

# systemd 简介

Aron Xu

2014-11-16@THU

It takes too long...

# PID 1: init

- 启动系统 (startup)
- 看守状态 (babysitting)

# PID 1: init

想启动快？

- Start **less**  
Start **more** in parallel

# 几种常用的 init

- sysvinit + insserv
- upstart
- systemd

# 不“那么”常用的 PID 1

- busybox-init
- launchd
- sysvinit + OpenRC
- emacs/vim

# sysvinit 的问题

- 软硬件即插即用需求变多，要求变高
- 启动可靠性、启动速度
- 服务和会话的管理与监控

此处省略一万字.....

# systemd 的答案

- event-driven init
- 尽可能避免 shell 脚本
- cgroup
- Let's move it to systemd!

# event-driven init

原因：

- Linux 采用 dynamic/event-based 架构
- This looks cool

# event-driven init

socket/dbus activation:

- systemd 创建所有的各种 “socket”
- 类似的想法早在 inetd 中就有了
- event-driven desktop

# shell 脚本为何不靠谱

- 执行结果易受环境变量、变种和版本的影响
- 如 Debian/Ubuntu 从 bash 切换到 dash 时因为各种 bashism 而经历了相当大的痛苦

# shell 脚本为何不靠谱

再来做个实验：

- `$ time for i in `seq 1 100000`; do echo 1 >/dev/null; done`
- `$ time for i in `seq 1 100000`; do /bin/echo 1 >/dev/null; done`

# shell 脚本为何不靠谱

我的结果：

- *real* 0m1.454s *user* 0m1.052s *sys* 0m0.384s
- *real* 15m4.262s *user* 0m18.461s *sys* 13m2.609s

# shell 脚本为何不靠谱

推论：

- 进程创建太昂贵，应使第一个用户进程的 PID 尽可能小 (echo \$\$)

# shell 脚本为何不靠谱

联想：

- 当年 Ubuntu (6.10) 启动速度变快是 dash 的功劳，和 upstart 没多少关系

# cgroup: 管着服务的“生老病死”

现在的情况：

- double-forking
- 日志：syslog/stdout/stderr
- 权限和资源

# cgroup: 管着服务的“生老病死”

cgroup: simple job tracking

- 进程及其子孙无法从中逃脱
- 能记录进程组的全部输入输出
- 按进程组设置权限、控制资源分配

# cgroup: 管着服务的“生老病死”

- 可靠地查看服务状态、启动 / 杀死服务
- 日志收集不用操心了 (+ no daemon opt)
- 基本权限设置 (systemd.{cgroup|exec})

# move it to systemd

- 预谋已久：  
systemd = system daemon
- udev, logind, syslog, util-linux, networkd, readahead, ...

# move it to systemd

- 万恶的，不符合传统 Unix 哲学
- 确实简化了设计

# move it to systemd

系统开发有东西搞不定时，请说

- Let's move it to systemd!

# systemd 的几点其他事项

- .desktop 风格的启动配置文件  
(systemd.unit)
- 系统状态快照  
(systemd.{snapshot|isolate})

# 几个有用的 systemd 命令

- systemctl 全部服务的状态
- systemctl  
start/stop/reload/kill/status/show/enable/  
disable/... ..
- systemctl  
default/rescue/emergency/halt/poweroff/  
reboot/kexec/stop/... => run levels?

# 几个有用的 systemd 命令

- `systemd-cgls`, `systemd-cgtop`
- `systemd-analyze [blame|plot]`
- `systemd-journalctl`
- `systemd-detect-virt`
- `systemd-nspawn`

And it takes too long...



Debian 大法好！

