Nouveau

Recap, on-going and future work

Martin Peres, Lucas Stach & the Nouveau community

Ph.D. student at LaBRI, B.Eng. student at HfTL

January 19, 2012

Summary

History: NVIDIA – a new hope

- 1998(?): NVIDIA releases a Linux open-source 2D-only driver(nv)
- 1998: Obfuscation commit (release only pre-processed source)

After we already finalized XFree86-3.3.3 NVIDIA forced The
XFree86 Project to replace the sources we had with sources that
were partly run through the C preprocessor in order to remove
some of the names that NVIDIA thought might give away IP from
NVIDIA. This resulted in unreadable and unmaintainable code.

The XFree86 Project is strongly opposed to such obfuscated code.
We do not regard this as free software according to our standards.
Due to the extremely late date of this decision from NVIDIA we
decided to include the code as offered by NVIDIA. We are
considering to remove support for the later NVIDIA chips in a
future release, though.

Xfree86, commit message from 11/18/98

History: The open-source strikes back

- 2005: Stephane Marchesin improves nv and works on 3D
  - Project named Nouveau after an unfortunate automatic spelling correction
  - Back-port nv updates to Nouveau
- 2008: Open Arena runs on nv40
- 2009: KMS driver based on TTM
- 2010: Merged in Linux 2.6.33
- 2010: Nv is deprecated by NVIDIA, say "use VESA".

History: The return of the Jedi

"It's so hard to write a graphics driver that open-sourcing it would not help [...] In addition, customers aren't asking for opensource drivers."
Andrew Fear, NVIDIA software product manager, April 2006

Summary

Chipset families

- NV03: RIVA 128
- NV04: RIVA TNT/TNT2
- NV10: Geforce2/4 MX
- NV20: Geforce 3, Geforce 4 Ti
- NV30: Geforce FX
- NV40: Geforce 6/7 (the first real family)
- NV50: Geforce 8/9/100/200/300 (the biggest family)
- NVC0: Geforce 400/500, AKA Fermi
- NVD9: Half-Kepler chipset

NV30/40 microcodes

- HWSQ: very limited use (LVDS), no flow control
- CtxProgs on nv40: Context switching

NV50 microcodes

- CtxProgs: Context-switching microcode
- HWSQ v2 (formerly PMS): memory reclocking
- nv98+ Flexible MicroCode($F\mu c$): a general-purpose microcode. Broad usage (PCrypt, PDaemon, Vdec, ...)

NVC0-D9

- nvc0: PGRAPH (the rendering engine) is converted to $F\mu c$
- nvd9: some remaining engines are converted to $F\mu c$

The 4 parts of Nouveau

- Linux Kernelmodule
- libdrm-nouveau
- DDX: xf86-video-nouveau
- Mesa3D drivers:
    - nouveau-vieux
    - nvfx
    - nv50
    - nvc0

Linux kernel module Nouveau

- Resource management
  - GPU Channels
  - Memory Management
- Command-submission
- Kernel Mode Setting
- Power Management

Power management

- Readings: Temperature & clocks
- vbios parsing: mostly
- setting clocks: unreliable, potentially dangerous

WIP

- nvc0: setting clocks
- nv40-d9: memory timings (almost ready)
- nv30-c0: reliable clock changes (almost ready)
- AGP/PCIE, clock gating: reverse engineering
- performance counters, dynamic reclocking: WIP

libdrm-nouveau

- buffer management
    - everything is a buffer!
- wraps around the IOCTL interface

work in progress

- currently rewritten
- designed with nv40 class hardware in mind
- pushbuffer replay

DDX – xf86-video-nouveau

- EXA (2D acceleration)
- X-Video
- supports full range of GPUs (NV03 - NVD9)

Mesa3D: drivers for 3D and more

- nouveau-vieux: classic Mesa3D driver for NV04,NV1x,NV2x
    - only supports fixed function OpenGL
- nvfx: Gallium driver for NV3x,4x class GPUs
    - will hopefully be soon replaced by rewritten driver
- nv50: Gallium driver for NV50 class GPUs
    - maybe first free software driver to allow hardware OpenCL
- nvc0: Gallium driver for NVC0 class GPUs

work in progress

- rewrite of the nvfx driver (will be called nv30 after this)
- merge reworked shader compiler for nv50
- add new features to both nv50 and nvc0
- hardware video decoding on nv40-nvc0

Summary

The nouveau community

- The largest xorg-related IRC channel
- Composed of:
  - One paid developer
  - Former developers
  - Student developers
  - Enthusiasts

The nouveau maintainer

Nouveau is maintained by Ben Skeggs(darktama):

- Hired by Red Hat in 2009
- Located in Brisbane, Australia (GMT+10)
- Works on almost everything

Gallium's nouveau maintainer

Christoph Bumiller (calim) maintains the nv50-c0 gallium driver:

- Physics master student at the University of Vienna (Austria, GMT+1)
- Main Nouveau Gallium contributor
- nv50/c0 3D support
- performance improvements

Marcin Kościelnicki (mwk)

- Polish master student at the university of Warsaw (GMT+1)
- Implemented most of the GPGPU-oriented PSCNV nouveau fork
- Reversed most of the Fermi's architecture and video decoding

History	Architecture	The Nouveau community	Demos
	00000000	000●00	

Composition

Pekka Paalanen (pq)

- Finnish (GMT+2)
- Worked on mmiotraces, a register DB
- Does some communication-related work

Francisco Jerez (curro)

- Spanish Physics student (GMT+1)
- Worked on page-flipping and nouveau_vieux

Marcin Slusarz (joi)

- Polish SQL/C++/Java developer (GMT+1)
- Fixes software bugs (mesa + kernel), maintains Valgrind-MMT

Emil Velikov (xexaxo)

- Bulgarian student at the University of Nottingham (England, GMT+0)
- Reverse engineering of some PM-related vbios table
- Other PM-related implementation work
- Debugging/Testing

Roy Spliet (RSpliet)

- Dutch master student at the Delft University (GMT+1)
- Memory timings reverse engineering

Martin Peres (mupuf)

- Engineer, Ph.D. student at LaBRI (France, GMT+1)
- Reclocking process
- Thermal-zones & thermal management
- Minor reverse engineering

Maxim Levitsky (MaximLevitsky)

- Student at the Technion University of Haifa (GMT+2)
- Important reverse engineering work on reducing power consumption
- Stability-related reverse engineering
- New-comer

Maarten Lankhorst (mlankhorst)

- Netherlands (GMT+1)
- Implemented XVMC support
- Work towards an open VDPAU
- New-comer

Summary

MPEG1/2 video decoding

- using MPlayer
- Kernel-side: done
- Mesa: merged 3 days ago

Dynamic reclocking

- see clocks changing according to the load
- performance improvements in OpenArena
- performance improvements in xvmc