BPF User Tracepoints

Enhancing User Defined Tracepoints

Matheus Marchini

@mmarkini

Goals

- Instrument user-space applications
  - Trace with BPF tools (e.g., bpftrace)
- Sophisticated Argument Types (structures)
  - Possibly integrated with BTF
- Static and dynamic instrumentation

Current Approach

Systemtap SDT [1]

- Only static instrumentation
  - Workaround: libstapsdt [2]
- Simple (numeric) argument type
  - Workaround: casting arguments

Extending Systemtap SDT

(or: creating a new user-space tracepoint format)

- ELF Note Section → New ELF Section
- Revamped Arguments
  - Linux Tracepoint-inspired format
    - **field:** *BTF_type_id name*; **offset:** *arg_offset*;
      - Arguments located with offsets instead of strings
        - (might have perf impact, but increases portability)
- Dynamic tracepoints installed in a separate file
  - /proc/PID/events/*provider*/*probe*/[address,format]
  - or inspired in *perf* JIT support (on /tmp/ or $PWD/)

https://sourceware.org/systemtap/wiki/UserSpaceProbeImplementation

https://docs.oracle.com/cd/E23824_01/html/819-0690/chapter6-18048.html

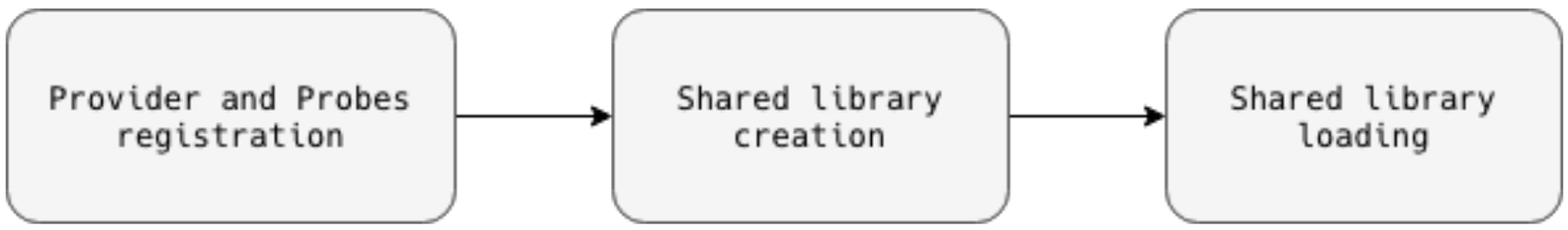
	+0	+1	+2	+3	
namesz	8				
descsz	24 + len(provider) + len(probe) + len(arguments)				
type	3				
name	s	t	a	p	
	s	d	t	\0	
desc	0x…				**Probe PC**
	0x…				**Link-time sh_addr**
	0x…				**Link-time semaphore variable address**
			…	\0	**Provider name (null-terminated)**
			…	\0	**Probe name (null-terminated)**
			…	\0	**Arguments (space-separated, null-terminated)**

```
$ readelf -n ./example

Displaying notes found in: .note.stapsdt
  Owner                 Data size       Description
  stapsdt               0x00000047      NT_STAPSDT (SystemTap probe descriptors)
    Provider: example
    Name: second_probe
    Location: 0x0000000000006a8, Base: 0x0000000000000754, Semaphore:
0x0000000000000000
    Arguments: 8@%rax -4@$1 16@-48(%rbp)
```

libstapsdt

- Creates shared libraries on runtime
  - With empty functions
  - Probes instrument those functions
  - *dlopen* the shared library



https://github.com/sthima/libstapsdt