



IBM

Virtualizing Older Hardware

lguest and kvm-lite

Anthony Liguori – aliguori@us.ibm.com
Open Virtualization
IBM Linux Technology Center

November 1st, 2007

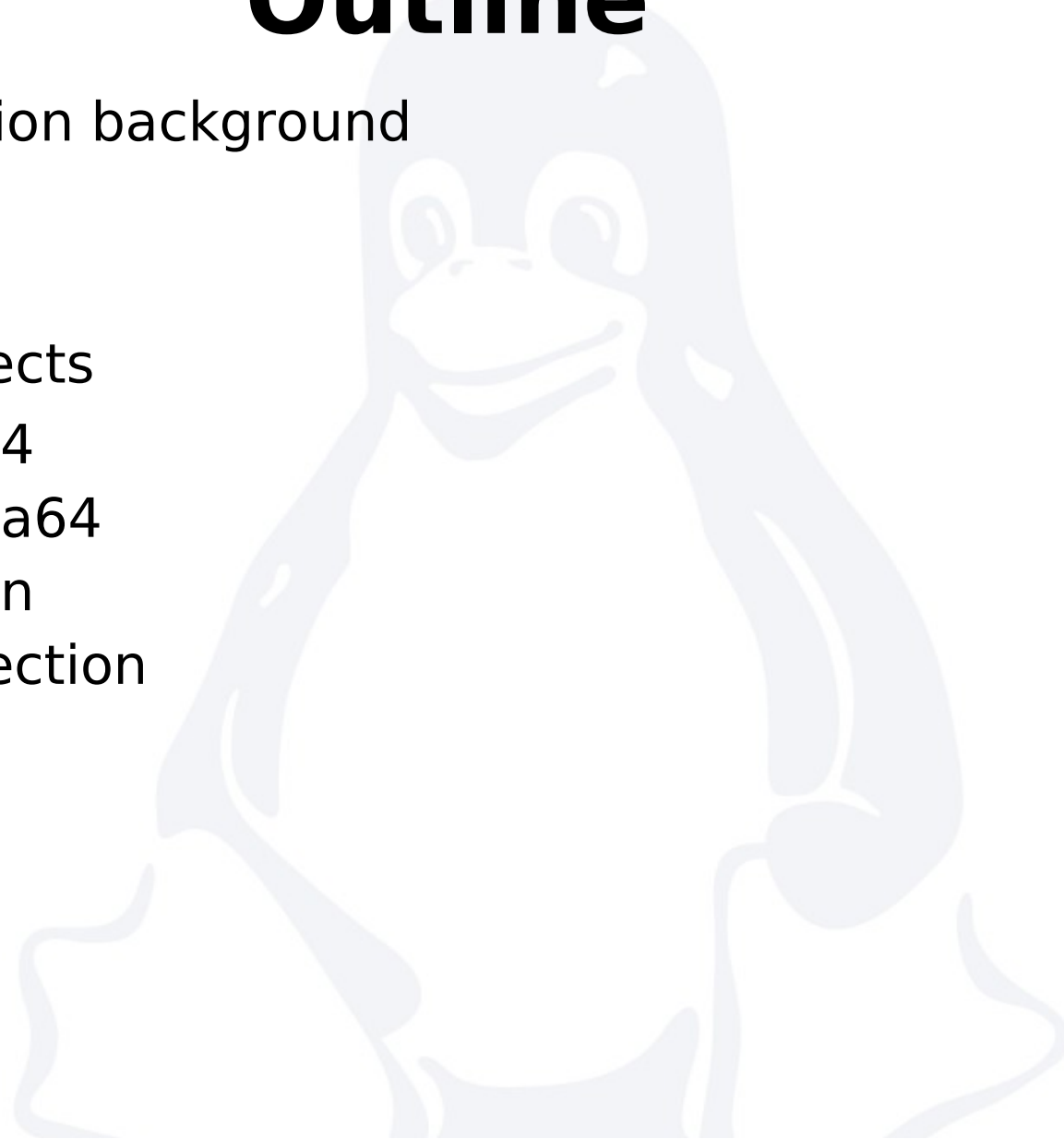
Linux is a registered trademark of Linus Torvalds.

IBM



Outline

- Virtualization background
- lguest
- kvm-lite
- other projects
 - lguest64
 - lguest-ia64
 - kvm-xen
- Future direction
- Questions



Full virtualization

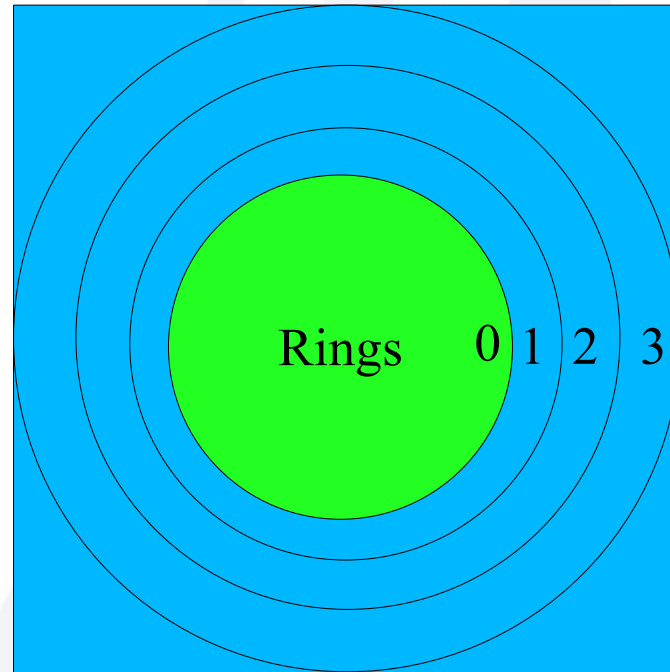
- Allows unmodified guests to run
 - Windows
 - Linux
 - BSDs
- Emulates all IO devices found on a traditional PC
 - Including devices from the 80s that are no longer used
 - Legacy DMA controller
 - RTC - BCD



