It's Time to Debloat the Cloud with Unikraft

Felipe Huici, Simon Kuenzer, Sharan Santhanam – NEC Laboratories Europe GmbH
Alexander Jung – Lancaster University

USENIX LISA 2021, June 1st-3rd

felipe.huici@neclab.eu

The Private/Public Cloud

200

Cloud computing with AWS

Amazon Web Services (**AWS**) is the world's most comprehensive and broadly adopted **cloud** platform, offering over 200 fully featured **services** from data centers globally.

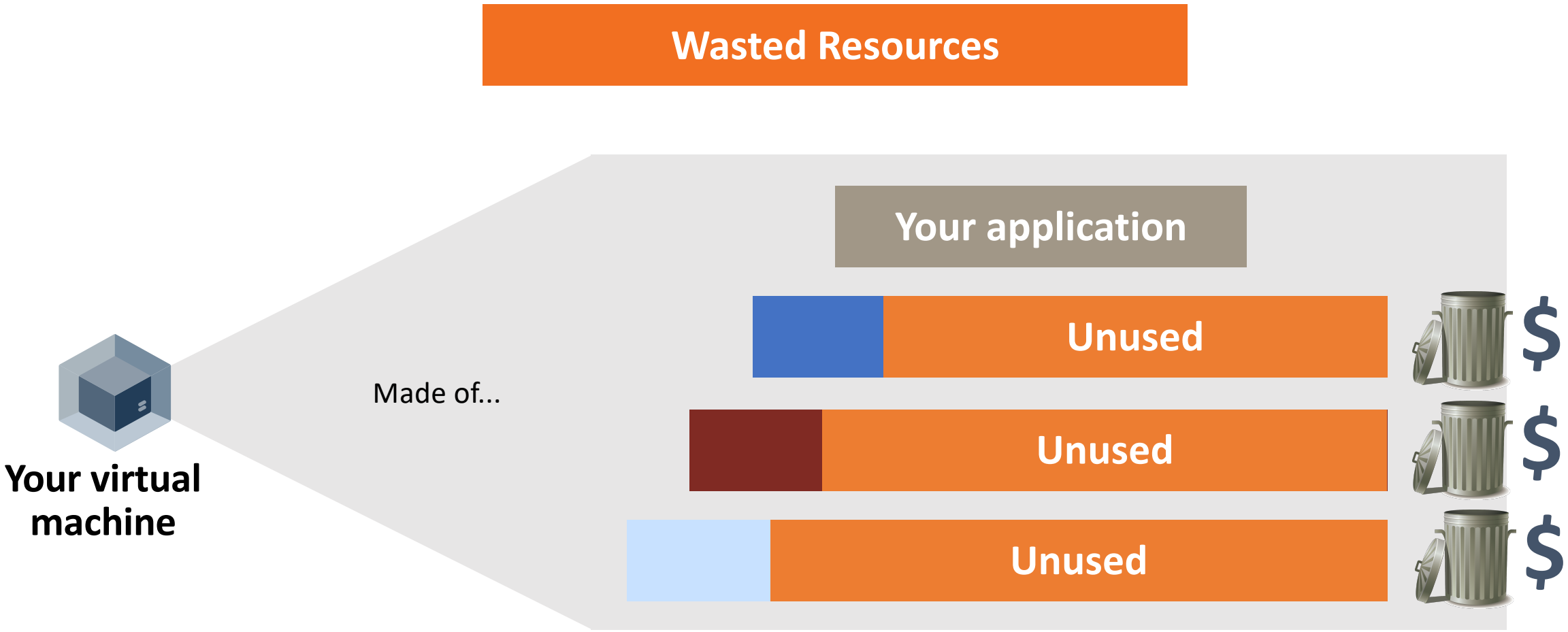
Multitude of services

Bad for the environment

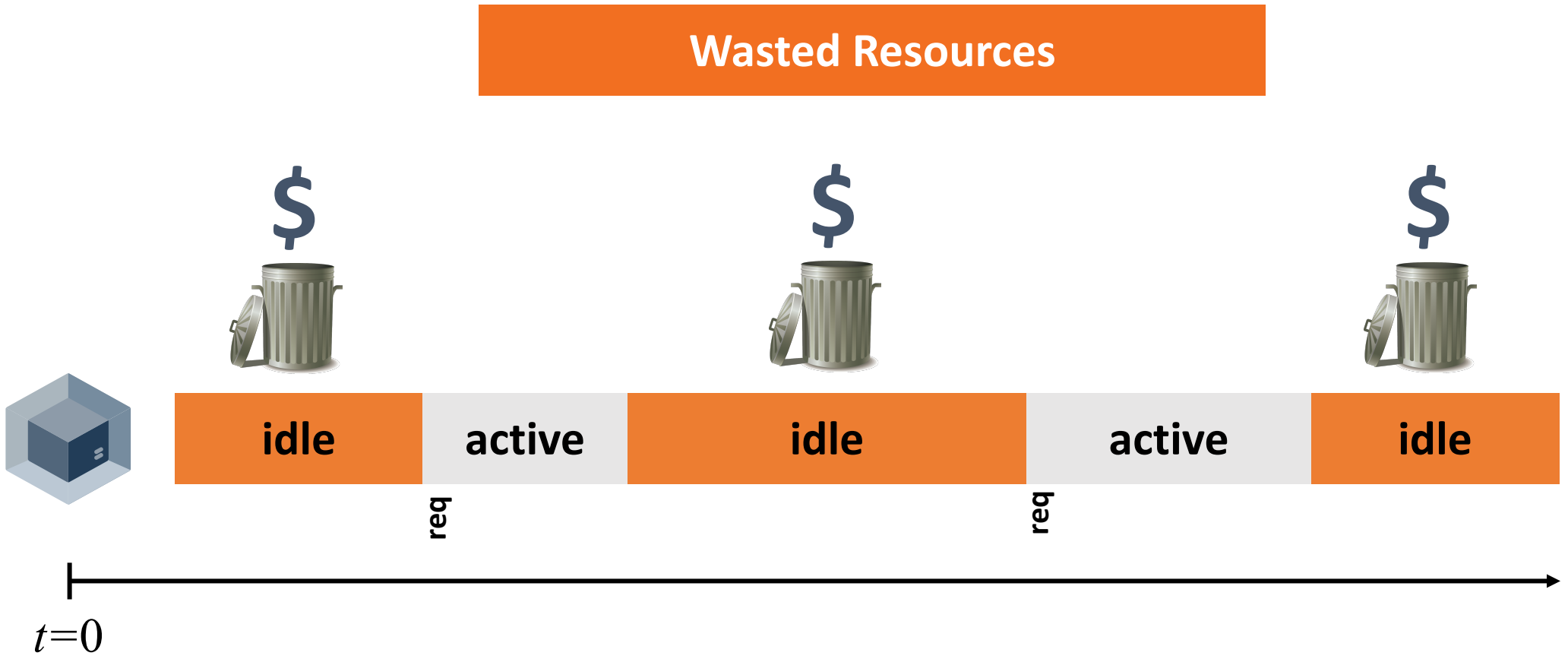
We need to debloat our VMs

We need to stop idle VMs

The Problem Part 1: Size



The Problem Part 2: Time



Specialization = High Efficiency

Unikernels = Specialized Virtual Machines

GOALS

- **Easy to build and run**
- **Easy or no app porting**
- **Great performance**

Unikernel Power

Fast start/stop/migration times

- 10s of milliseconds or less (and as little as 2.3ms)
  REFS: *LigthVM [Manco SOSP 2017], Jitsu [Madhvapeddy, NSDI 2015]*

Low memory footprint

- Few MBs of RAM or less
  REFS*: ClickOS [Martins NSDI 2014], Unikraft [Kuenzer, Eurosys 2021.* ***Best Paper Award****]*

