## Beyond Init: systemd

linux.conf.au 2011

## Lennart Poettering

## January 2011

▲ □ ► < □ ►</p>

< ≣⇒

"systemd is a system and session manager for Linux,

・ロ・ ・ 日・ ・ 日・ ・ 日・

"systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts.

(4回) (4回) (4回)

"systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities,

・ 同 ト ・ ヨ ト ・ ヨ ト

"systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services,

伺 ト イヨト イヨト

"systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons,

伺 ト イヨト イヨト

"systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons, keeps track of processes using Linux cgroups,

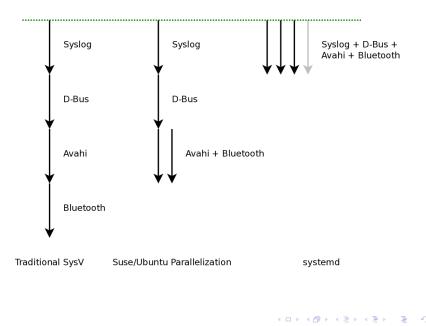
"systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons, keeps track of processes using Linux cgroups, supports snapshotting and restoring of the system state, "systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons, keeps track of processes using Linux cgroups, supports snapshotting and restoring of the system state, maintains mount and automount points "systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons, keeps track of processes using Linux cgroups, supports snapshotting and restoring of the system state, maintains mount and automount points and implements an elaborate transactional dependency-based service control logic. "systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons, keeps track of processes using Linux cgroups, supports snapshotting and restoring of the system state, maintains mount and automount points and implements an elaborate transactional dependency-based service control logic. It can work as a drop-in replacement for sysvinit."

## init(8)

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへで

### Parallelization

・ロ・・(四・・)を注・・(注・・)注



Socket-Based Activation

・ロ・・(四・・)を注・・(注・・)注

#### Socket-Based Activation

### The kernel orders and buffers requests for us!

- - 4 回 ト - 4 回 ト

Socket-Based Activation The kernel orders and buffers requests for us! Implicit dependencies!

回 と く ヨ と く ヨ と

Socket-Based Activation The kernel orders and buffers requests for us! Implicit dependencies! Patching daemons

白 ト く ヨ ト く ヨ ト

**Bus-Based Activation** 

・ロ・・(四・・)を注・・(注・・)注

## Starting Less: On-Demand Loading

◆□ > ◆□ > ◆臣 > ◆臣 > ○

Parallelizing File System Jobs

◆□→ ◆□→ ◆三→ ◆三→

## Parallelizing File System Jobs autofs!

< □ > < □ > < □ > < □ > < □ > .

## Shell is evil

・ロ・・(四・・)を注・・(注・・)注

## Shell is evil

### Move to systemd, daemons, kernel, udev, ....

・ロン ・四と ・ヨン ・ヨン

# Shell is evil Move to systemd, daemons, kernel, udev, ... Provide proper debugging facilities

・回 と く ヨ と く ヨ と

The best babysitter.

・ロン ・四と ・日と ・日と

The best babysitter. Control Groups!

・ロト ・回ト ・ヨト ・ヨト

The best babysitter II

ヘロン 人間と 人間と 人間と

#### The best babysitter II

Environment, resource limits, working directory, chroot(), umask, OOM adjustment, nice level, IO priority and class, CPU scheduler priority and policy/reset-on-fork, CPU affinity, timer slack, stdio to syslog/tty/null/kmsg, uid, gid, supplementary groups, file system namespacing (r/o file systems, inaccessible systems, mount propagation, private /tmp), capabilities (inherited set, bounding set, secure bits), ...

伺下 イヨト イヨト

Unit types: service, socket, device, mount, automount, target, snapshot, timer, swap, path

(4回) (4回) (4回)

Don't reinvent the wheel:

◆□> ◆□> ◆目> ◆目> ◆目> 目 のへで

Don't reinvent the wheel:

Read SysV/LSB init script headers, read /etc/fstab, support traditional inetd modes, support /dev/initctl, utmp, wtmp, support double-fork()ing daemons.

同 と く ヨ と く ヨ と

Don't reinvent the wheel:

Read SysV/LSB init script headers, read /etc/fstab, support traditional inetd modes, support /dev/initctl, utmp, wtmp, support double-fork()ing daemons.

.desktop files.

回 と く ヨ と く ヨ と

## Snapshots

Transaction System

・ロ・・(四・・)を注・・(注・・)注

## D-Bus!

▲□▶ ▲圖▶ ▲≣▶ ▲≣▶ 三重 - のへぐ

systemadm

◆□▶ ◆□▶ ◆三▶ ◆三▶ 三三 - のへで

## systemd as basic OS building block

< ロ > < 回 > < 回 > < 回 > < 回 > :

systemd as basic OS building block systemd for cross-distribution standardization

(4回) (4回) (4回)

・ロン ・雪 と ・ ヨ と ・ ヨ と

Status: almost made Fedora 14. Substantial coverage of basic OS boot-up tasks,

・日・ ・ヨ・ ・ヨ・

Substantial coverage of basic OS boot-up tasks, including fsck,

▲□ ▶ ▲ □ ▶ ▲ □ ▶

Substantial coverage of basic OS boot-up tasks, including fsck, mount,

▲□ ▶ ▲ □ ▶ ▲ □ ▶

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota,

▲□ ▶ ▲ □ ▶ ▲ □ ▶

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock,

回 と く ヨ と く ヨ と

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead,

同 と く ヨ と く ヨ と

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles,

同 ト く ヨ ト く ヨ ト

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed,

< ∃⇒

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog, plymouth,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog, plymouth, shutdown,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog, plymouth, shutdown, kexec,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog, plymouth, shutdown, kexec, SELinux,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog, plymouth, shutdown, kexec, SELinux, initrd+initrd-less boots,

Substantial coverage of basic OS boot-up tasks, including fsck, mount, quota, hwclock, readahead, tmpfiles, random-seed, console, static module loading, early syslog, plymouth, shutdown, kexec, SELinux, initrd+initrd-less boots, cryptsetup, ...

## Status: 7s on openSUSE. Less than 14s on fully-featured Fedora. (SSD)

## Status: 7s on openSUSE. Less than 14s on fully-featured Fedora. (SSD)

Lots of room for improvement.

(1日) (1日) (日)

Adoption: Fedora, openSUSE, Mandriva, Debian, Gentoo, ArchLinux, ...

Next: Fedora 15.

▲□▶ ▲圖▶ ▲≣▶ ▲≣▶ 三重 - のへぐ

Next: Fedora 15. read-only root,

◆□ > ◆□ > ◆臣 > ◆臣 > ○

Next: Fedora 15. read-only root, session manager,

Next: Fedora 15. read-only root, session manager, automatic initrd fallback.

・ 母 と ・ ヨ と ・ ヨ と

Say No! to Copyright Assignment.

That's all, folks.

・ロン ・雪 と ・ ヨ と ・ ヨ と

That's all, folks. Any questions?

・ロン ・四と ・ヨン ・ヨン

## systemd

http://www.freedesktop.org/wiki/Software/systemd
http://0pointer.de/blog/projects/systemd
git://anongit.freedesktop.org/systemd
#systemd on irc.freenode.org

回 と く ヨ と く ヨ と