x86 Full Virtualization

Host Save
Area

Guest Save
Area



New Instructions

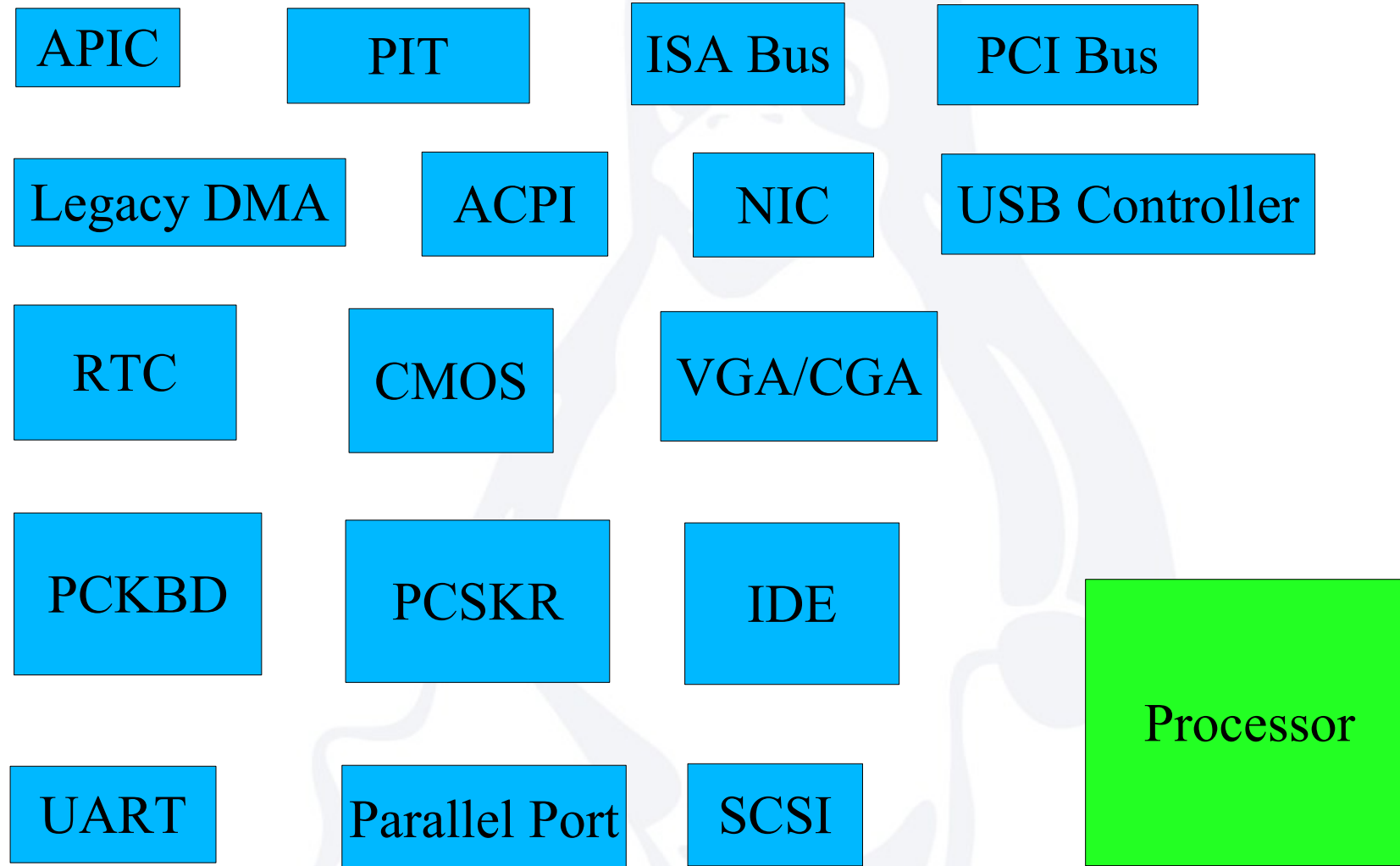
- Guest Switch
- Set Host Area
- Call into Host

