
Glauber Costa, Lead Engineer
glommer@cloudius-systems.com

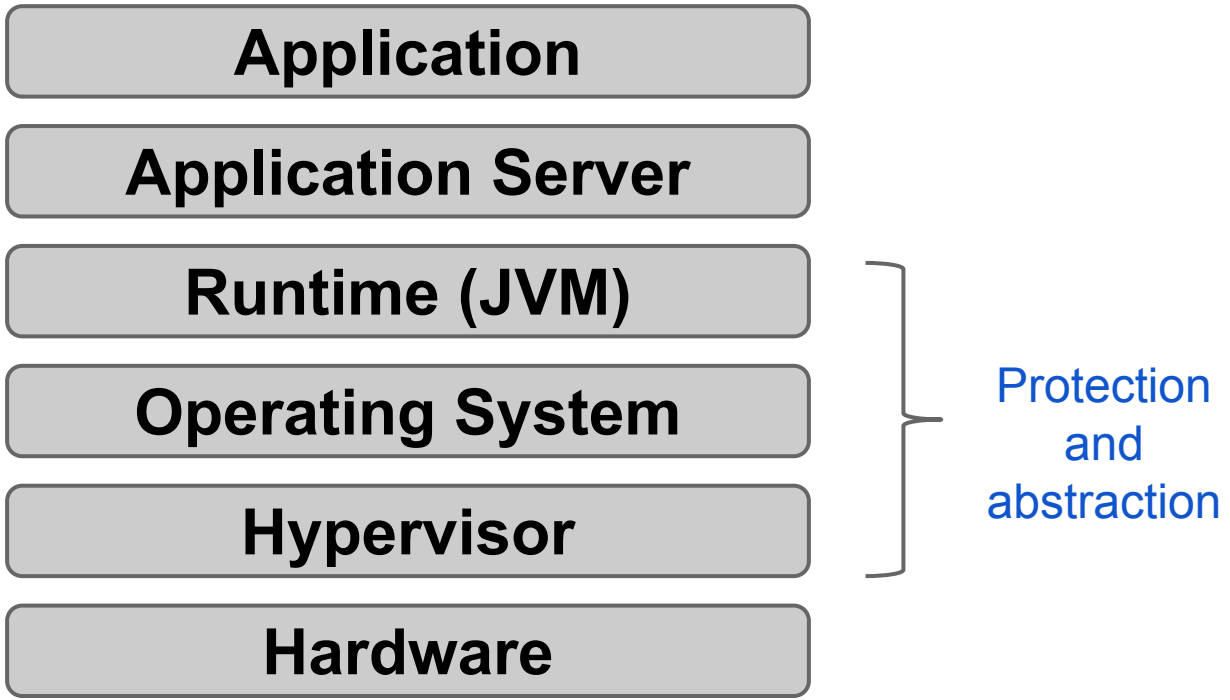
OSv
designed for the cloud
alpha

# Who are we

- Small startup founded by Avi Kivity (Creator of KVM)
- Around 15 people, including some Linux veterans
  - 10 countries
- Headquarters in Israel
- Our mission is to build the default OS for public and private Clouds.

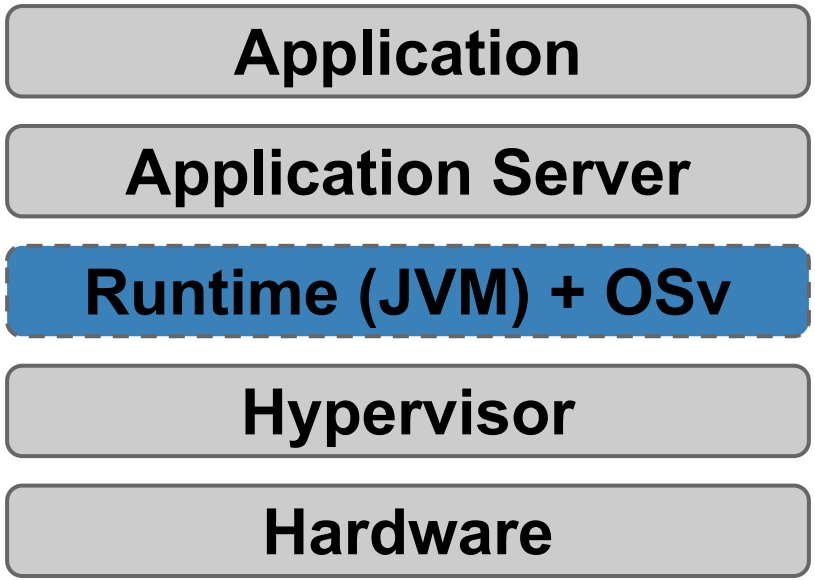
OSv
designed for the cloud
alpha

# Typical Cloud Stack



OSv
designed for the cloud
alpha

# “Library OS”



OSv
designed for the cloud
alpha

# OSv

- LibraryOS written from scratch in C++11
  - Actually network stack came initially from FreeBSD, heavily modified
  - ZFS filesystem from Open Solaris
- BSD Licensed.
- Implements most of the (Linux) POSIX API
- Better if we have a runtime (Java at the moment)
- It supports only one application.
- KVM, Xen, VirtualBox, VMWare
- Image size as low as ~15 Mb
- Beta expected in a couple of months.

