GNU/Hurd DDE userland device drivers

Samuel Thibault

2014 February 2nd
It's all about freedom #0

“The freedom to run the program, for any purpose”

I.e.:

- Freedom from sysadmin!
  - WTH is fdisk/mke2fs/... hidden in /sbin?
  - I should be able to just work with my disk/network access

- Freedom to innovate
  - Experimental filesystem, personal work-flow, new kind of process combination,...

- Also provide freedom from misbehaving programs and drivers
It's all about freedom #0

From: xxx <xxx@yyy.fr>
Subject: Network expertise
Date: Thu, 31 Jan 2013 12:37:34 +0100

[...] Would it be possible to route to my VPN the traffic of only one application?

Actually, also well-known classical issue of full-VPN: traffic of the VPN itself shouldn't go through the VPN!

And yet, here root capabilities!!

Spoiler: Yes, GNU/Hurd can already do it. Without even asking root.
It's all about freedom #0

Extensibility for the user

- Mount one's own files
  - Access archives content
  - Access remote files
  - Experiment with filesystems
- Access one's own network
  - Access remote networks / VPN
  - Access virtual machine network
- Redirect one's sound
  - Through network
  - Sound effects
  - Recording
- ...
- and Flexible hardware support
Outline

- Hurd architecture Overview
- Network flexibility
- DDE stack
- Console support
- Hardware support
- Releases & future
Micro-kernel layering

Kernel

Tasks, memory, IPC

root

proc

auth

ext2fs

pfinet

user

sh

cp
Micro-kernel layering

Kernel: Tasks, memory, IPC

ext2fs -> auth -> sh
pfinet -> proc
root

user

sh -> cp
Micro-kernel layering

- Server crash? Not a problem
  - “Computer bought the farm” is just an error, not something-of-the-death
- Easier to debug/tune
  - Just run gdb, gprof, …
- Can dare crazy things
  - The Hurd console has dynamic font support
    - See chinese support in pseudo-graphical mode (actually pure VGA textmode!) of Debian installer.
- Kernel only handles Tasks, memory, IPC
Hurd possibilities

- ext2fs
- pfinet
- ftpfs
- auth
- proc
- root
- user
- isofs
- sh
- cp

Kernel
Hurd possibilities

€ settrans -c ~/ftp: /hurd/hostmux /hurd/ftpfs /
(just once for good)

€ settrans -a ~/mnt /hurd/iso9660fs

€ ls ~/mnt

README-or-FAIL

...

- Only downloads what is needed.
- Can be permanently stored in ext2fs

€ settrans ~/.signature /hurd/run /usr/games/fortune
How does it work?
Rationale

- **Everything** is an (interposable) RPC
- Translators exposed in the FS
  - The user gets to decide what/how to interpose
    - Without need for costly ptrace or fragile libc symbols interposition.
    - **Native** fakeroot/chroot
    - Fully virtualized and fine-grained interface
  - Just need to use what's provided by the admin, e.g.
    - $HOME/
    - TCP/IP stack
  and pile over it
But also

- \( ~/\text{remap/remap.sh} \) /bin/sh \$HOME/bin/sh
- \( ~/\text{remap/remap.sh} \) /bin \$HOME/unionbin

...  

- Check out Stow/Nix/Guix!
Hurd possibilities (cont'ed)

Kernel

- ext2fs
- vpn
- auth
- proc
- open
- ftpfs
- part
- sh
- cp
- ext2fs
- pfinet
- user
- isofs
Hurd possibilities (cont'ed)

i.e. ISO image inside a partitioned disk image on ftp over a VPN
Hurd possibilities (cont'ed)

i.e. ISO image inside a partitioned disk image on ftp over a VPN
Hurd userland network support

```
/servers/socket/2
```

```
pfinet
```

```
eth0
```

```
Kernel
```

```
root
```

```
w3m
```

```
user
```
Hurd userland network support

pfinet

DDE layer
Linux 2.6.32 drivers
/dev/eth0

eth0

w3m

Kernel

root

user
Hurd userland network support

Kernel

root

<table>
<thead>
<tr>
<th>eth-filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>pfinet</td>
</tr>
</tbody>
</table>

w3m

user
Hurd userland network support

Kernel

/servers/socket/2
pfinet
eth-filter
eth0

openvpn
pfinet
~/servers/socket/2

root

w3m

~/servers/tun0

user
Hurd userland network support

```sh
$ settrans -ca ~/servers/socket/2 \
  ~/bin/pfinet -i ~/servers/tun0 \
  -a 80.67.176.254 -p 80.67.179.1

$ vpn.sh &

$ ~/remap/remap.sh \
  /servers/socket/2 ~/servers/socket/2 \
  /etc/resolv.conf ~/resolv.conf

$ wget www.gnu.org
```