Control Area

- Control intercepts
- Store host/guest state



IO Emulation

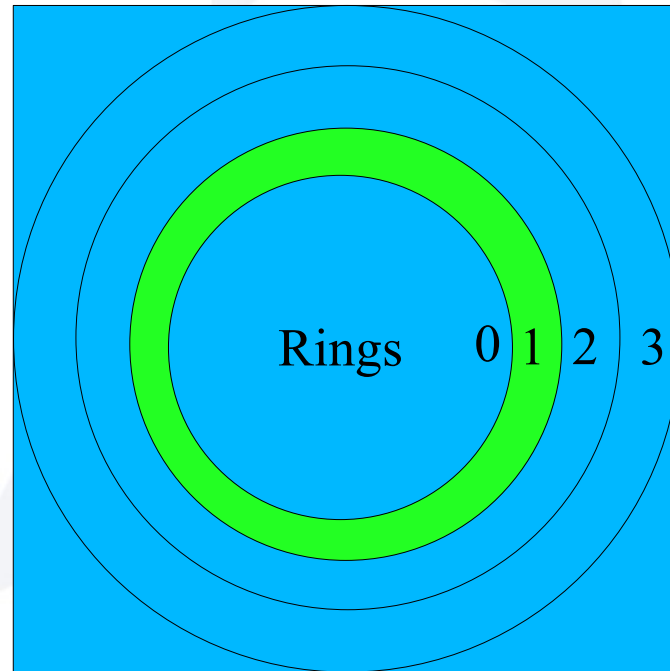
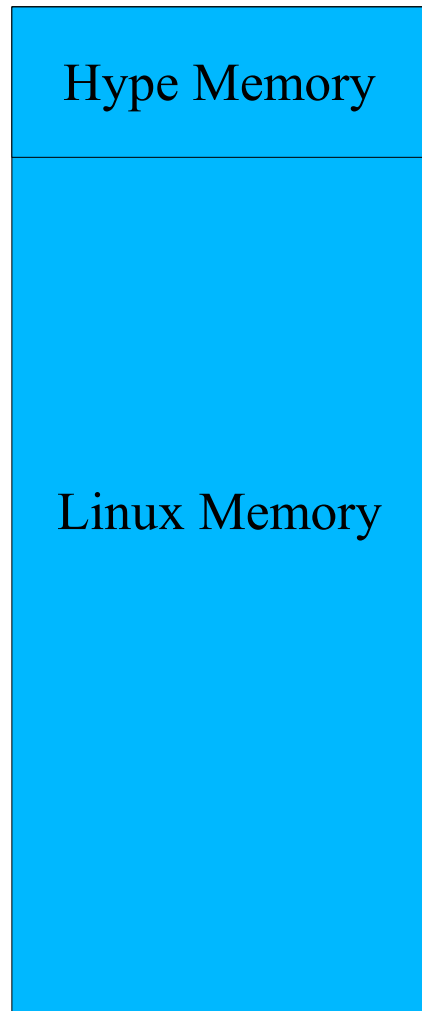


Paravirtualization

- Modify the guest Operating System
 - Certain Linux versions
 - A few versions of FreeBSD/NetBSD
 - Some minimal changes are needed to allow virtualization to work at all (VMI)
 - Why stop there? Do not emulate any legacy hardware devices



x86 Paravirtualization



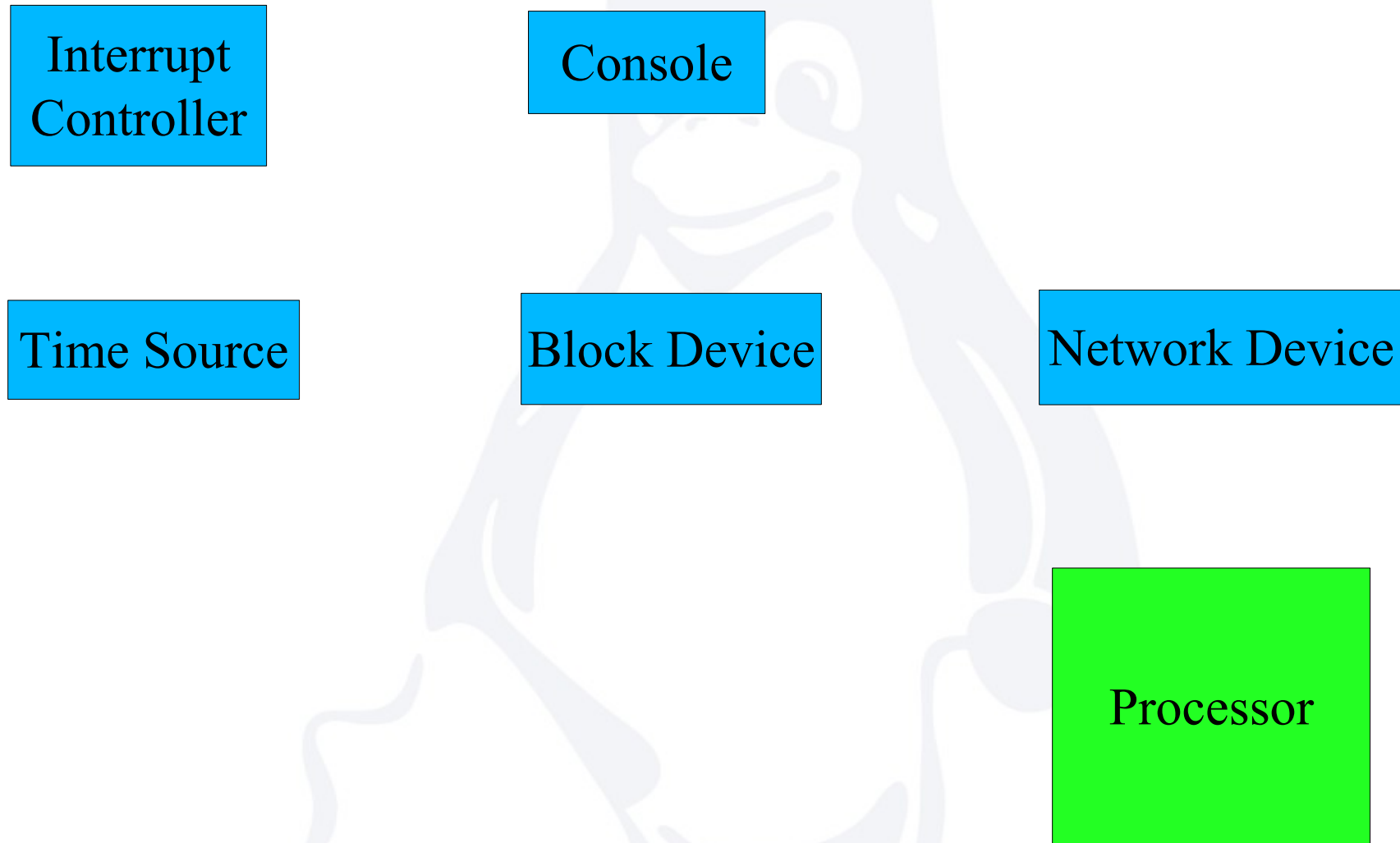
```
pushf
mov $42, %eax
sti
sysexit
```



```
vmcall HYPE_PUSHF
mov $42, %eax
vmcall HYPE_STI
sysexit
```



