

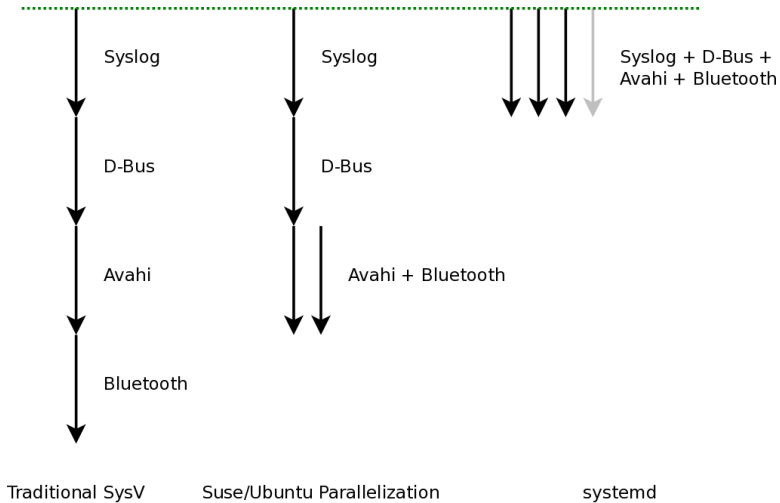
Beyond Init: systemd

Linux Plumbers Conference 2010

Kay Sievers
Lennart Poettering

November 2010

Triggers: Boot, Socket, Bus, Device, Path, Timers, More



Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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

Status: almost made Fedora 14.

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,

Status: almost made Fedora 14.

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,

Status: almost made Fedora 14.

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,

Status: almost made Fedora 14.

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.

Status: 7s on openSUSE. Less than 18s on full-featured Fedora.
(SSD)

Status: 7s on openSUSE. Less than 18s on full-featured Fedora.
(SSD)

Lots of room for improvement.

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

Next: Fedora 15. cryptsetup,

Next: Fedora 15. cryptsetup, read-only root,

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

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

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

Storage Assembly Daemon?

As session manager:

Redefine D-Bus session bus:

Be honest, give up on multiple graphical logins per user, per machine. Don't claim D-Bus was attached to Display. To reallocate multiple sessions per home dir, per machine, attach multiple displays to bus, differentiate by bus name suffix. Redefine session as time from first login to last logout.

As session manager:

Redefine D-Bus session bus:

Be honest, give up on multiple graphical logins per user, per machine. Don't claim D-Bus was attached to Display. To reallocate multiple sessions per home dir, per machine, attach multiple displays to bus, differentiate by bus name suffix. Redefine session as time from first login to last logout.

Analogous XDG_RUNTIME_DIR.

Or, introduce additional user bus?

Or, introduce additional user bus?
Relation to gnome-session?

Or, introduce additional user bus?

Relation to gnome-session?

Handling of KDE style dlopen()-exec()?

Or, introduce additional user bus?

Relation to gnome-session?

Handling of KDE style dlopen()-exec()?

Handling of gdm/kiosk problem?

Or, introduce additional user bus?

Relation to gnome-session?

Handling of KDE style dlopen()-exec()?

Handling of gdm/kiosk problem?

Handling of user services when nobody is logged in?

“systemd is a system and session manager for Linux,

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

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

“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,

“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,

“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,

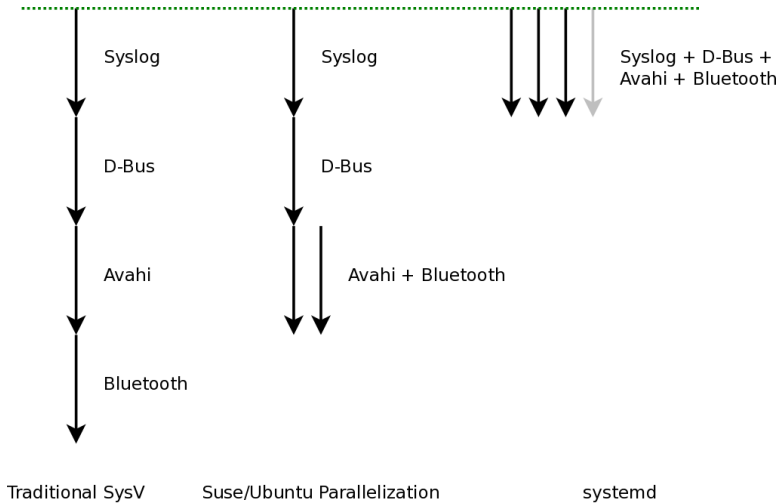
“systemd is a system and session manager for Linux, compatible with SysV and LSB init scripts. systemd provides aggressive parallelization capabilities, uses socket and D-Bus activation for starting services, offers on-demand starting of daemons, keeps track of processes using Linux cgroups, supports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

Parallelization



Socket-Based Activation

Socket-Based Activation

The kernel orders and buffers requests for us!

Socket-Based Activation

The kernel orders and buffers requests for us!

Implicit dependencies!

Socket-Based Activation

The kernel orders and buffers requests for us!

Implicit dependencies!

Patching daemons

Bus-Based Activation

Starting Less: On-Demand Loading

Parallelizing File System Jobs

Parallelizing File System Jobs autofs!

Shell is evil

Shell is evil

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

Shell is evil

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

Provide proper debugging facilities

The best babysitter.

The best babysitter.
Control Groups!

The best babysitter II

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), ...

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

Don't reinvent the wheel:

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.

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.

Snapshots

Transaction System

D-Bus!

systemadm

systemd as basic OS building block

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

systemd in the distributions

Future: managing sessions

Say No! to Copyright Assignment.

That's all, folks.

That's all, folks.
Any questions?

systemd

<http://www.freedesktop.org/wiki/Software/systemd>

<http://0pointer.de/blog/projects/systemd>

<git://anongit.freedesktop.org/systemd>

#systemd on irc.freenode.org