Plasma, GNOME Shell.
- Display manager: Program that calls the X server, usually gives a login screen, and also lets you change the Desktop Environment that you want to start after logging in.
- Compositor: A program that *composites* a desktop based on instructions given by X server. Example: Mutter, KWin, picom, etc
- Window Manager: A program that manages the look and feel of the windows around client applications.



KEYLOGGERS ON LINUX DESKTOP (GUI)

- Keyloggers that work on X11
- Work on Wayland

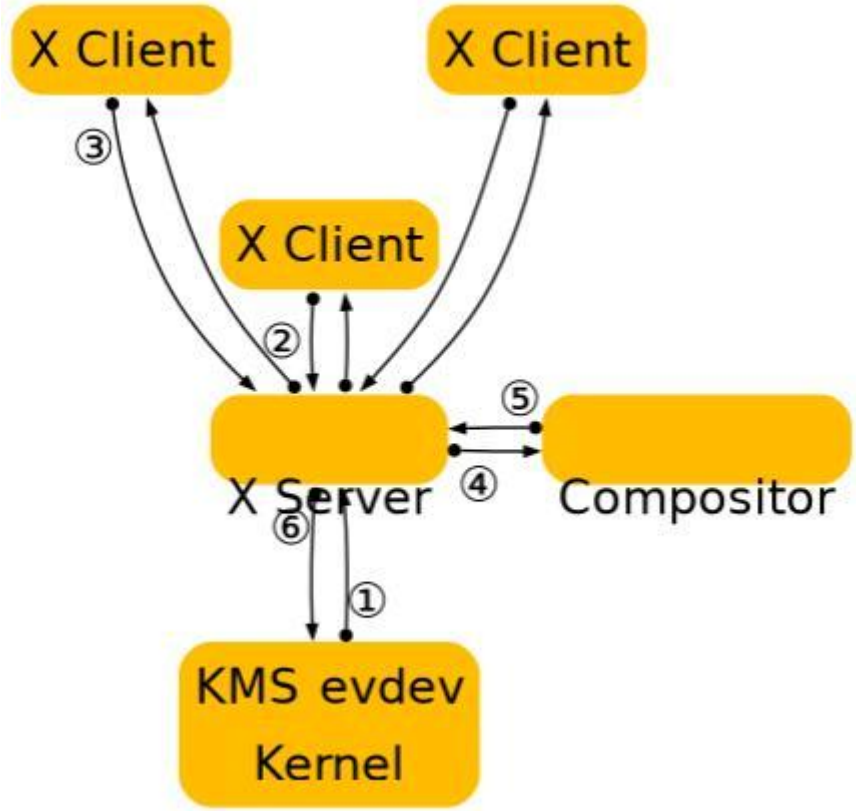
X SERVER SECURITY

- X Server is a server in the sense that any computer that can connect to a remote X server can be used to display graphical applications from a remote computer.
- For example: X forwarding over SSH, Xming and MobaXTerm for Windows
- Therefore, the security that’s built in is usually related to remote computers, and there’s no limitation on what a GUI program can ask for in localhost

X SERVER ARCHITECTURE

- X Server has a server/client model.
- Every GUI program running is a client that connects to a local X server.
- Specifically about input, the X server has no idea where the mouse is in the desktop, so it relies on the compositor to tell the active window.
- Programs like virtual keyboards are treated like normal keyboards in X, therefore they can receive events from the physical keyboard (through X) and send keyboard events like physical keyboards.
- Therefore, you can query X for all keyboard events.

X ARCHITECTURE



Demo

X KEYLOGGERS

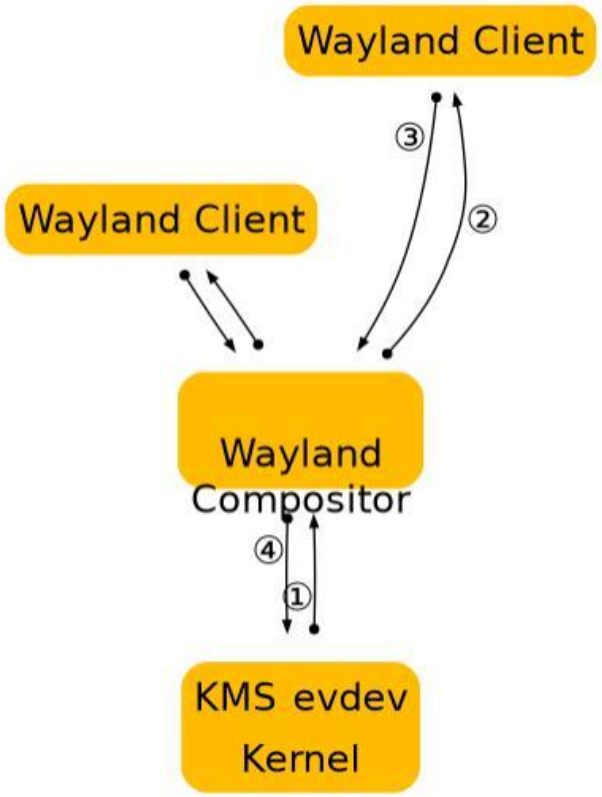
WAYLAND

- Wayland is not a display server like X, it’s a communication protocol.
- Programs built using the Wayland protocol are called Wayland compositors, or Wayland compositing window managers.
- Since Wayland uses compositors by default, it knows where every window is in a display. Therefore, keyboard events will only be sent to a client program if it’s in focus, i.e. the pointer has hovered to it or it was selected by alt-tabbing or similar actions.
- Keyloggers will have to try harder.