IO Paravirtualization



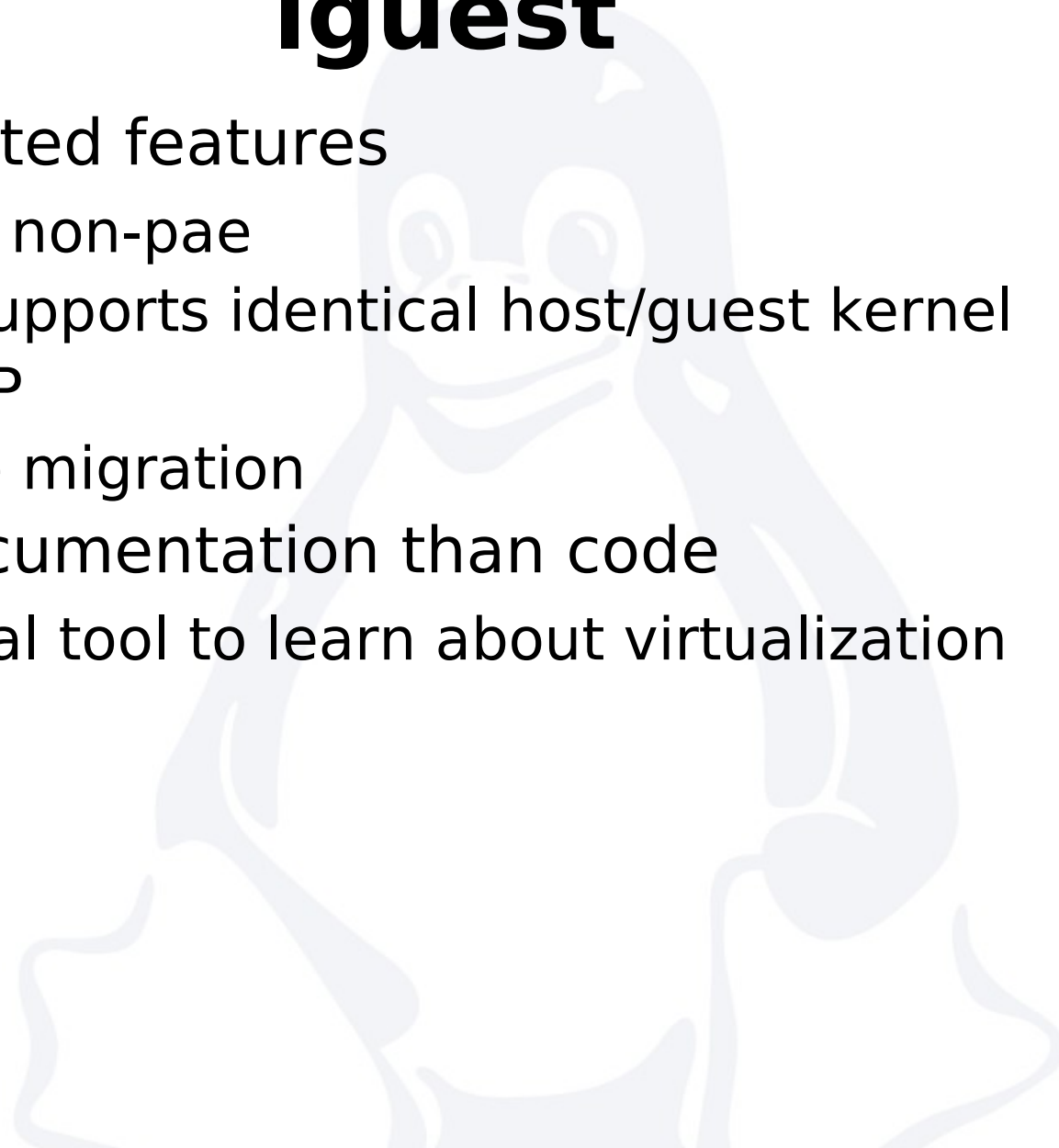
Iguest

- Written by Rusty Russell (IBM)
- Originally developed to validate paravirt_ops
 - At the time, VMI was the only implementation of paravirt_ops
- Help develop kernel community interest in virtualization
- Puppies!