High density

- 8k guests on a singlex86 server
  REFS*: LigthVM [Manco SOSP 2017]*

High Performance

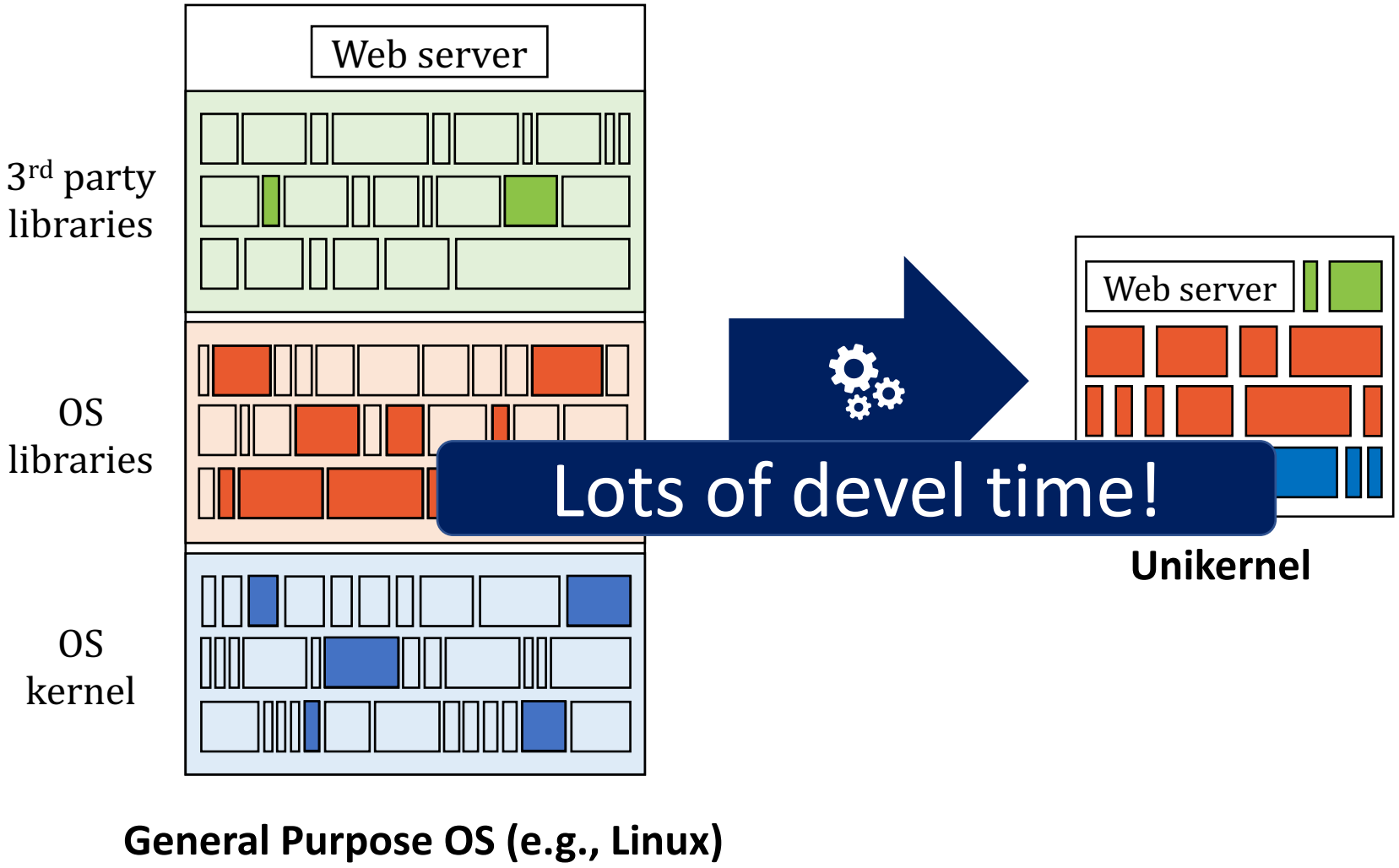
- ~300K reqs/sec nginx with a single guest CPU
  REFS*: Unikraft [Kuenzer, Eurosys 2021.* ***Best Paper Award****], Elastic CDNs [Kuenzer VEE 2017]*

Security Features

- Small trusted compute base
- Strong isolation by hypervisor
- Per-library isolation

REFS*: FlexOS [Lefeuvre HotOS 2021], CubicleOS [Sartakov ASPLOS 2021])*

Unikernels in One Slide



How do we *transparently* build efficient and POSIX-compliant unikernels?

Design Principles

1. Fully **modular** kernel

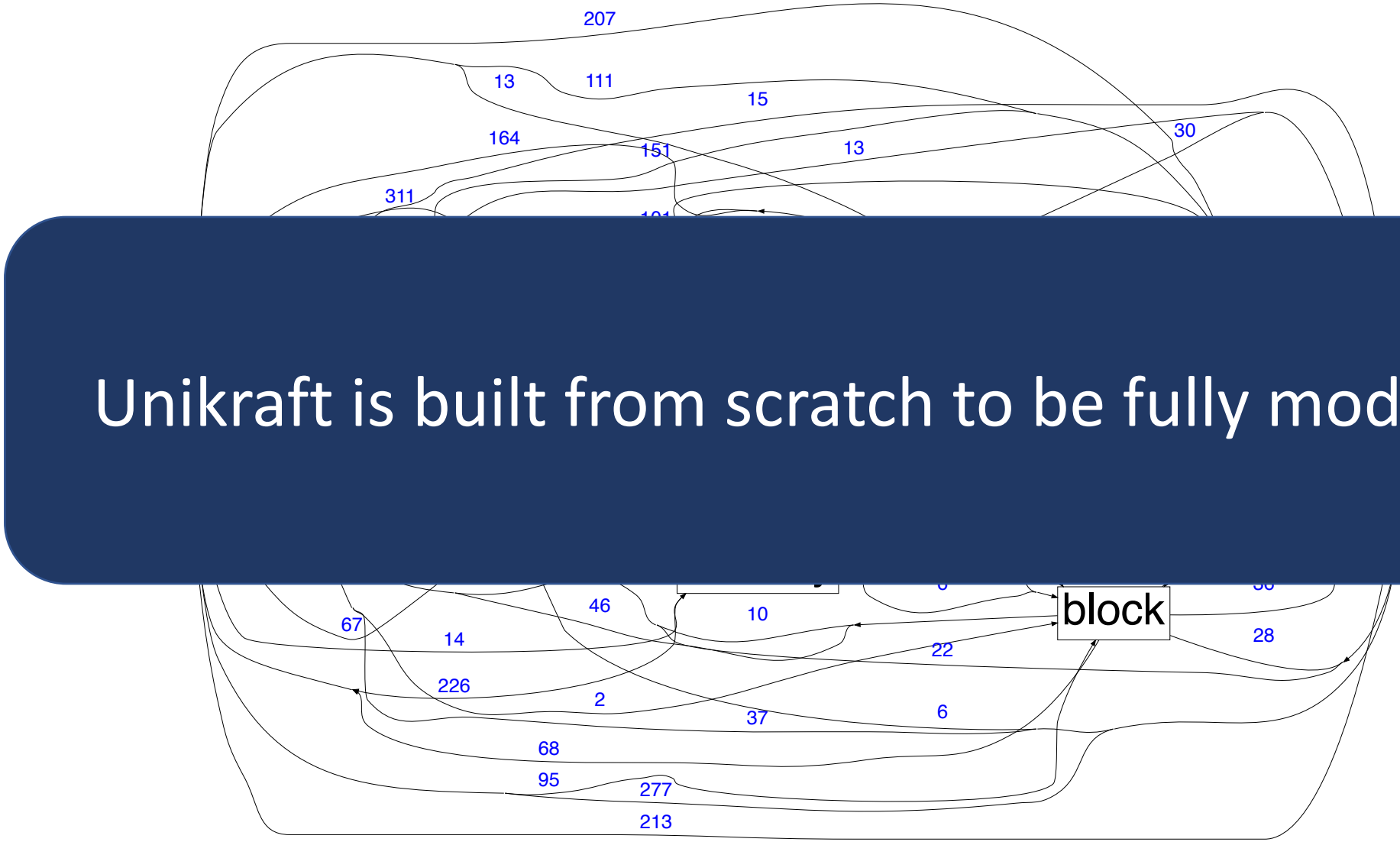
2. Provide high performance **specialized APIs**

Design Principles

1. Fully **modular** kernel

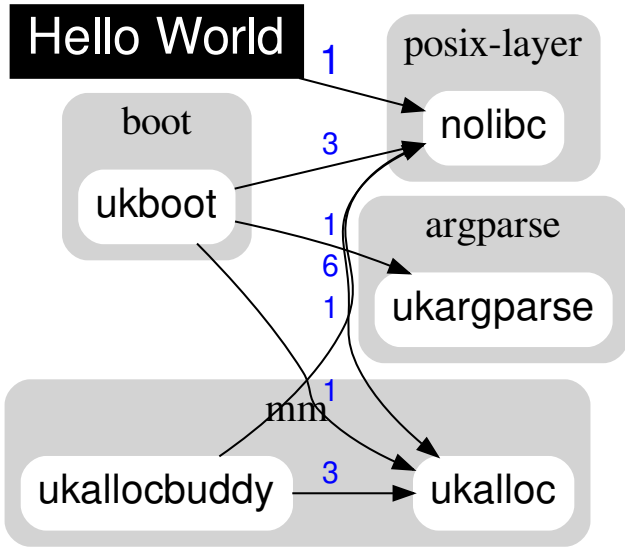
2. Provide high performance **specialized APIs**

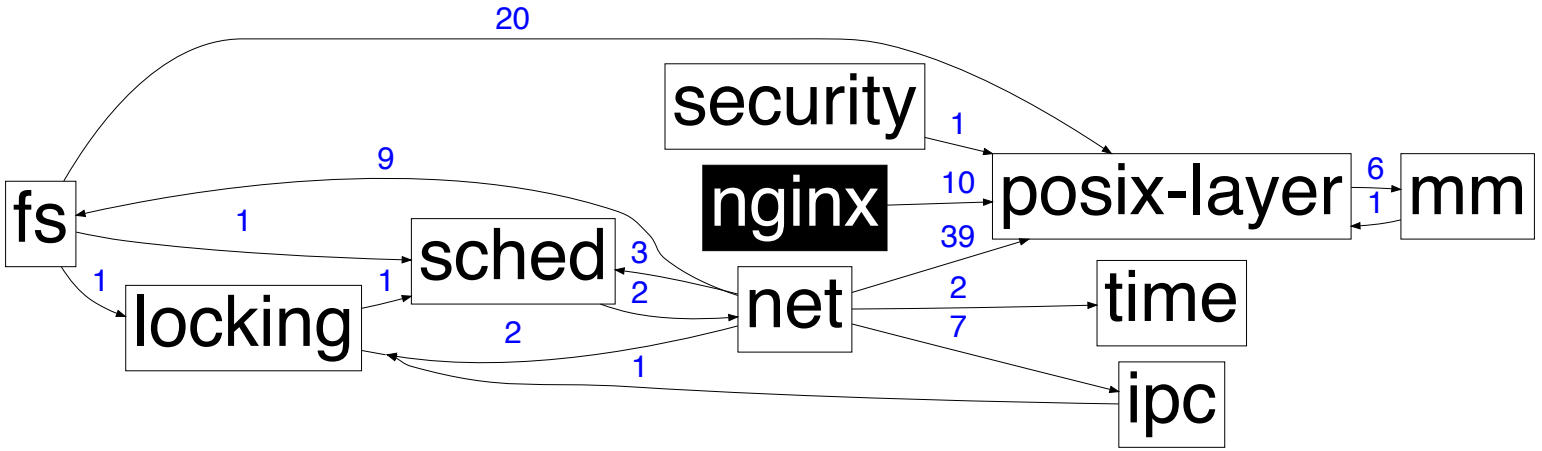
Why not Linux?