WAYLAND

- But what about virtual keyboards? The implementations are fragmented.
- Example: Gnome and Phosh can use Squeekboard, but KDE Plasma can’t, because Gnome has it’s own unstable API for virtual keyboards. KDE has another existing API.
- What about screen sharing?
  - Currently supported by using PipeWire. Not implemented by Wayland itself.
  - KDE has its own protocol.
  - COSMIC has its own protocol.
- What about screenshots?
  - Pretty much everyone has a competing implementation.

WAYLAND ARCHITECTURE



- Compared to X11, Wayland's architecture is clearly much simpler.
- This simplicity is because the Wayland compositor here does much of the actual work that was done with multiple different layers.
- Also means that far less data is exposed to potential clients
- For example, key strokes:
  - In X, you're supposed to tell the server to send you particular keystrokes by "binding" it.
  - No such luck in Wayland. Your keystrokes will be sent to the GUI program only when the window is in focus.

WAYLAND KEYLOGGER

- Wayland compositors and X need to read the keyboard data from somewhere.
- /dev/input/ provides the devices corresponding to physical hardware.
- Therefore, if we are able to be a man-in-the-middle and snoop into this process, we can get access.
- Two well-known processes:
  - LD_PRELOAD
  - Trace the compositor

LD_PRELOAD KEYLOGGER

- LD_PRELOAD is a way to write functions that get executed in priority, compared to the rest of the program.
- LD_PRELOAD can be used to be a man in the middle of any function running in the wayland compositor.
- The keylogger hooks into `wl_proxy_create` and `wl_proxy_marshal_array_constructor`.
- Can be run as non-root.

STRACE KEYLOGGER

- The ptrace syscall is a Linux kernel-level syscall to trace a particular process and list what syscalls are being used.
- Using ptrace, the attacker can hook into the read syscall, and use it to decode bytes read from keyboards by Wayland as and when it happens.
- Since the ptrace API does not allow for tracing of programs running in a different user, you need to be root for this attack.
- Still very hard to protect against in a fundamental level.

Demo

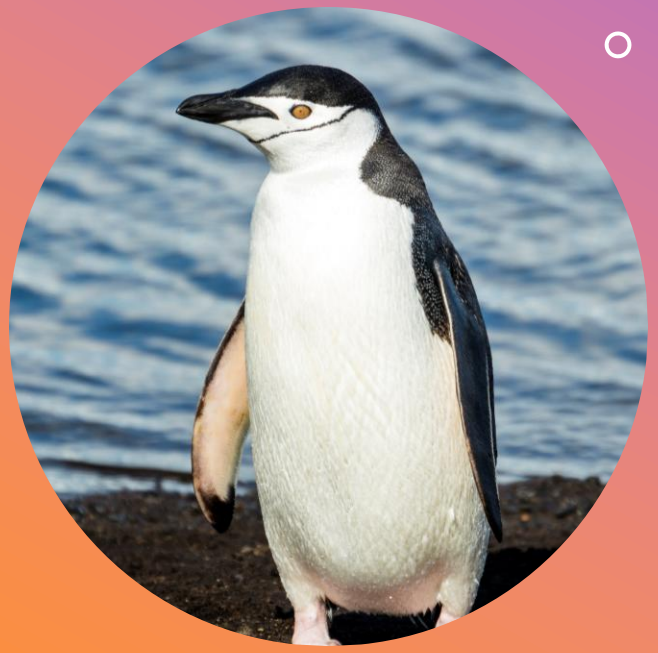
WAYLAND KEYLOGGERS

Any questions?

Q&A

REFERENCES

- https://www.x.org/wiki/guide/concepts/#index1h4
- https://fedoramagazine.org/pipewire-the-new-audio-and-video-daemon-in-fedora-linux-34/
- https://wearewaylandnow.com/
- https://wayland.app/protocols/
- https://github.com/anko/xkbcat
- https://github.com/Aishou/wayland-keylogger
- https://github.com/schauveau/sway-keylogger



THANK YOU

Nischay Hegde

hello@nischay.me

https://www.linkedin.com/in/thatloststudent/

https://twitter.com/thatloststudent

https://infosec.exchange/@thatloststudent