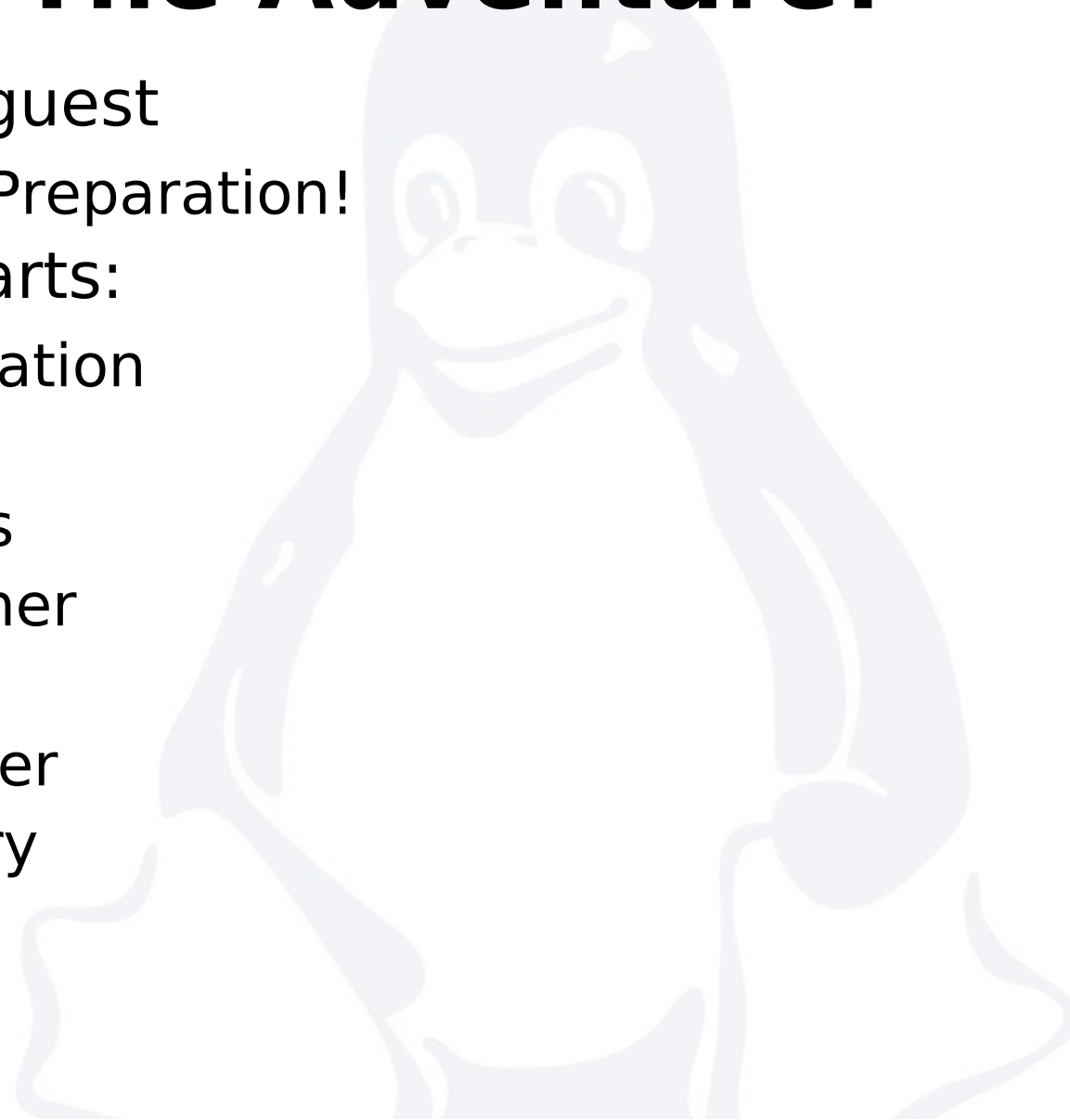
Iguest

- Very limited features
 - 32-bit, non-pae
 - Only supports identical host/guest kernel
 - No SMP
 - No live migration
- More documentation than code
 - An ideal tool to learn about virtualization



The Adventure!

- drivers/lguest
 - make Preparation!
- Seven parts:
 - Preparation
 - Guest
 - Drivers
 - Launcher
 - Host
 - Switcher
 - Mastery

