

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Virtio on early kernels

How does it work & implementation

Nahim 'Naam' El Atmani

EPITA

July 16, 2015

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion



Problem

What you have: basic lvl1 weapon



Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion



Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Solution: Virtio

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

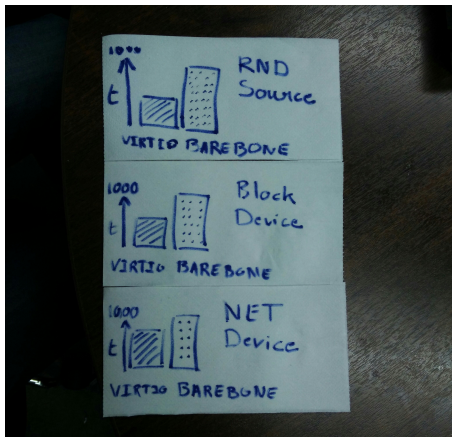
Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion



Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

libvirt wiki's

Virtio is a virtualization standard for network and disk device drivers where just the guest's device driver "knows" it is running in a virtual environment, and cooperates with the hypervisor.

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Great variety of solutions for doing virtualization under Linux:

- KVM (Kernel-based Virtual Machine, kvm.ko)
- QEMU (emulator with device virtualization)
- LXC (OS level)
- Iguest (paravirtualization, lg.ko)
- ...

Idea

- Each of these solutions may virtualize devices
- “OS is taxed on the base of their independent needs”

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

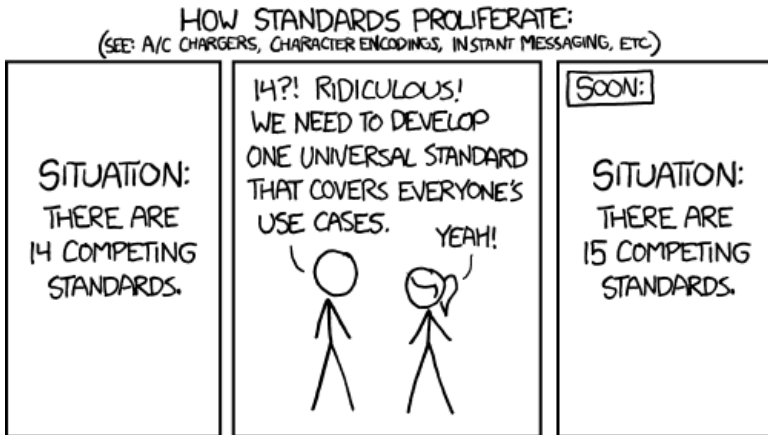


Figure: xkcd 927

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Discovery

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

- Common front end
- Code reuse

How

One wants to implement generic front end guest drivers for the Virtio's API.

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Subsystem device ID	Virtio Device	Specification
1	network card	Appendix C
2	block device	Appendix D
3	console	Appendix E
4	entropy source	Appendix F
5	memory ballooning	Appendix G
6	ioMemory	-
7	rpmmsg	Appendix H
8	SCSI Host	Appendix I
9	9P transport	Appendix J
10	mac80211 wlan	-
11	rproc serial	-
12	virtio CAIF	-

Virtio on early kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

STOS Support	Before	After
Block device	X	✓
Network	X	✓
Console	✓	✓
Random source	X	✓
Memory ballooning	X	X

You said paravirtualization?

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

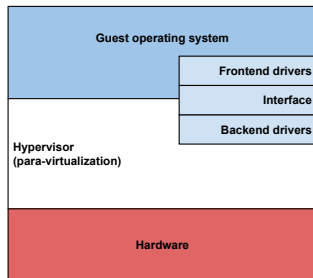
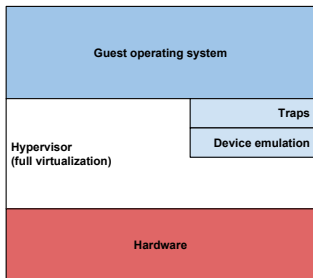
With
myKernel?

Benchmarks

Conclusion

Virtio provide a way to standardize the communication interface between:

- hypervisor: backend drivers / emulated devices
- guest: front end drivers



Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Virtio internals

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution: Virtio

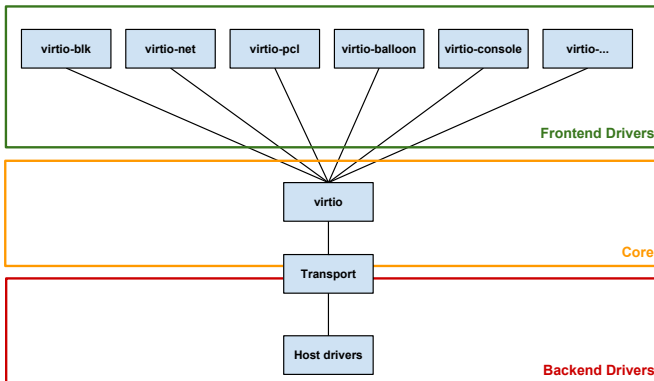
Discovery

Virtio internals

With myKernel?

Benchmarks

Conclusion



Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution: Virtio

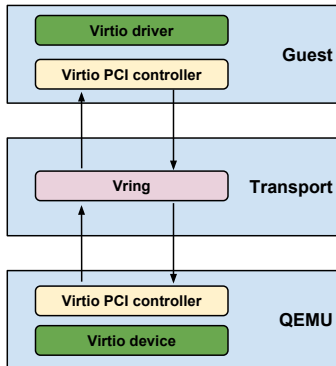
Discovery

Virtio internals

With myKernel?