Unikraft is built from scratch to be fully modular

With Unikraft





Doing it with existing unikernels?

(1) Require significant expert work to build

Unikraft is built from scratch (with borrowing)

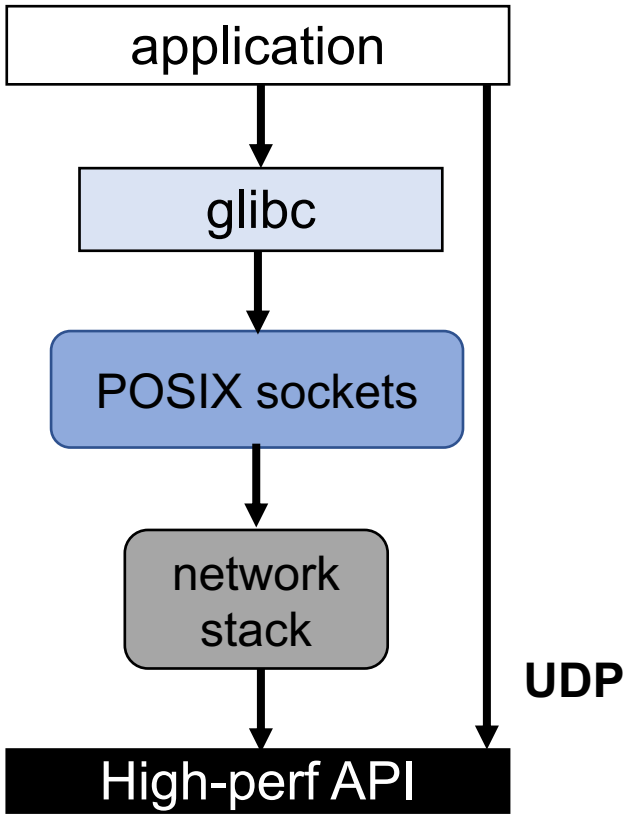
(3) The (uni)kernels are *still* monolithic

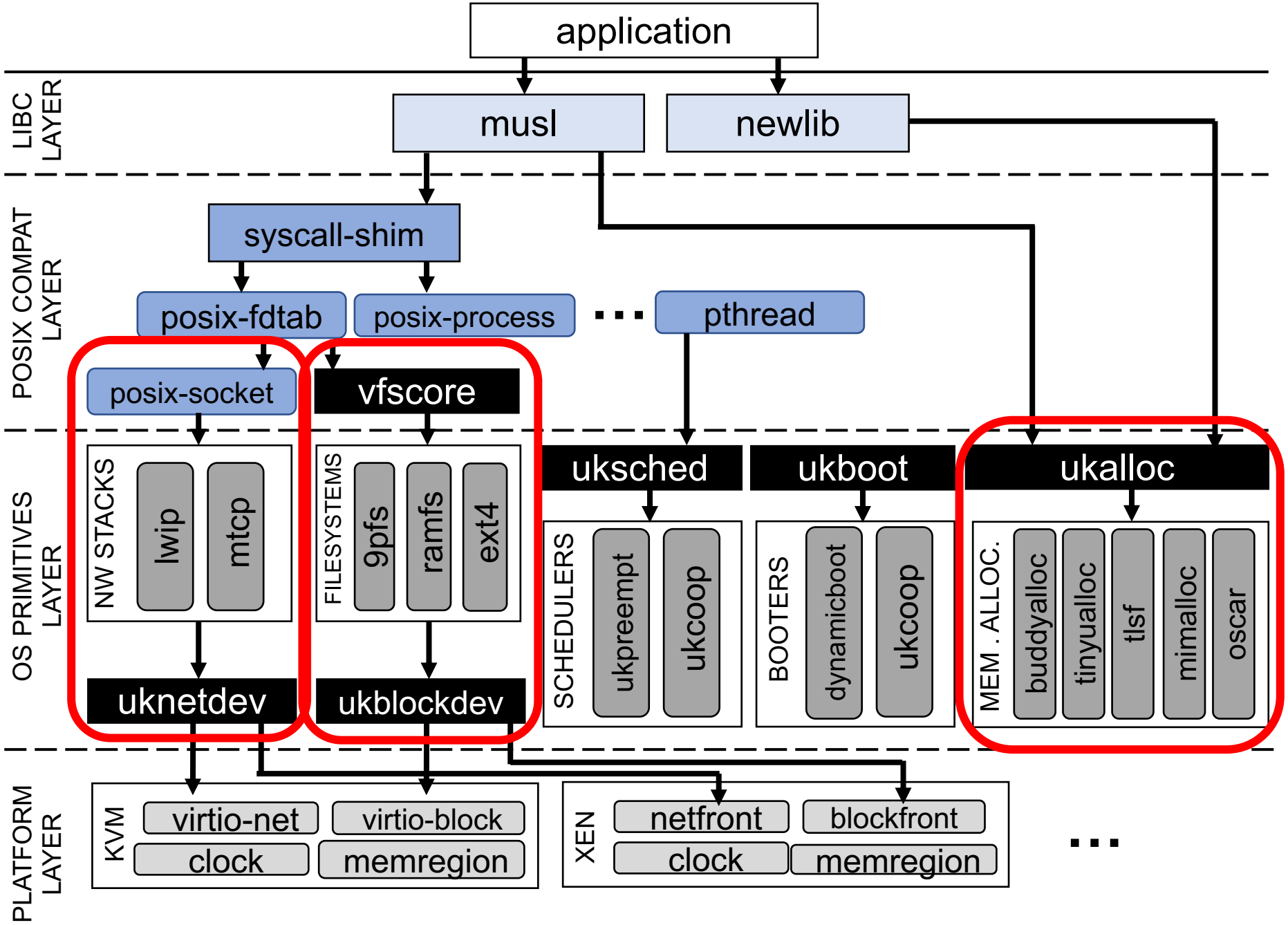
Design Principles

1. Fully **modular** kernel

2. Provide high performance **specialized APIs**

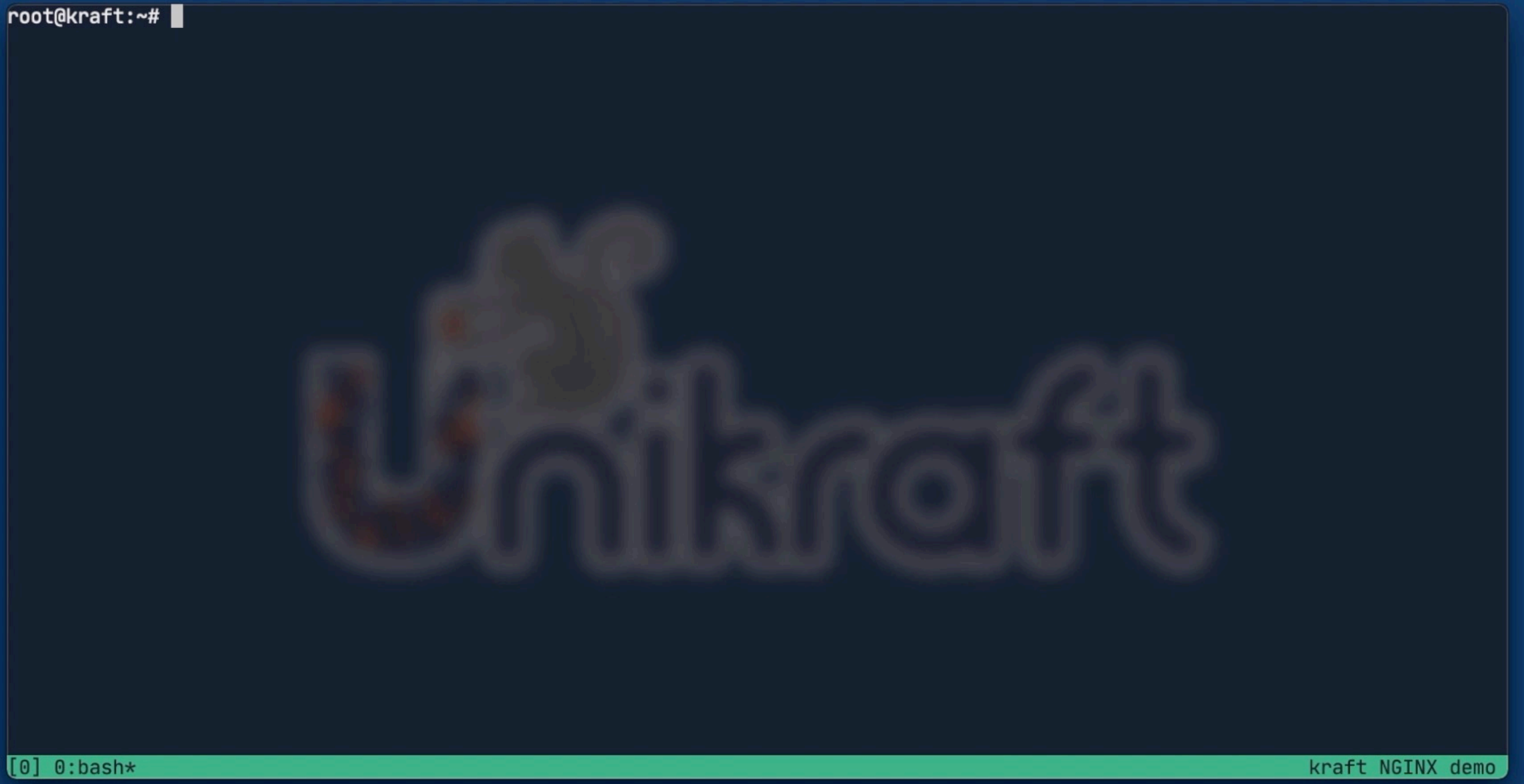
Specialized API Example




application
LIBC LAYER
musl
newlib
POSIX COMPAT LAYER
syscall-shim
posix-fdtab
posix-process
pthread
posix-socket
vfscore
OS PRIMITIVES LAYER
NW STACKS
lwip
mtcp
FILESYSTEMS
9pfs
ramfs
ext4
uksched
ukboot
ukalloc
SCHEDULERS
ukpreempt
ukcoop
BOOTERS
dynamicboot
ukcoop
MEM . ALLOC.
buddyalloc
tinyualloc
tlsf
mimalloc
oscar
uknetdev
ukblockdev
PLATFORM LAYER
KVM
virtio-net
virtio-block
clock
memregion
XEN
netfront
blockfront
clock
memregion