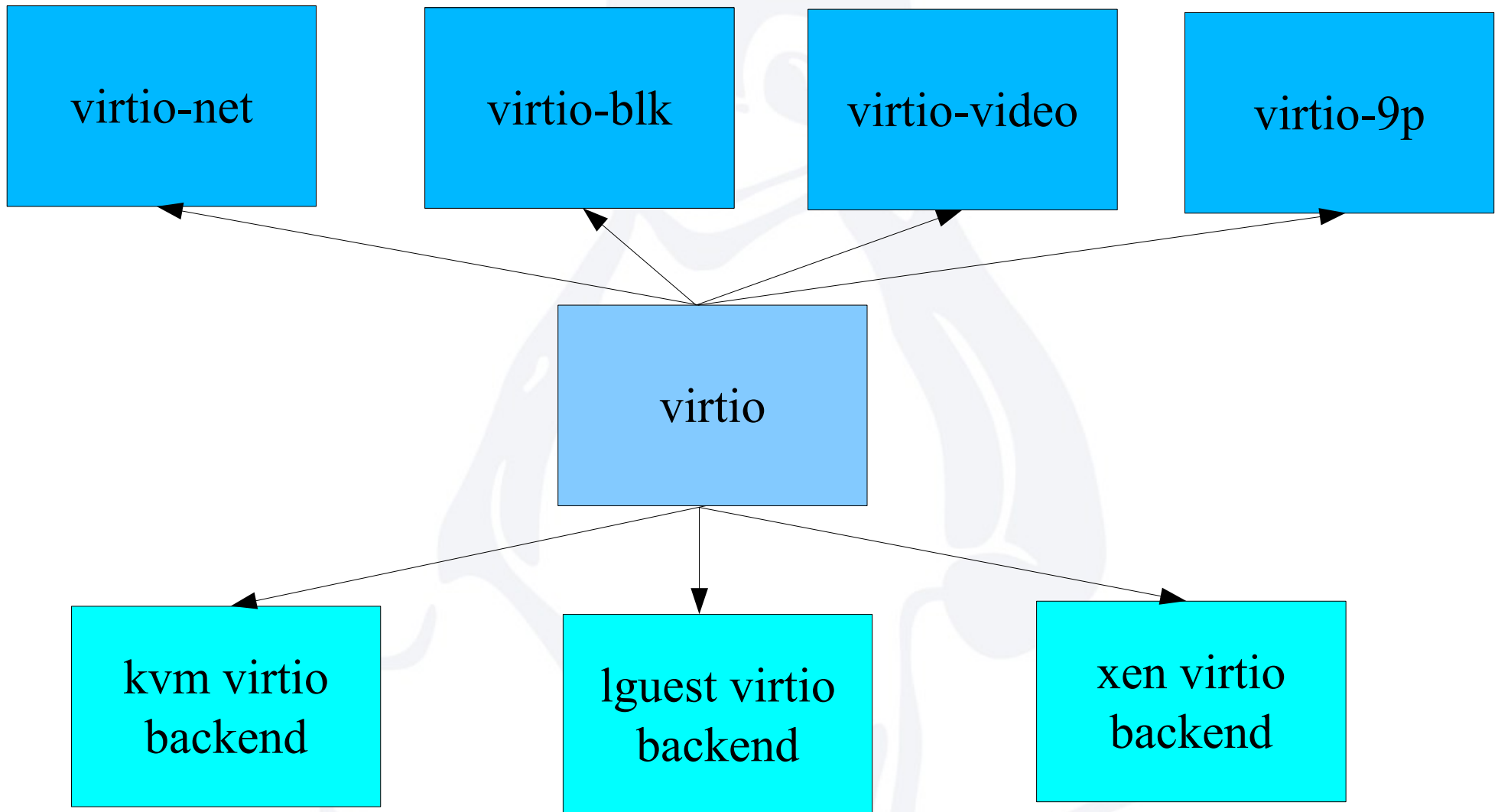


virtio

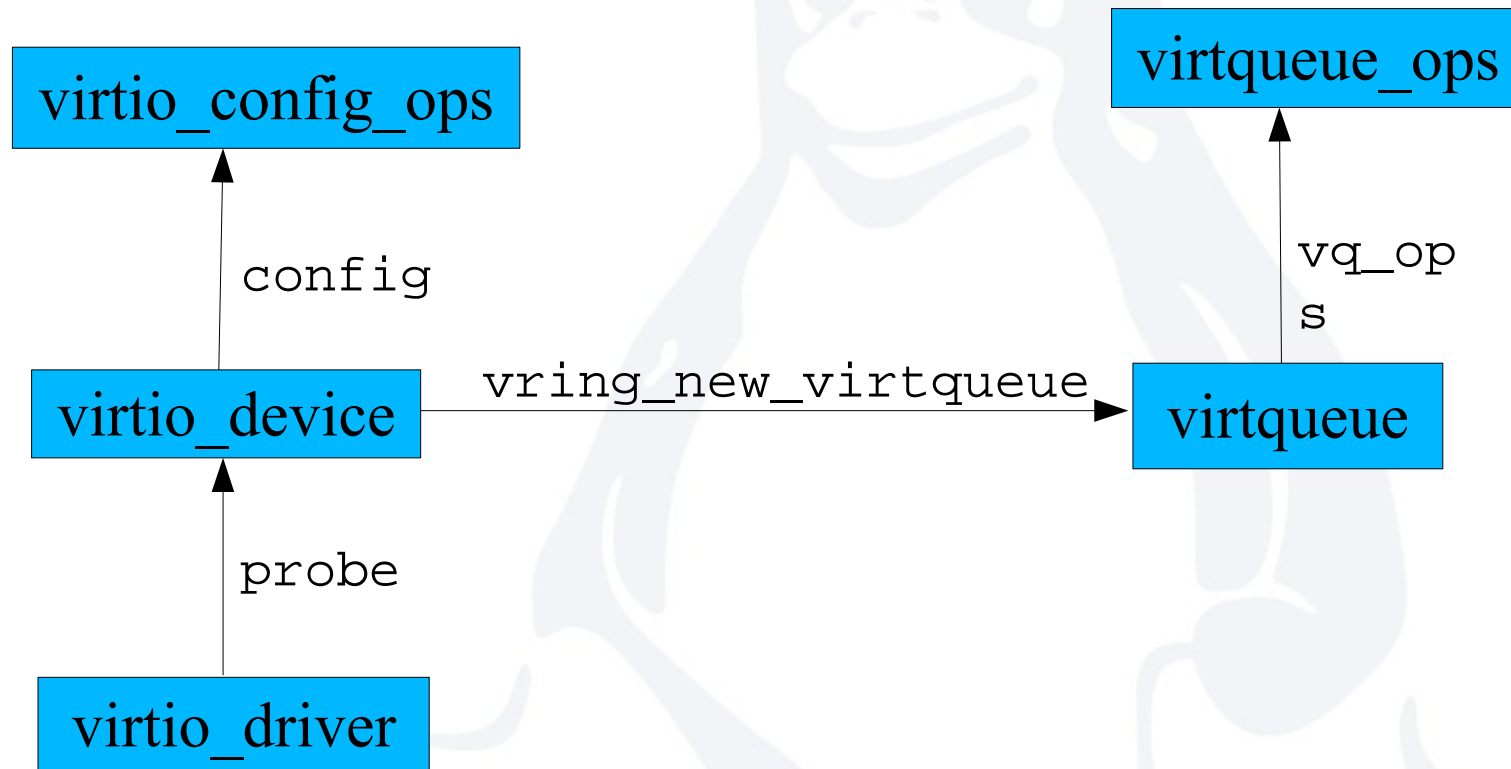
- Linux will support lguest, KVM, Xen, KVM-lite, PHYP, VMware, Viridian, and possibly more
 - If each has 4-5 PV drivers, that's 35 new drivers!
 - All drivers would be doing the same thing
- virtio is an abstraction of the common mechanism of VMMs
 - A single driver could, with little modification, run on many different VMMs
- Especially important for “small” drivers (entropy driver, CPU hotplug, ballooning, etc.)