- My own translators
- Only wget accesses my pfinet (well, the shell too :) )
DDE stack

Based on TU-Dresden's DDE stack

- Zheng Da's GSOC
- Ported to Mach kernel
- Ported to Mach device interface
- Updated libdde_linux26 for long-term-supported linux 2.6.32
  - Most drivers (and mostly the really useful ones) just work without patches
  - We should cooperate on upgrading that part
- Used by default by Debian GNU/Hurd
DDE stack

Linux drivers

Linux API

libdde_linux26

ddatekit API

libddeenkit

pkg_xmit

rx_callback

netdde

kernel

intr

vm_allocate_contiguous

libmachdev

pdevnet

device_read

device_write
DDE stack

- Only two additions to the kernel
  - Interrupt delivery and masking
  - Physically-contiguous memory allocation
  - (Direct I/O access was already available)
- Performance similar to in-kernel driver
- Driver in a separate process
  - Can just crash and be happy with it...
  - Can easily debug and profile them
  - Stack smashing protection ;)
  - Could benefit from I/O MMU for better isolation.
    - For now drivers can just access all RAM...
More DDE?

- Disk DDE is supposed to be working
  - Should be not very complex
    - device_read / device_write
  - Zheng Da said he didn't manage to make it work

- USB/sound DDE was mentioned as experimental
  - I don't know the status?
  - We would definitely love to have that

- Rather use Rump kernels?
Hurd userland console support

Modular design similar to screen

- Server running virtual ttys and gettys on them
- Client with drivers
  - Keyboard + mouse + VGA,
  - or ncurses,
  - or whatever
Hurd userland console support

Keyboard driver

- Gets keyboard/mouse events from kernel
- Translation done through xkb
  - No need to maintain our own keymaps any more
Hurd userland console support

VGA driver

- Directly drives VGA board in VGA text mode
- 256/512 dynamic glyphs support
  - 32-126 static ASCII characters for compatibility
  - Other glyphs dynamically allocated from BFD font
  - GNU greets user!
- Double-width glyph support
  - Can print kanjis in text mode!
Recent software support

- GCJ, GNAT
- GCC go: ongoing GSOC, issues with its own thread implementation
- Fixed lots of testsuite failures (perl, python, ...)
  - POSIX corners
  - Around the 99% figure now
- Languages for translators
  - Now using libpthread $\rightarrow$ python, perl, whatever...
Current State

Hardware support

- i686
- start of 64bit support
  - Kernel boots completely, now missing RPC 32/64bit translation
- DDE Linux 2.6.32 drivers layer for network boards
  - In userland netdde translator!
- IDE, Xorg, …
- AHCI driver for SATA (up to 2TiB disk support btw)
  - Needs more testing, perhaps bug fixing (no trouble on qemu)
- Xen PV domU
  - Required GNU Mach changes only
- No USB, no sound yet
Current State

Software support

- Quite stable
  - Have not reinstalled boxes for years.
  - Debian buildds keep building packages, usually hang after some weeks, out of some remaining memory leak.
- ~79% of Debian archive builds out of tree
  - XFCE, almost gnome, almost KDE
  - Firefox (aka iceweasel), gnumeric, …
- Standard *native* Debian Installer
Releases

- Nice 0.401 release on April 2011.
- Arch Hurd LiveCD release on August 2011.
- Released Debian-unofficial wheezy/sid snapshot CDs on May 2013 \o/
- Hurd 0.5 released on 2013 Sept 27th \O/
  - Just in time for GNU's 30th birthday!
Future work

- Xen PVH support, X86_64 support
- Language bindings for translators
- Read-ahead
- {hdd,sound,usb}dde?
- GNU system: Guix/Hurd?
- Debian GNU/Hurd Jessie?
- Your own pet project?
Thanks!

- http://hurd.gnu.org/
- http://www.debian.org/ports/hurd/
- The increasing irrelevance of IPC performance for microkernel-based Operating Systems
  