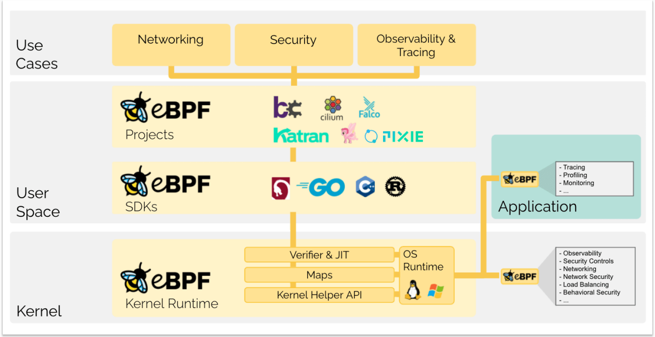
Introduction to eBPF

Théophile Dubuc

January 15, 2025

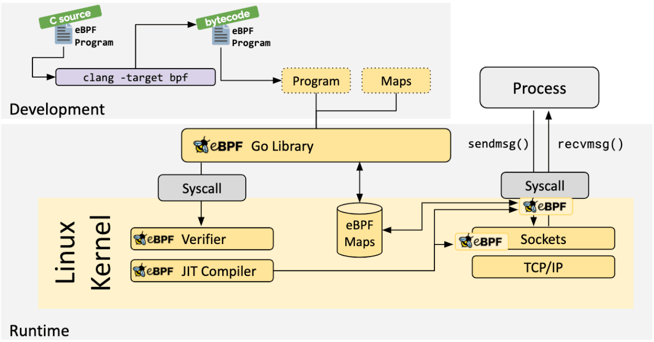
Introduction: What is eBPF?

What is eBPF?



- 1990’s: Berkeley Packet Filter (BPF) to filter network packets in Unix systems
- 2014: extended BPF (eBPF) for the whole kernel
- Now, BPF = eBPF and cBPF = old BPF.
- Run sandboxed programs in response to kernel events.
- Uses cases in observability, security, kernel programming.

How eBPF Works



- Load and attach eBPF programs from user space (bpftool, libbpf, BCC, etc.).
- eBPF program interacts with kernel resources.
- Optionally returned to user space via BPF maps, like ring buffers.

The BPF programming language

The BPF programming language: C with constraints

BPF security relies on constraints while programming:

- No standard C libraries, use BPF helpers instead
- Limited number of instructions
- No ”regular” loops
- No dynamic memory
- Strict pointer checking
- No floating point operations
- ...

Types of eBPF Programs

Tracepoints: Built-In Instrumentation

- Predefined kernel events.
- Lower overhead than kprobes.
- Example: monitoring time spent in a syscall.

kprobes: Attaching to Kernel Functions

- Dynamically attach to almost any kernel function.
- Used for:
  - Measuring function execution times.
  - Tracking kernel resource usage.
- Example: Hooking `vfs_read`.

Attaching to network interfaces

Use eBPF like cBPF and perform packet processing by attaching programs to:

- XDP (express data path): very fast, early in the stack
- tc (traffic control): more flexible

With them you perform many operations on ingress and egress packets:

- Read them
- Re-write them
- Redirect them
- Drop them

And you can offload your programs to compatible NICs!

Communication Mechanisms

Using bpf_printk

- Debugging tool for eBPF programs.
- Prints messages to the kernel log (dmesg).
- Lightweight but not suitable for production.

Maps: Sharing Data Between Kernel and User Space

- Key-value storage accessible by both eBPF and user-space programs.
- Types:
  - Hash maps
  - Arrays
  - Per-CPU maps
- Example: Counting system calls by process ID.

Ring Buffers: Streaming Data

- Efficient mechanism for sending structured data to user space.
- Commonly used for profiling and tracing applications.
- Example: Streaming syscall durations.

Conclusion

Summary

- eBPF provides powerful, low-overhead tools for kernel performance evaluation, kernel programming, and packet processing.
- Key types of programs:
    - kprobes
    - Tracepoints
    - Network filters
- Communication mechanisms include:
    - `bpf_printk`
    - Maps
    - Ring buffers

Lets try it! (Questions before?)