virtio



virtio



virtqueue_ops

```
struct virtqueue_ops {
    int (*add_buf)(struct virtqueue *vq,
                  struct scatterlist sg[],
                  unsigned int out_num,
                  unsigned int in_num,
                  void *data);

    void (*kick)(struct virtqueue *vq);

    void *(*get_buf)(struct virtqueue *vq, unsigned int*len);

    bool (*restart)(struct virtqueue *vq);

    void (*shutdown)(struct virtqueue *vq);
};
```



kvm-lite

- lguest is not meant to have features
 - A reference; not meant for average user
- kvm-lite will provide support for running KVM on non-VT/SVM processors
- Currently only a pending patch-series to KVM
- Only use paravirtual IO drivers
 - More easily verifiable
 - Potentially could perform better
- Current series is lguest ported to KVM



Integration Issues

- Some challenges integrating into KVM
 - Paravirtualization requires a guest “memory hole”
 - Should kvm-lite run along side kvm-amd/kvm-intel?
 - The paravirtualization ABI should be shared with the rest of KVM



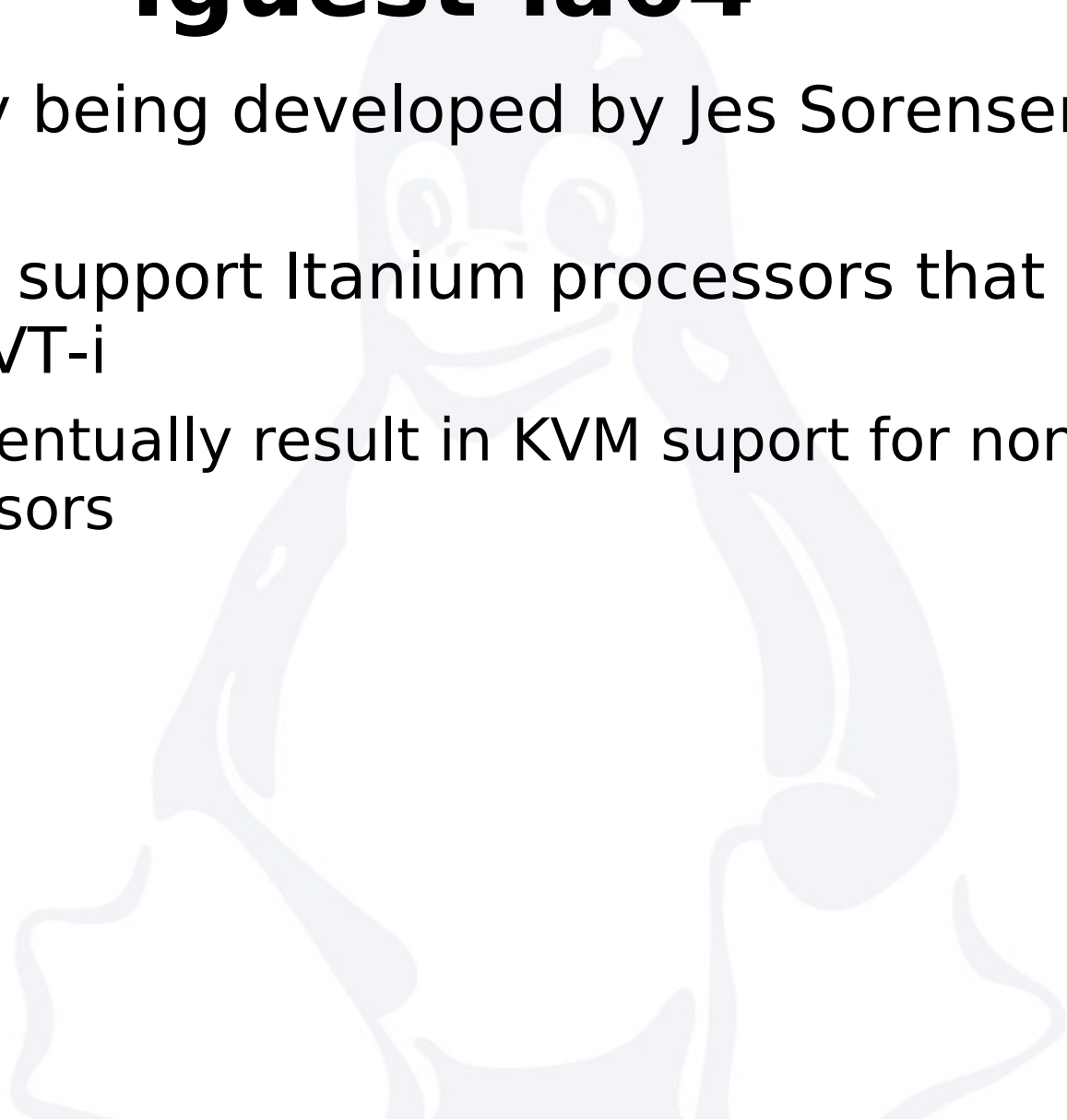
Iguest64

- A project from Steve Rostedt and Glauber de Oliveira Costa of RedHat
- Help define what's needed for paravirt_ops on x86_64
- Some unique challenges
 - x86_64 does not enforce segment limits for fs/gs



Iguest-ia64

- Currently being developed by Jes Sorensen (SGI)
- Meant to support Itanium processors that predate VT-i
 - Will eventually result in KVM support for non-VT-i processors



kvm-xen

- Research project by Ryan Harper and Anthony Liguori (IBM)
 - Run unmodified Xen paravirt guests under KVM
- No longer under development
 - kvm-lite provides a better solution (Xen ABI is cumbersome)
 - Diminishing returns – more processors have VT/SVM every day
- Rik van Riel is working on another project to run Xen guests under KVM
 - Not too many details yet.



