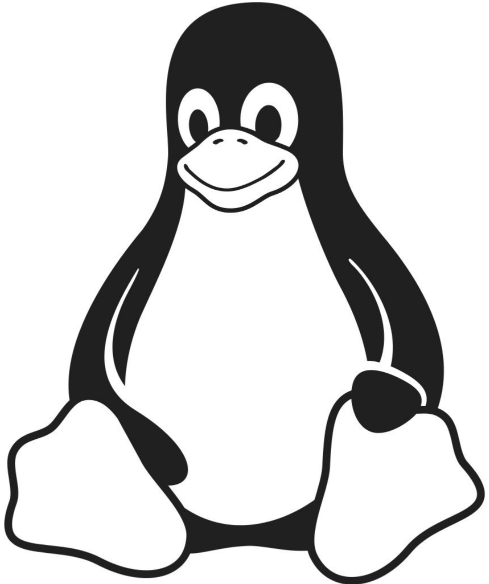
Time Namespace

Andrei Vagin, Dmitry Safonov

Linux Plumbers Conference 2018

User Timing API in Linux

- Timers (setitimer, timer_create, timer_fdcreate, alarm)
- gettimeofday, settimeofday
- adjtimex
- nanosleep
- Posix clock-s
  - clock_{set,get}time
  - clock_adjtime
  - clock_nanosleep
- Proc files, file attributes, etc



POSIX clocks

- CLOCK_REALTIME

  System-wide clock that measures real (i.e., wall-clock) time. Setting this clock requires appropriate privileges. This clock is affected by discontinuous jumps in the system time.
- CLOCK_MONOTONIC

  Clock that cannot be set and represents monotonic time since some unspecified starting point.
- CLOCK_MONOTONIC_RAW

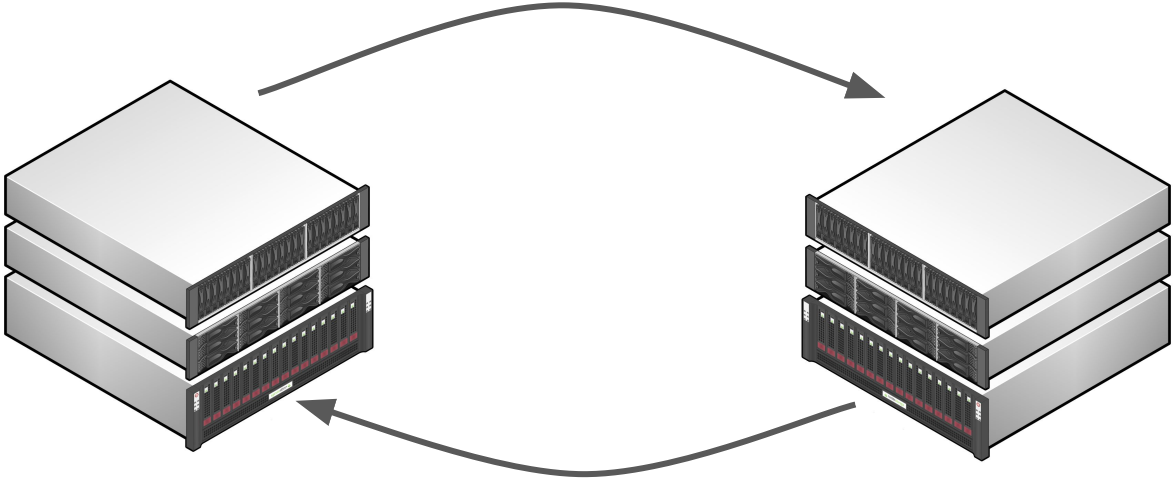
  Similar to CLOCK_MONOTONIC, but provides access to a raw hardware-based time that is not subject to NTP adjustments
- CLOCK_BOOTTIME

  Identical to CLOCK_MONOTONIC, except it also includes any time that the system is suspended.
- CLOCK_PROCESS_CPUTIME_ID,

  CLOCK_THREAD_CPUTIME_ID

Checkpoint/Restore and Migration

- CLOCK_REALTIME is already synchronized between hosts
- Need to handle monotonic and boottime clocks
  - Must never go backwards
  - Should be smooth



CRIU

Lightweight virtualization

Per-namespace offsets for system clocks

Pros

- Small overhead
- Easy to use
- No need to adjtimex (ntpd)

Use cases:

- Checkpoint/Restore
- Testing

Implementation

- Per-namespace clock offsets
- VDSO: VVAR page
- clock_gettime
- Timers
- Zero-overhead w/o timens

Full virtualization

A separate time domain for each namespace

Nov 15, 4:00 Multiple Time Domains (Thomas Gleixner)

Pros

- Running ntpd (adjtimex())
- Smoothing of leap seconds (UTC-SLS)
- A custom length of a second

Use cases

- Testing
- ???

Questions?