GOALS

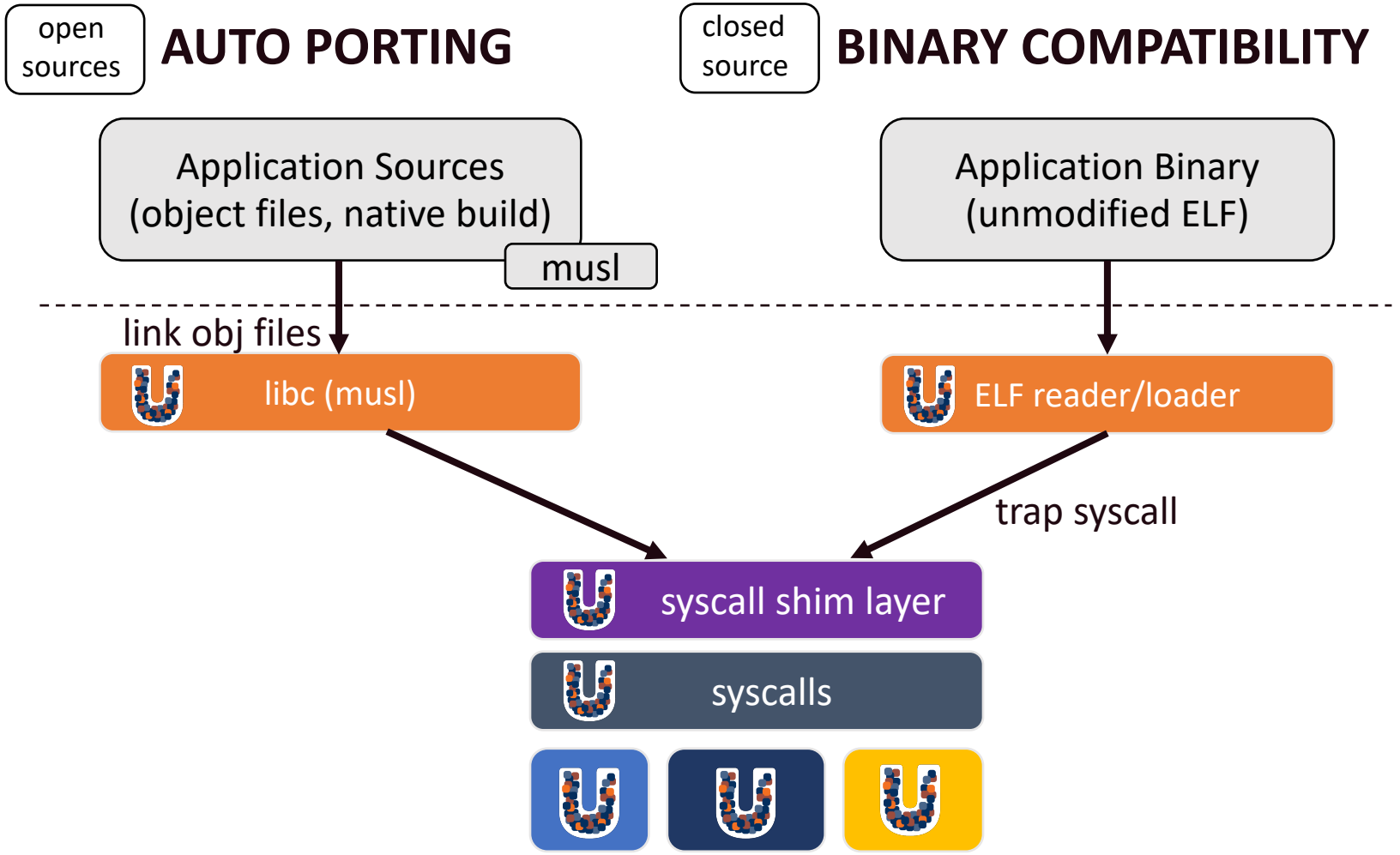
- **Easy to build and run**
- **Easy or no app porting**
- **Great performance**


root@kraft:~#
unikraft
[0] 0:bash*
kraft NGINX demo

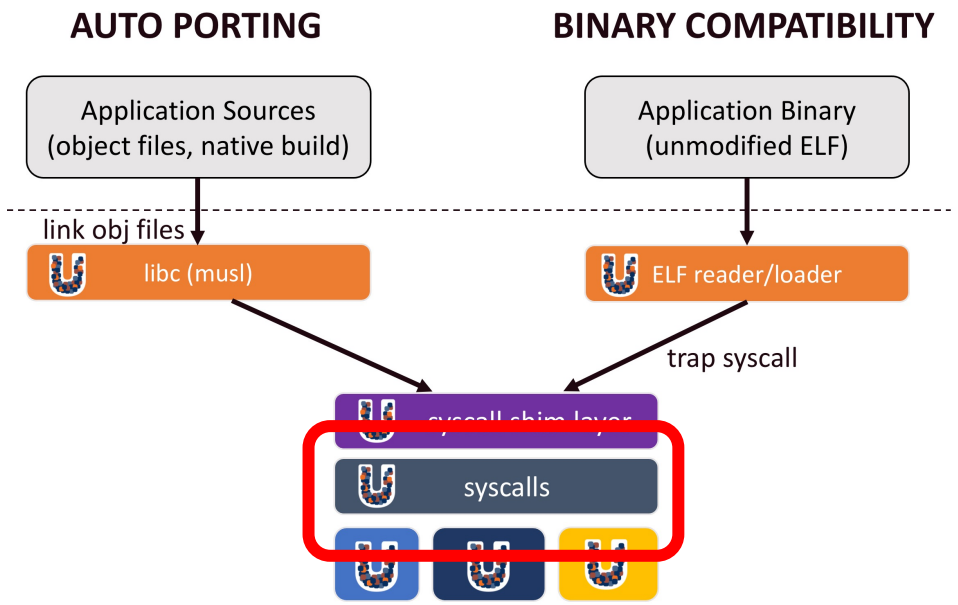
GOALS

- Easy to build and run
- **Easy or no app porting**
- Great performance

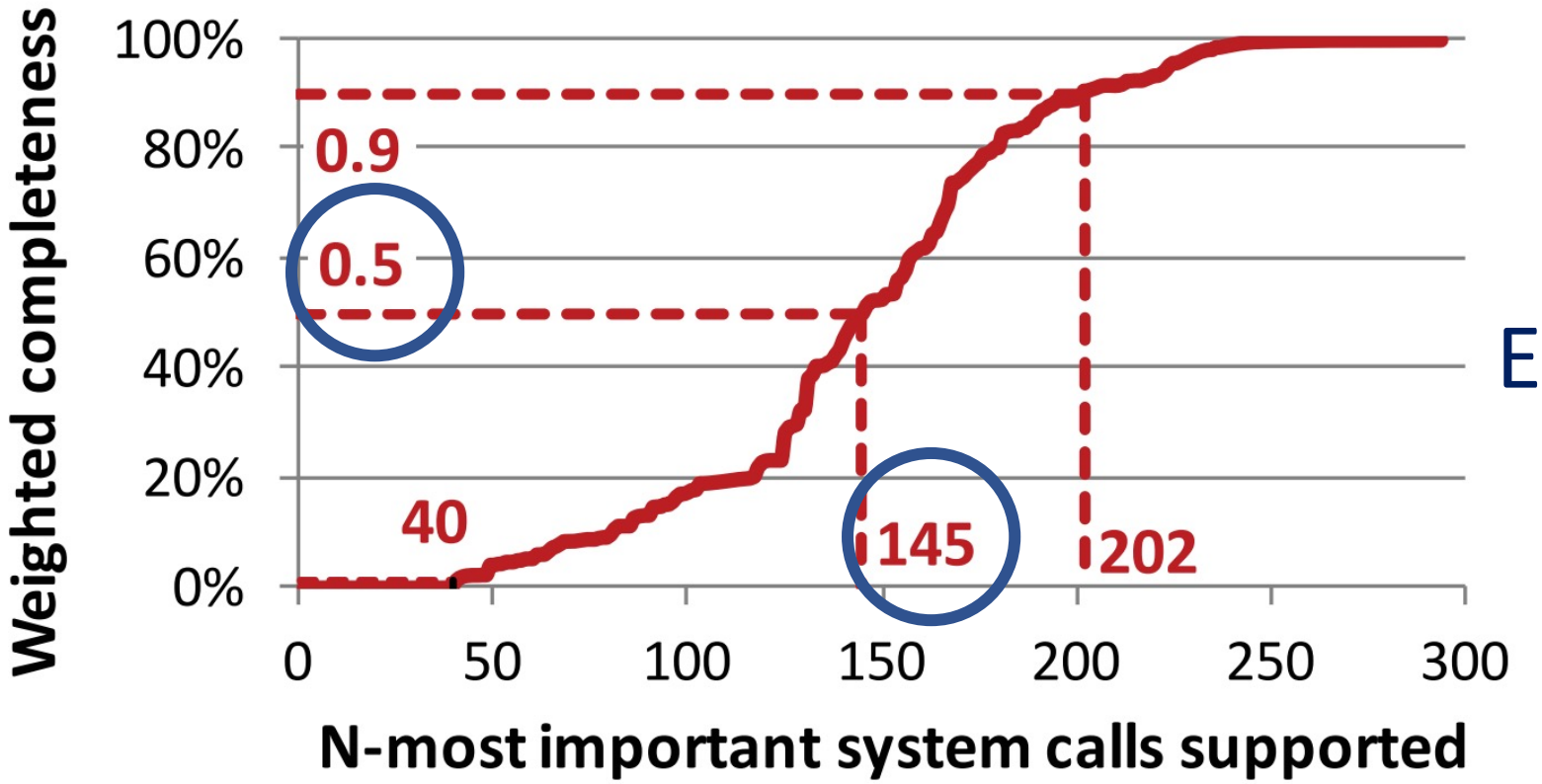
POSIX Compatibility – Two Approaches



What about syscall support?