Future

- KVM is the interface for Linux virtualization
 - kvm-lite will provide support for older hardware
 - lguest will continue to exist as a reference
 - lguest-ia64 will eventually be merged into KVM
- A simpler userspace will be developed for KVM
 - Easier for validation
- There will probably be some easy way for Xen users to transition to KVM



Questions

- Questions?



Answers :-)

- Can three virtualization technologies co-exist (lguest, KVM, Xen)
 - Of course! lguest currently serves as a test-bed for virtualization technology in Linux. KVM and lguest will share most significant code (like paravirt IO).
 - lguest's functionality will also be integrated in KVM
 - KVM is fundamentally different from Xen. They certainly aren't mutually exclusive.



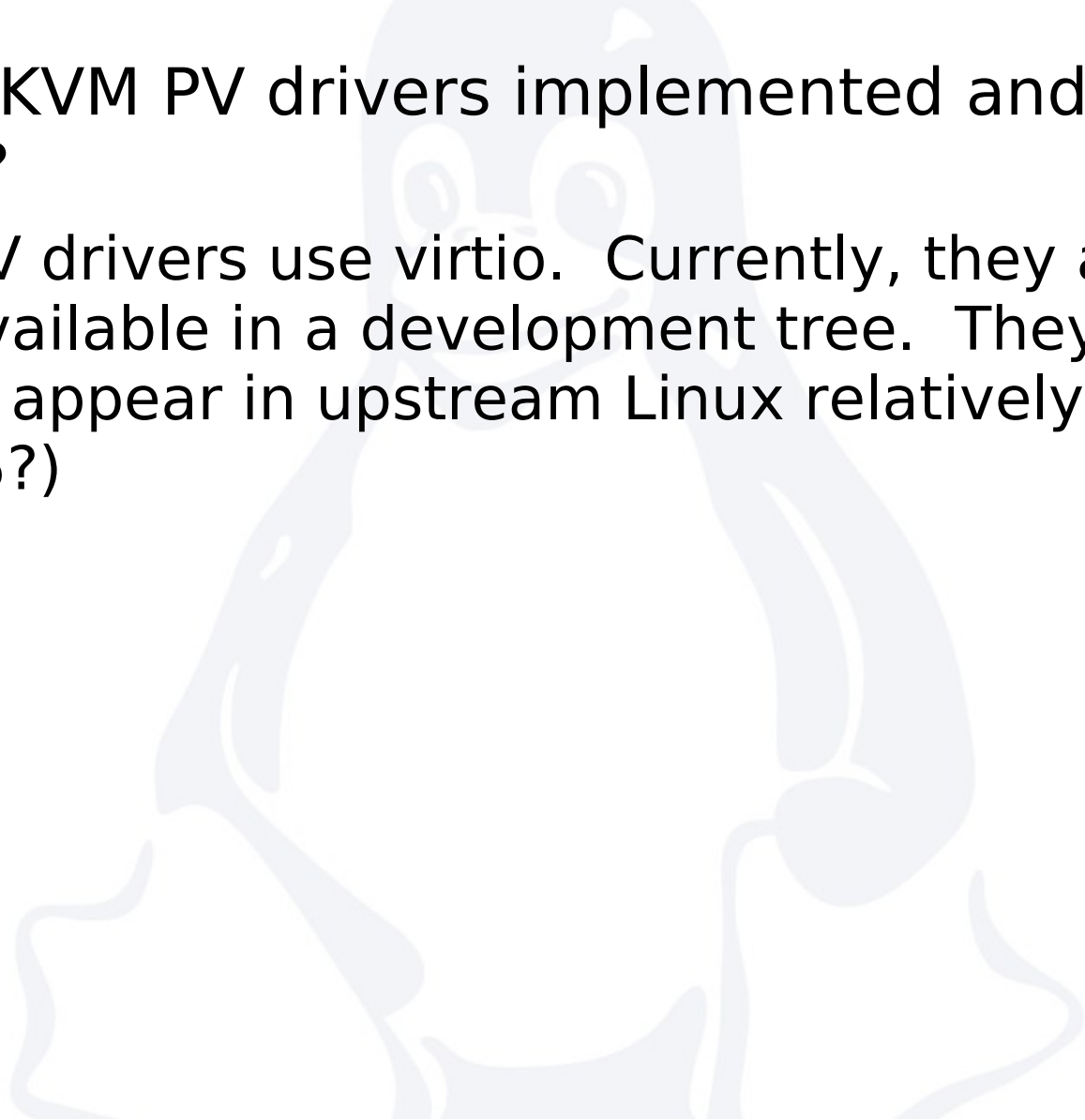
Answers

- How do you install lguest
 - Build a host kernel with CONFIG_LGUEST
 - Build a guest kernel with CONFIG_LGUEST_GUEST
 - build the lguest utilities in Documentation/lguest
- Run lguest:
 - Documentation/lguest/lguest 64 vmlinux –
tunnet=192.168.19.1 –block=rootfile
root=/dev/vda
- See Documentation/lguest/lguest.txt



Answers

- How are KVM PV drivers implemented and how to setup?
 - KVM PV drivers use virtio. Currently, they are only available in a development tree. They should appear in upstream Linux relatively soon (2.6.25?)



Answers

- What is a small hypervisor? What areas are kvm-lite useful for?
 - Very interesting question! Depending on time, we can discuss TCB or just about the paravirtual IO verses emulated IO.
 - kvm-lite is primarily useful for older hardware that does not have VT/SVM.



Answers

- What is the best use of Iguest?
 - Learning how to become a virtualization hacker!