OSv
designed for the cloud
alpha

# OSv and NoSQL

- Think of us as "No-OS"
- You can distribute a copy of your application that will run directly on a hypervisor/cloud
- Cassandra
  - Due to its Java focus
- in memory redis
  - Due to its pervasiveness
- memcached
  - Due to its simplicity
- mongodb
  - Because I wanted to.

OSv
designed for the cloud
alpha

# What do people think?

Roman Shaposhnik, Bigtop/Hadoop

*"Mark my words, GridGain [...] and OSv [...] are going to be excitingly disruptive in the next few years. […] And, by the way, if, after reading this blog, you are not dropping everything and porting your cloud application to OSv, I don't know what's wrong with you."*

http://blog.gopivotal.com/pivotal/features/pivotal-debuts-at-apachecon-north-america-2014-thanks-for-having-us

OSv
designed for the cloud
alpha

# Performance

- System calls are free
  - context switches are really cheap ( x 4 Fedora* Linux)
- Network performance significantly faster
  - around 20 % with netperf over Fedora Linux
  - more than 50 % for some UDP workloads
- SpecJVM between 3 and 5 % faster.
  - Not a lot actually expected
- Boots < 1 second.

* Fedora 19, roughly 6 - 8 months old Linux Kernel

OSv
designed for the cloud
alpha

# Real savings: sysadmins

- No command line,
  - except for compatibility
- no graphical interface either.
- REST API for full automation.
- Almost zero configuration
  - Compare with almost 200 in a very stripped down Linux
  - or even registry.

OSv
designed for the cloud
alpha

# Capstan

- Capstan is a tool for building and running your application on OSv (under a hypervisor)
- Docker-like command line interface
  - capstan run cloudius/cassandra
- Multi-platform
  - Linux, OS X, Windows
  - QEMU/KVM, VirtualBox, VMware
  - Google Compute Engine, Amazon EC2

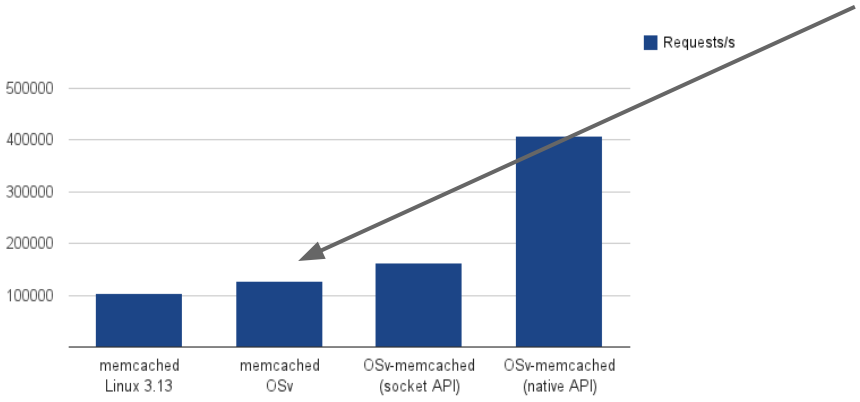
OSv
designed for the cloud
alpha

# mmap

- mmap is a system call for mapping files to memory
  - NoSQL databases rely on it for persistence, caching, I/O, and copy-on-write
- Page cache behavior is important for database performance
  - Recent Linux & PostgreSQL fsync() problems
- ZFS + mmap = :-(
  - ZFS ARC cache is not integrated with page cache.
  - We are trying to fix that.

# memcached

- unmodified:
- modified:
  - that showcases the real performance opportunity behind OSv.




That's already ~20 %
more at http://osv.io/benchmarks/

OSv
designed for the cloud
alpha

# Cassandra

- A bug in our mmap implementation creates some instability
- Performance is on pair with Linux
  - But room for improvement is very big.
- There is, however, less need to worry about JVM parameters due to JVM Ballooning and friends

OSv
designed for the cloud
alpha

# MongoDB, redis and more

- recently started to meddle with them
  - No trusted numbers for them yet - follow us for news!
- Mongo has a problem with mmap as well as Cassandra
  - Should be ready soon
- Redis is way too dependent on COW/fork for persistence
- Mongo and Cassandra are quite dependent on mmap
  - There is still work to do with that, but over time using our own APIs can help there as well

OSv
designed for the cloud
alpha

# To know more

http://osv.io

https://github.com/cloudius-systems/osv

@CloudiusSystems

osv-dev@googlegroups.com

OSv
designed for the cloud
alpha

# Virtualization Oriented

- No spinlocks in the whole kernel
  - Because of lockholder preemption
- No complicated hardware model
- Avoid things that are traditionally more expensive in HV
  - Like IPIs and timer setting
- We also feature a fair scheduler, support multiple page sizes transparently and avoid page metadata overhead.