Syscall Support

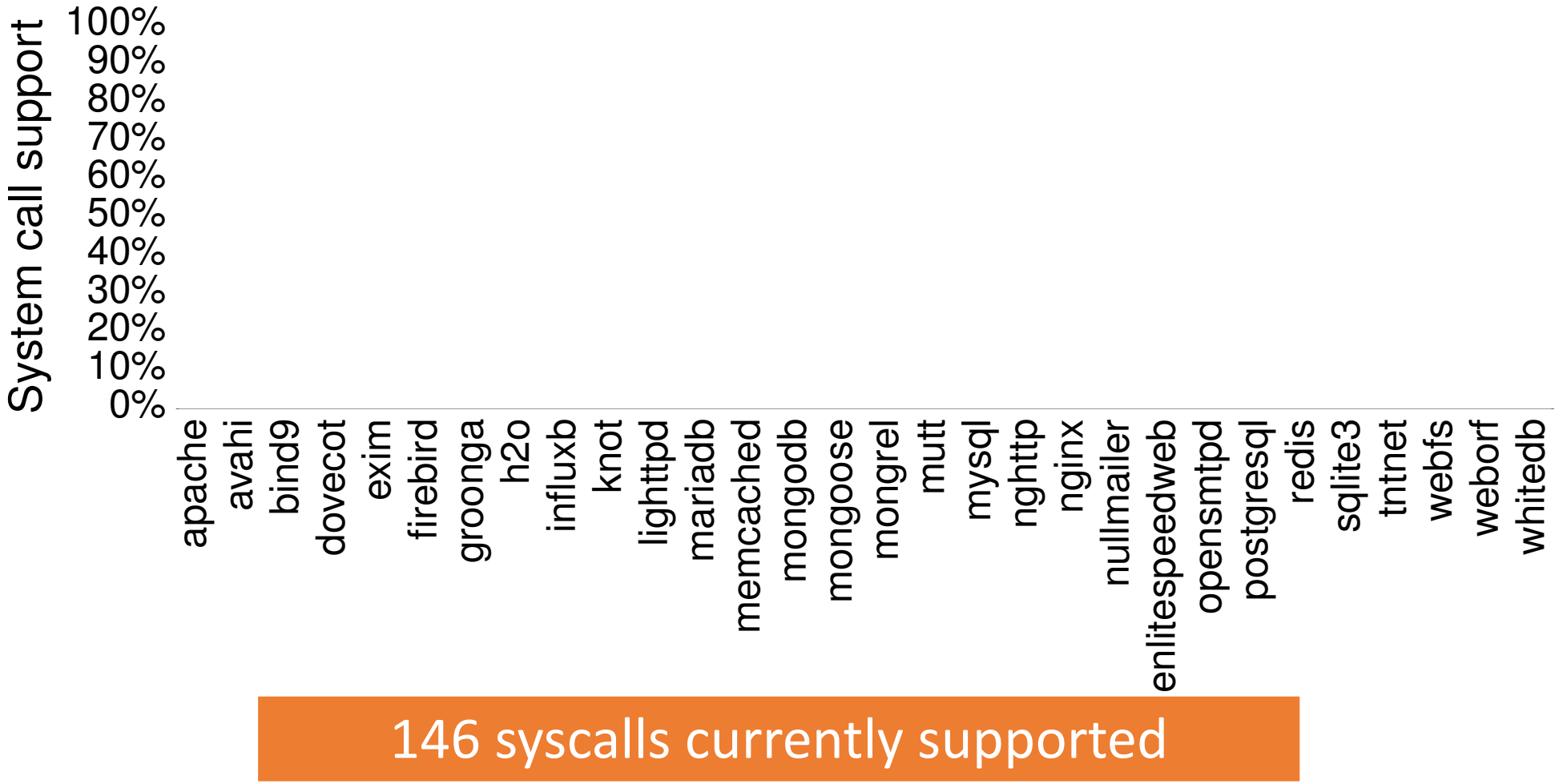


Eurosys 2016

146 syscalls currently supported

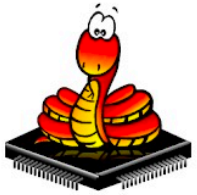
Linux: ~350 syscalls

Syscall Support
Top 30 Debian Popcon Apps



What Unikraft Supports (sample)





(ongoing)

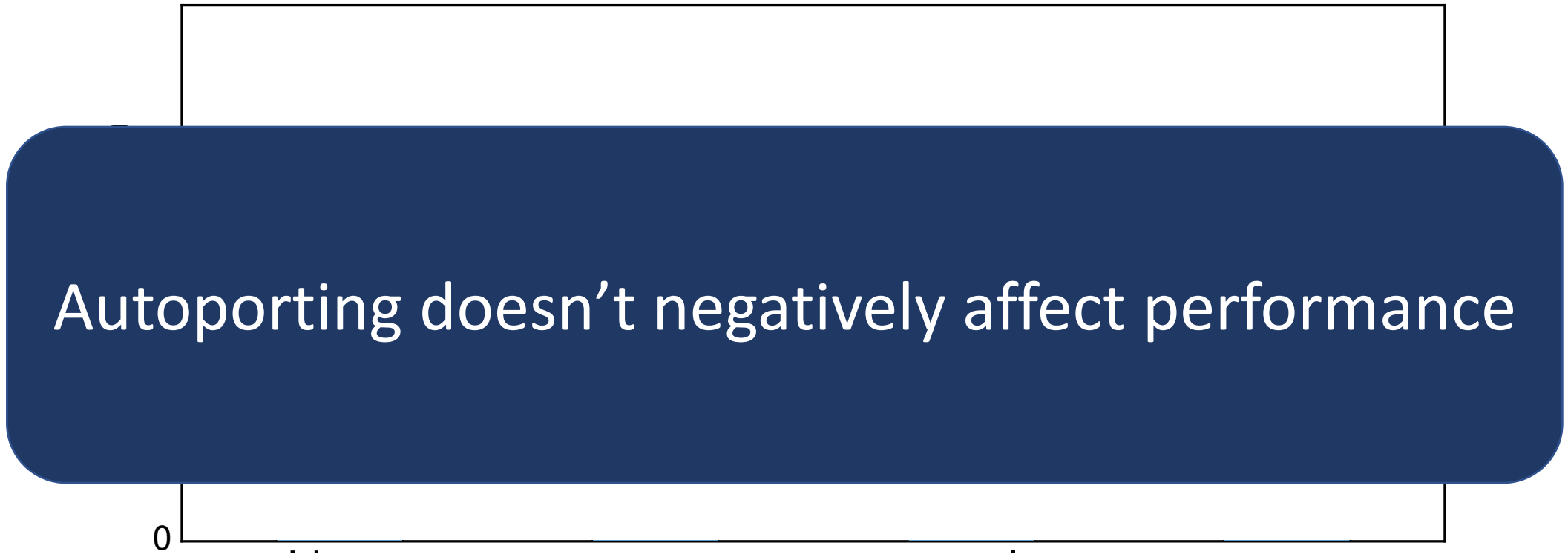
(ongoing)

GOALS

- Easy to build and run
- Easy or no app porting
- **Great performance**

Does autoporting sacrifice performance?

