Let’s port together. Debian fun for everyone.

”Most civilised people are out of touch with reality because they confuse the world as it is with the world as they think about it, talk about it and describe it.”

Peter De Schrijver
p2@debian.org

February 24, 2007
Overview 1

Portability issues
why?
C types
Bitfields
Endianness
Alignment

Accessing peripheral hardware
Example system architectures
Trends in system design
Out of order transactions
Non-coherent I/O
Userland hardware access
Let’s port together.
Debian fun for everyone.

Peter De Schrijver

Why ?

- Correctness
- Debian is "The Universal Operating System"
- Debian is the most used Embedded Distribution
- Hardware advances will make Debian feasible on new platforms
- It’s enlightening to see and play with other architectures/systems
C types

- **ANSI-C**
  - `sizeof(char) <= sizeof(short) <= sizeof(int) <= sizeof(long)`
  - `short` and `int` are at least 16bit
  - `long` is at least 32bit
  - `sizeof(ptr) != sizeof(int)`
  - Signedness of chars is arch dependent

- **Tips**
  - Use `int` as much as possible for computations, loop variables, ...
  - Use ISO C99 types (`u_int8`, `u_int16`, `u_int32`, ...) for external comms
  - Don’t abuse chars to ‘save memory’
  - Use the latest gcc version with `-Wall`
typedef struct bitfields {
    unsigned char bitfield0:3;
    unsigned char bitfield1:5;
}
Consider 0x12345678

Little endian : 0x78 0x56 0x34 0x12

Big endian : 0x12 0x34 0x56 0x78

PDP endian : 0x34 0x12 0x78 0x56

- External interfaces
- Use macros to convert between CPU and specific endianess
Alignment

- Most RISC CPUs require aligned accesses
- Unaligned accesses are trapped (mostly)
  - slow
  - not possible in kernel land
- Unaligned accesses are seldomly atomic with respect to SMP/other bus masters
- Better
  - avoid them
  - have the compiler generate the code
Let's port together. Debian fun for everyone.

Peter De Schrijver

Portability
Why?
C types
Bitfields
Endianness
Alignment

Peripherals
Architectures
Trends
Out of order
I/O
Userland

Intel style system

CPU

North bridge

Memory subsystem

Device

Device

Device
Intel style system

- Main components
  - CPU complex
  - Northbridge
  - Southbridge
  - Memory subsystem

- Main interfaces
  - Frontside bus
  - PCI
  - AGP
Let's port together. Debian fun for everyone.

Peter De Schrijver

Portability
Why?
C types
Bitfields
Endianness
Alignment

Peripherals
Architectures
Trends
Out of order
I/O
Userland

Opteron style system

Memory subsystem
CPU
HT PCIe bridge
Device
Device
Device
Device
Device
Device
CPU
HT PCIe bridge
Memory subsystem
Opteron style system

▶ Main components
  ▶ CPU
  ▶ Hypertransport - PCIe bridge
  ▶ PCIe - PCI bridge

▶ Main interfaces
  ▶ Hypertransport
  ▶ PCIe
  ▶ PCI
Trends in system design

Observations
- CPUs became much faster than memory
- Bus and memory bandwidth have gone up faster than latencies
- Parallel busses become very hard at high speeds

Solutions
- Caches
- Burstmode transfers
- Advanced DMA
- Multiple highspeed serial links
Out of order transactions

- examples
  - CPU
  - bus bridges
- use read/write barriers
  - CPU instructions
  - "magic" reads

possibly out of order:
  stw r20,0x20(r21)
  stw r22,0x24(r21)

always in order:
  stw r20,0x20(r21)
eieio
  stw r22,0x24(r21)
Non-coherent I/O

- Some systems do not support "bus snooping"
- Invalidate cachelines
  - network traffic
  - disk buffers
  - other kinds of streaming I/O
- non-cacheable memory
  - microcode
  - ring buffers
Addressing

- Virtual addresses
- Physical addresses
- Bus addresses
- Translation Physical to Bus addresses
  - identity mapped
  - fixed offset
  - page based translation
  - not memory mapped
    - IA32 I/O ports
    - PowerPC DCB
- Always access hardware via special functions
Transaction atomicity

- Multiple CPUs
- other busmasters (eg. on PCI)
- reads and writes are atomic only if aligned
- atomic read/modify/write is CPU specific
  - ia32: lock prefix on specific instructions
  - mips: il/sc
  - arm: swap
  - ppc: lwarx/stwcx
- bridges may break locks
Userland hardware access

- Hardware access from userland is problematic
- Separate command transport from driver logic
  - Firewire: libraw1394
  - USB: libusb
  - SCSI and ATAPI: scsi generic like ioctl
  - ...
- Provide abstraction layer for accessing hardware