AHCI

Doing storage right
About Me

- Alexander Graf
- KVM and Qemu developer
  - Server class PowerPC KVM port
  - S390x Qemu guest support
  - x86 Mac OS X in KVM
  - Nested SVM
  - Xenner
- ...

Dienstag, 16. August 2011
Storage
Naming

Host OS

Storage Controller

Hard Disk

IDE   AHCI   virtio-blk

PATA   SATA
Storage

Application

Host OS

Storage Controller

Hard Disk

write( )
Storage

Application

Host OS

Storage Controller

Hard Disk
Storage

Application
Host OS
Storage Controller
Hard Disk

RAM

Dienstag, 16. August 2011
Storage

Application

Host OS

Storage Controller

Hard Disk

RAM

Dienstag, 16. August 2011
Virtualization
Virtualization
Virtual Storage
Virtualized Storage

Application → Host OS → Storage Controller → Hard Disk
Virtualized Storage

Application → Guest OS → Storage Controller Emulation → Qemu → Host OS → Storage Controller → Hard Disk
Virtualized Storage

- Application
- Guest OS
- Storage Controller Emulation
- Qemu
- Host OS
- Storage Controller
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller

write( )

Dienstag, 16. August 2011
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller
Virtualized Storage

- Application
- Guest OS
- Storage Controller Emulation
- Qemu
- Host OS
- Storage Controller
- RAM
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller

RAM
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller

RAM

1

2

3

4
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller

RAM

write()
Virtualized Storage

- Application
- Guest OS
- Storage Controller Emulation
- Qemu
- Host OS
- Storage Controller

write()
Virtualized Storage
Virtualized Storage

Application

Guest OS

Storage Controller Emulation
Qemu

Host OS

Storage Controller

RAM

1
2
3
4
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller

RAM

1

2

3

4
Virtualized Storage

Application

Guest OS

Storage Controller Emulation

Qemu

Host OS

Storage Controller

RAM

1
2
3
4
Device Access

- Programs
- Kernel
- Hardware
Device Access

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware

Dienstag, 16. August 2011
Device Access

write()

Host
Qemu

Guest
Programs
Kernel

Kernel
kvm

Hardware
Device Access

process write() in kernel

Host

Qemu

Guest

Programs

Kernel

kvm

Kernel

Hardware
Device Access

trap device access

Host
- Qemu

Guest
- Programs
- Kernel

Kernel
- kvm

Hardware
Device Access

go to qemu to handle access
Device Access

dispatch access

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware

Dienstag, 16. August 2011
Device Access

handle access in device code

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware

Dienstag, 16. August 2011
Device Access

send data to hard disk

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware

Dienstag, 16. August 2011
Device Access

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware

finished writing data

Dienstag, 16. August 2011
Device Access

tell qemu data is written

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware

Dienstag, 16. August 2011
Device Access

enter `kvm` to run guest again

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware
Device Access

execute guest

Host

Qemu

Guest

Programs

Kernel

Kernel

Hardware

dienstag, 16. August 2011
Device Access

tell program data is written

Host

Qemu

Guest

Programs

Kernel

kvm

Hardware
Device Access

write() finished

Host

Qemu

Guest

Programs

Kernel

Kernel

kvm

Hardware
IDE

Kernel

Hardware

kvm

Kernel Programs

Guest

Qemu

Host

Dienstag, 16. August 2011
IDE

- set feature
- set nsector
- set sector
- set cylinder
- set hcylinder
- set drive

issue command = DMA WRITE

set DMA SG list address

set DMA command = START
AHCI

START PROCESSING files

Dienstag, 16. August 2011
Communication

old IDE

PIO

Dienstag, 16. August 2011
Communication
Communication

MMIO

DMA

AHCI

Dienstag, 16. August 2011
Communication

Dienstag, 16. August 2011
Multiple Requests
Multiple Requests

write
write
write
write

AHCI
(NCQ)

done
done
done
done
Multiple Requests

write
write
write
write

Dienstag, 16. August 2011
## Guest OS Support

<table>
<thead>
<tr>
<th></th>
<th>IDE</th>
<th>AHCI</th>
<th>Virtio-blk</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Linux</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>✔️</td>
</tr>
<tr>
<td><strong>Windows XP</strong></td>
<td>✔️</td>
<td>❌</td>
<td>external</td>
</tr>
<tr>
<td><strong>Windows Vista</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>external</td>
</tr>
<tr>
<td><strong>Mac OS X</strong></td>
<td>❌</td>
<td>✔️</td>
<td>❌</td>
</tr>
<tr>
<td><strong>BSD</strong></td>
<td>✔️</td>
<td>✔️</td>
<td>❌</td>
</tr>
</tbody>
</table>
Dienstag, 16. August 2011

(host) qemu-img create /dev/shm/test.raw 20G
(guest) dd if=/dev/Xda of=/dev/null bs=1M count=1000 iflag=direct
Benchmark ran on Intel Atom z530
Caching
Caching

write

write

write

write

write

write
cache=writethrough
cache=writeback

- write
- write
- write
- write
- flush
- write
cache=unsafe
Why AHCI

• Faster than IDE
• Compatible with most OSs
• Compatible with all modern devices (CD-ROMs)
• Only need to develop one side
Future ideas

- Default adapter in -M q35
- MSI-X
Questions?
蟻が十