SQLite: Manual vs. Auto Port



Autoporting doesn’t negatively affect performance

time for 60K insertions

Transparent Benefits – Boot, Memory, Size, Throughput

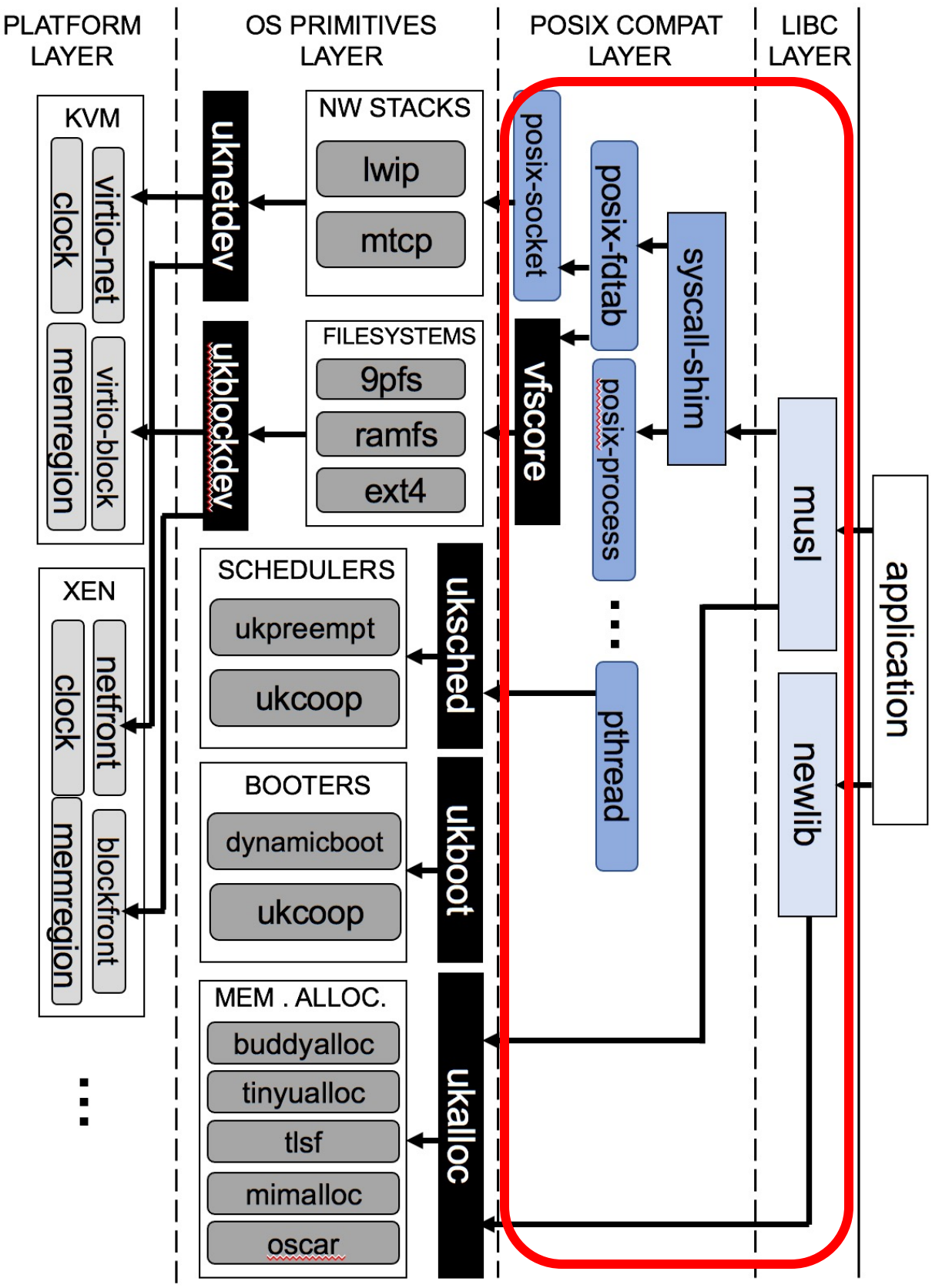
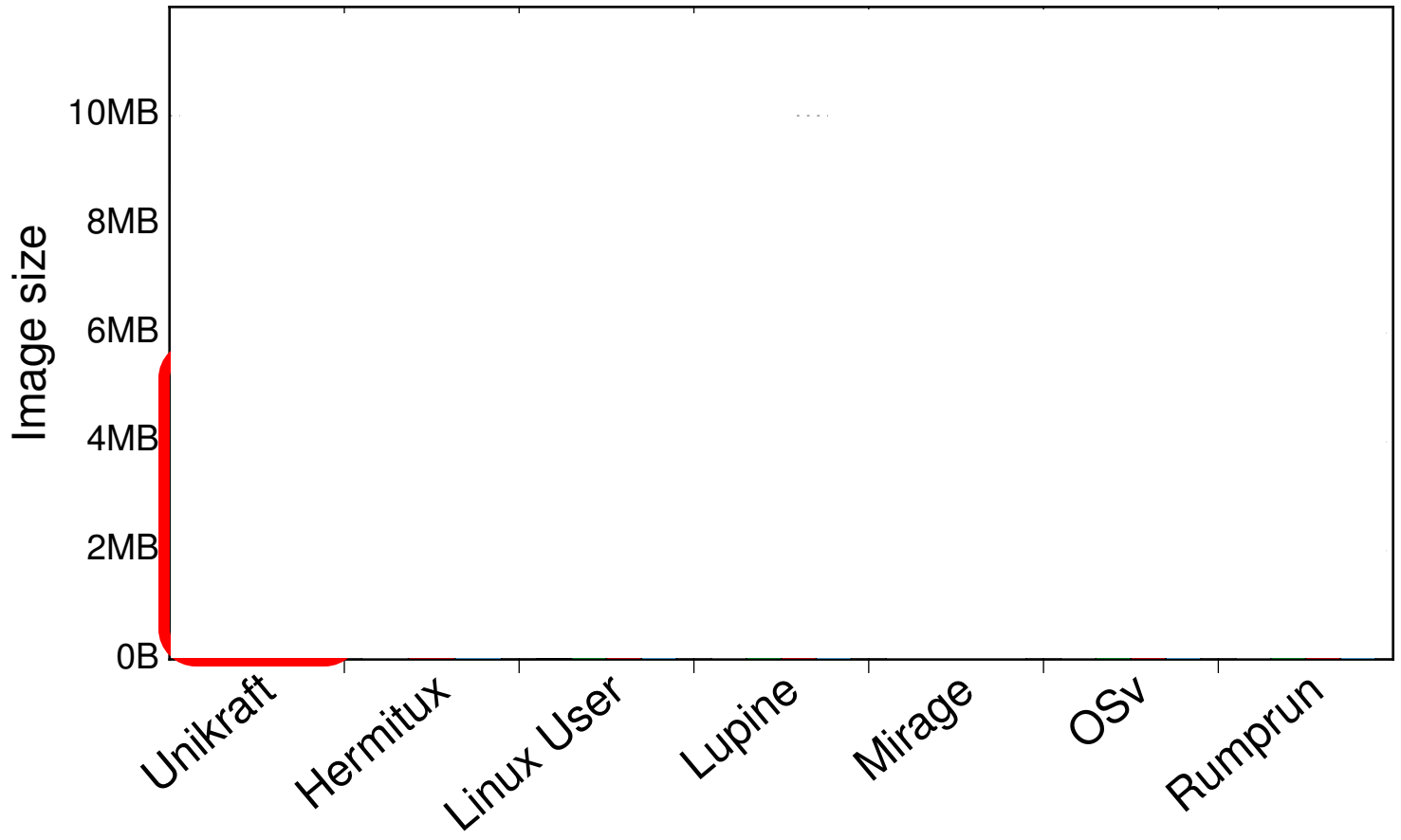
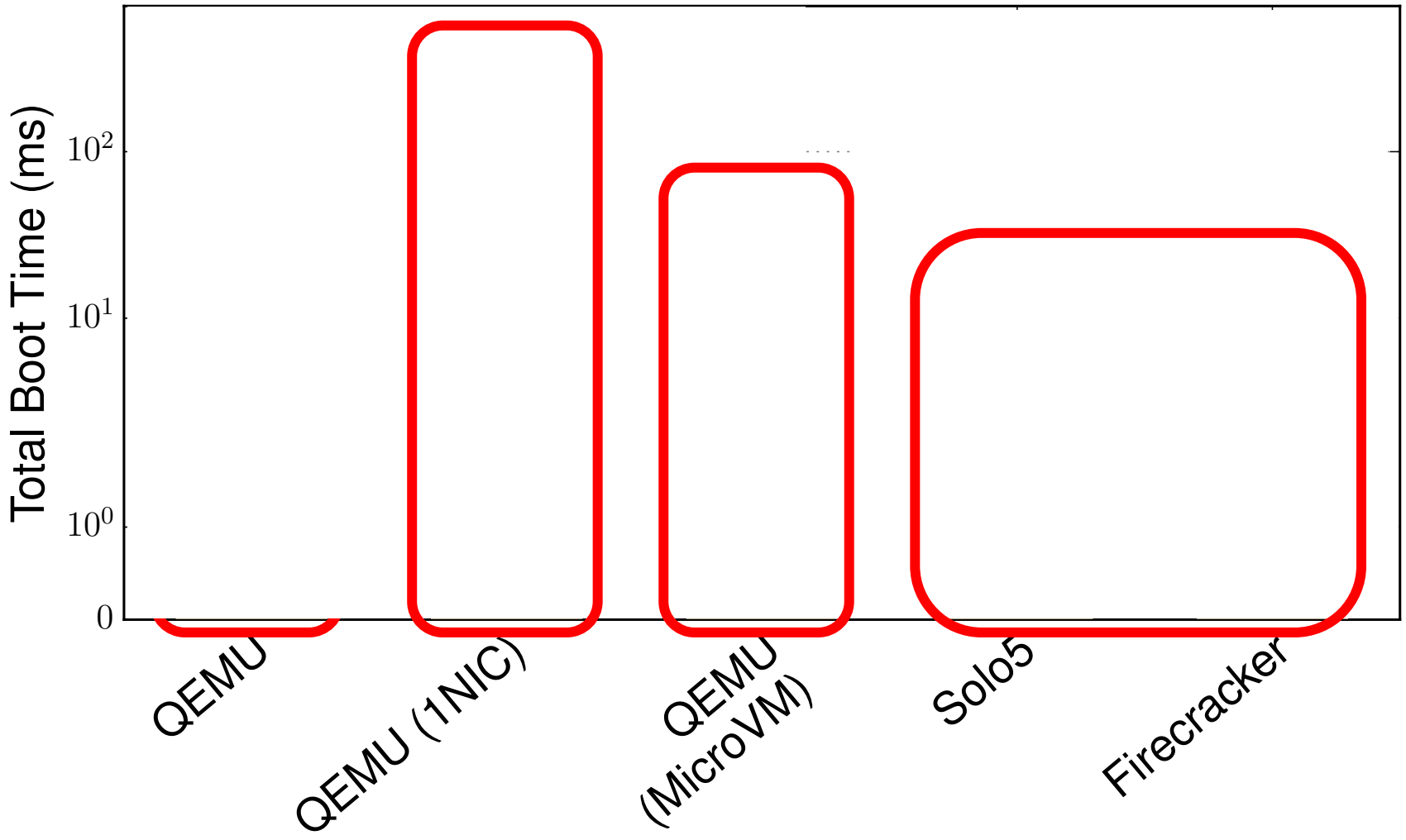