Benchmarks

Conclusion



Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Communication between host and guest are made through Virtio ring buffers (vrings).

- Predefined number of buffers
- Some are optionnals, others are mandatory

Usage of optional buffers are subject to the rules of common sense.

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Can be summarized as:

- 1 R/W features bits
- 2 R/W the configuration space
- 3 R/W status bits
- 4 Reset

There is a kind of handshake phase used to sync host and guest on used/available features. The host must now what feature bits are acked by the guest.

Host	\cap	Guest
VIRTIO_NET_F_CSUM	✓	VIRTIO_NET_F_GUEST_CSUM
VIRTIO_NET_F_HOST_UFO	✗	0
VIRTIO_NET_F_HOST_TSO4	✓	VIRTIO_NET_F_GUEST_TSO4
...	?	...

Virtio on early
kernelsNahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Properties:

- device specific-information
- R/W acces from guest

VIRTIO_NET_F_MAC feature bit specify if the host have a particular MAC address for the guest. The address is specified in this configuration space.

Linux example

```
/* Configuration may specify what MAC to use. Otherwise random. */  
if (virtio_has_feature(vdev, VIRTIO_NET_F_MAC))  
    virtio_cread_bytes(vdev,  
                       offsetof(struct virtio_net_config, mac),  
                       dev->dev_addr, dev->addr_len);  
else  
    eth_hw_addr_random(dev);
```

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Represent internal state of the device probe.

```
./include/uapi/linux/virtio_config.h
```

virtio_config.h

```
/* We have seen device and processed generic fields */  
#define VIRTIO_CONFIG_S_ACKNOWLEDGE 1  
/* We have found a driver for the device. */  
#define VIRTIO_CONFIG_S_DRIVER      2  
/* Driver has used its parts of the config, and is happy */  
#define VIRTIO_CONFIG_S_DRIVER_OK   4  
/* Driver has finished configuring features */  
#define VIRTIO_CONFIG_S_FEATURES_OK 8
```

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Does what you think it does:

- reset configuration
- reset status bit
- ...

Needed by modular drivers.

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

With myKernel?

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

- Work with not so dumb kernel
- Get some libs
- Review your roadmap
- PCI devices support

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

- Cut the specs
- Organize your work

Subsystem device ID	Virtio Device	Order	Specification
1	network card	4	Appendix C
2	block device	3	Appendix D
3	console	2	Appendix E
4	entropy source	1	Appendix F

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

- Cut does not means dumb cut
- Framework constraints
- Arch dependent features

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

- virtio-gpu / vga & Co
- virtio-input
- ...



Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Benchmarks

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

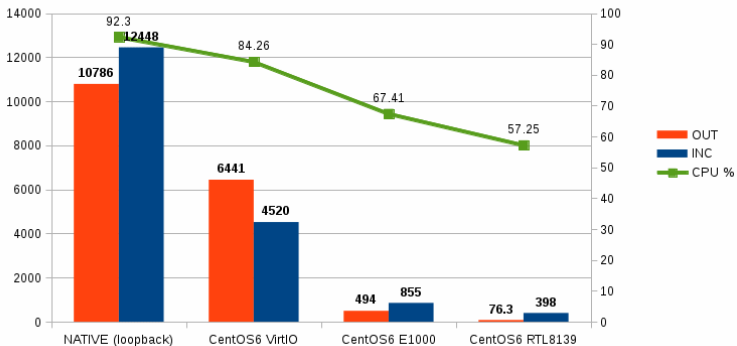
Virtio internals

With
myKernel?

Benchmarks

Conclusion

TCP transfer speed, Mbit/s and CPU load



Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

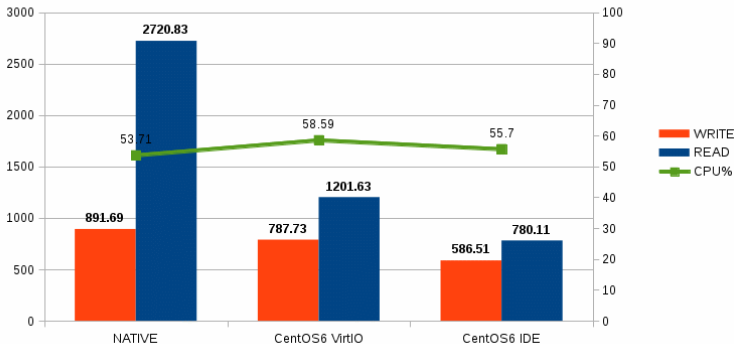
Virtio internals

With
myKernel?

Benchmarks

Conclusion

Disk sequential speed, MB/s and CPU load



Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Conclusion

Virtio on early kernels

Nahim 'Naam' El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

- Best of the full and para virtualization
- Dev boost
- Cool features

Virtio on early
kernels

Nahim 'Naam'
El Atmani

Problem

Solution:
Virtio

Discovery

Virtio internals

With
myKernel?

Benchmarks

Conclusion

Here is some links used for this talk:

- 1 Virtio: An I/O virtualization framework
- 2 Virtio: Towards a De-Facto Standard For Virtual I/O Devices
- 3 Benefits of paravirtualization
- 4 Network driver example
- 5 Benchmarks
- 6 STOS kernel