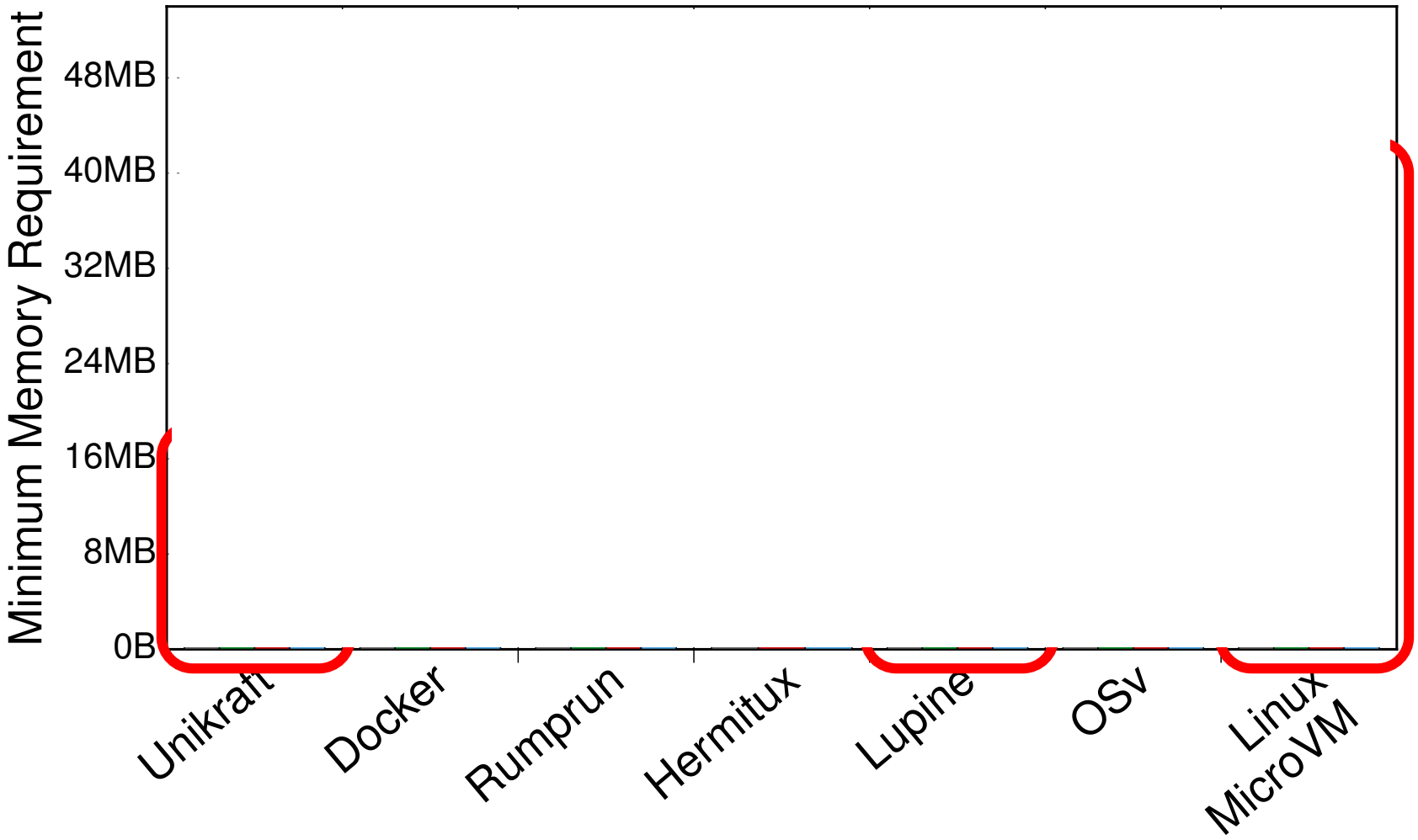
Image Sizes vs. other Projects



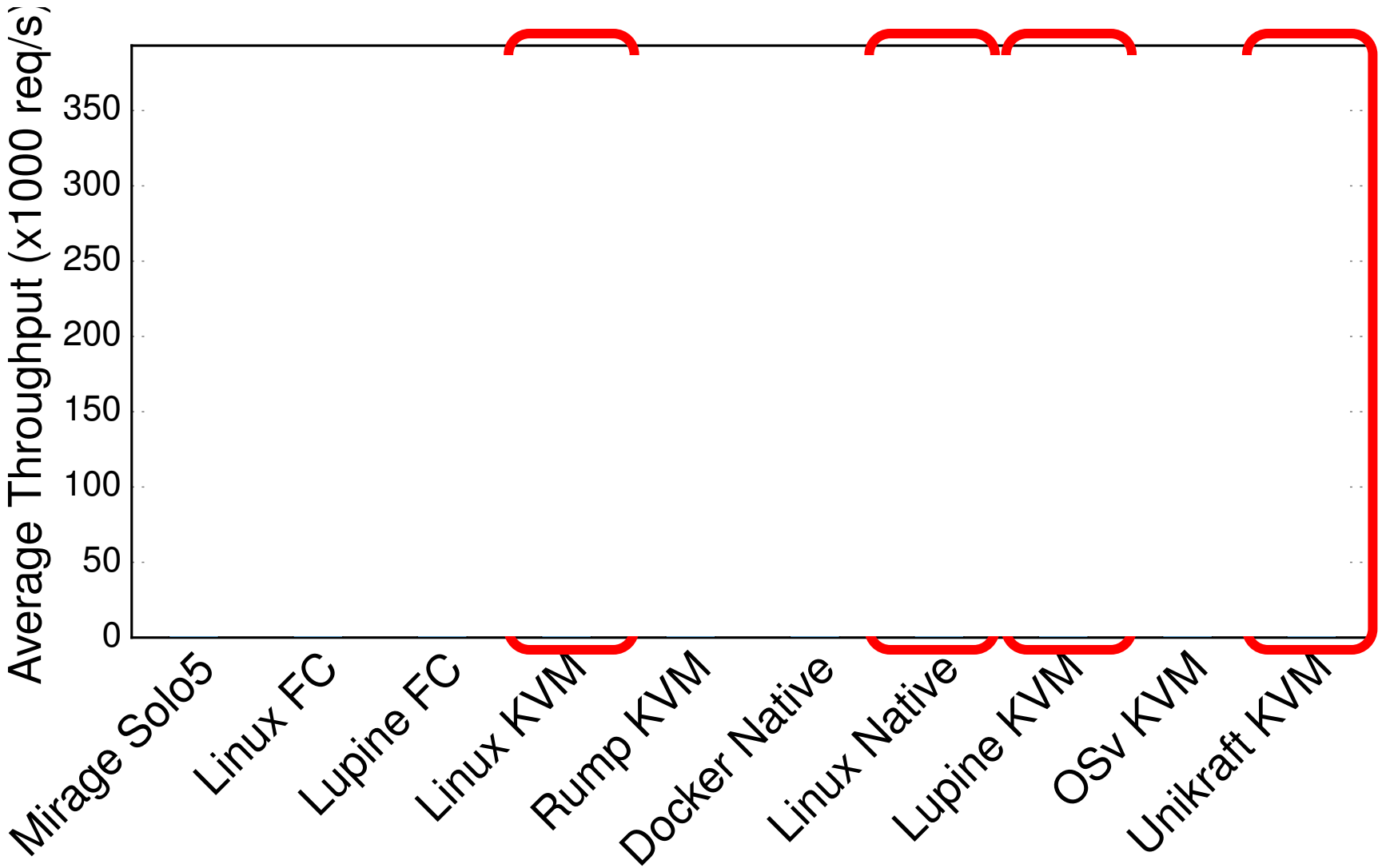
Unikraft Boot Times



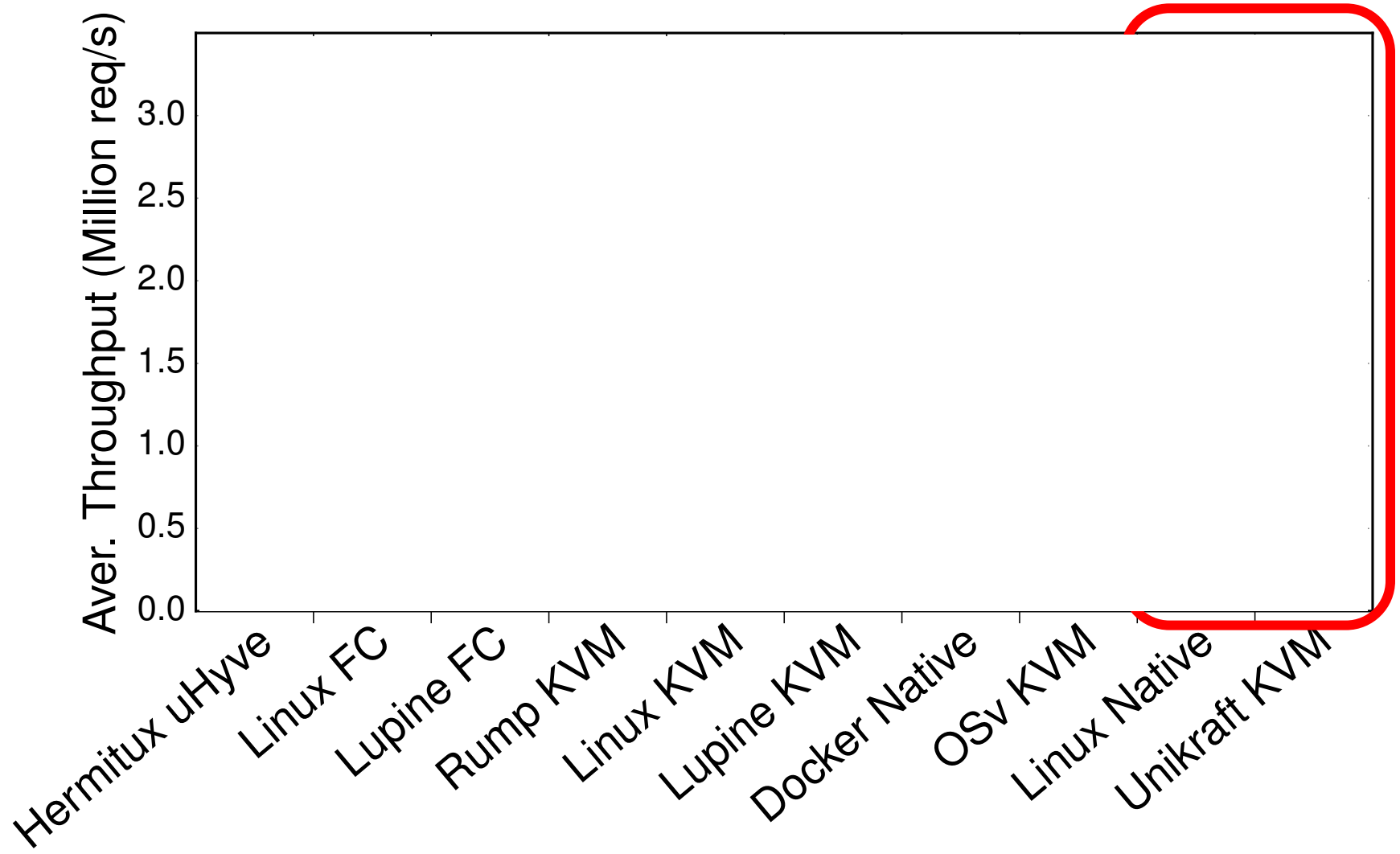
Minimum Memory Requirements



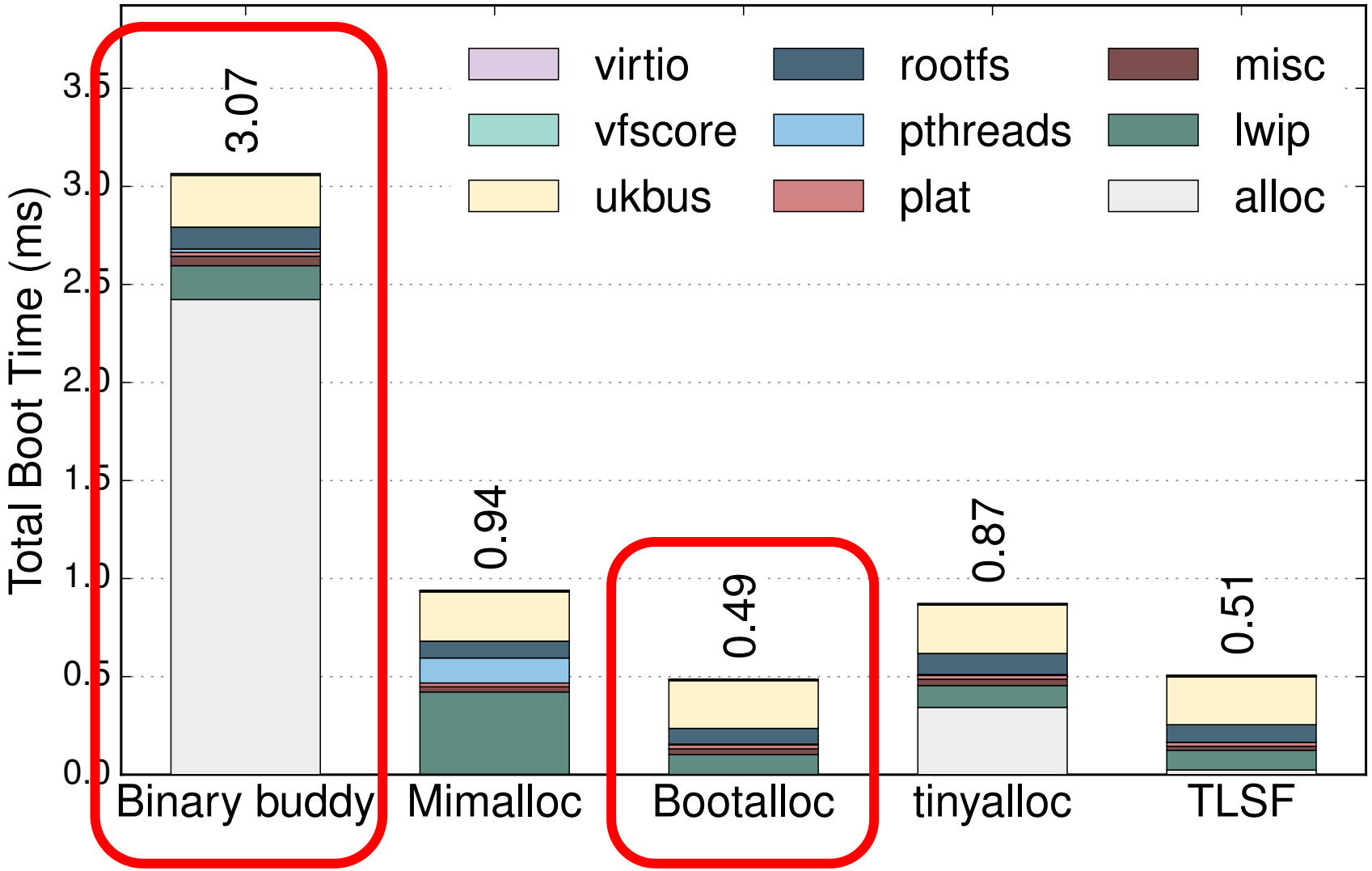
nginx Throughput



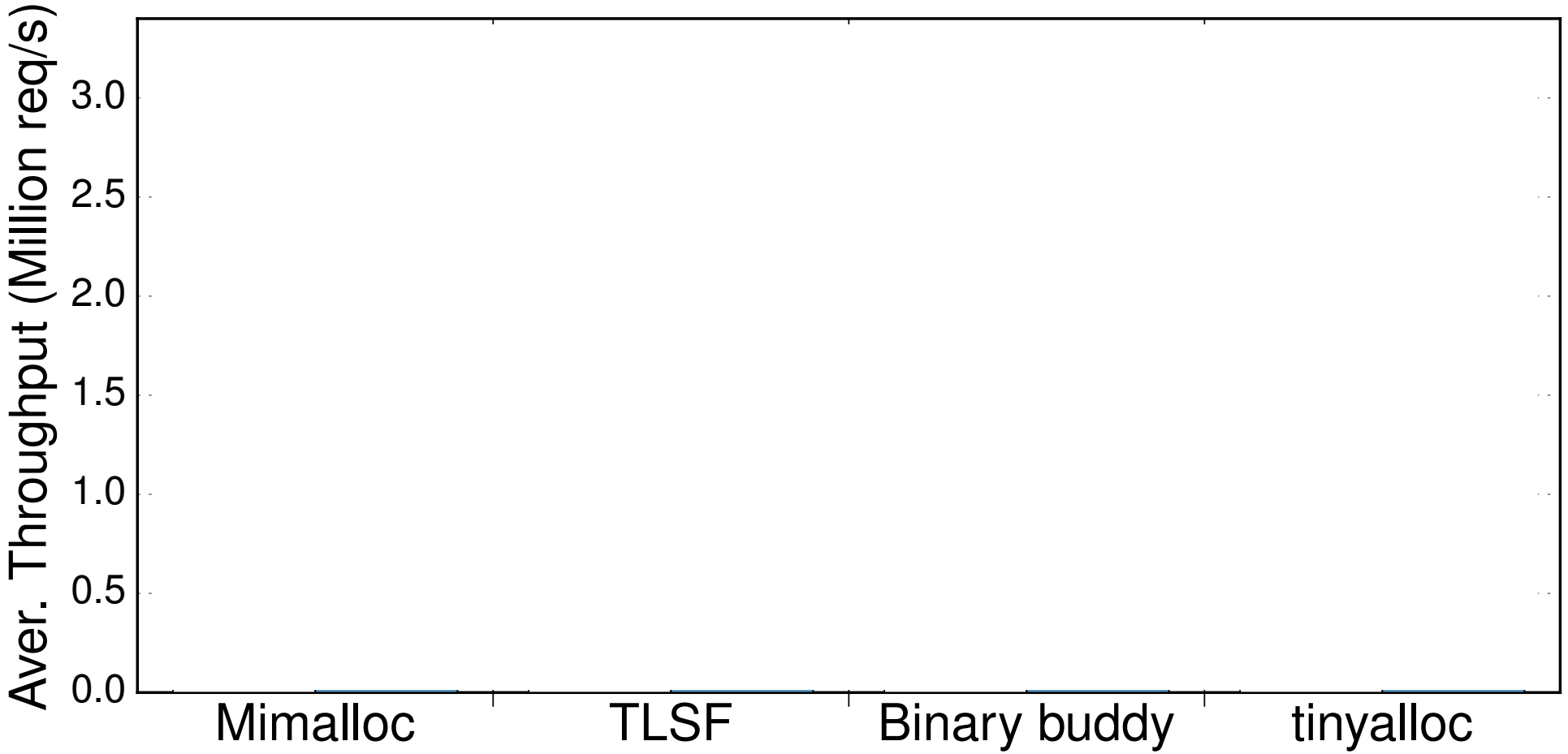
Redis Performance



Boot Times - Different Allocators

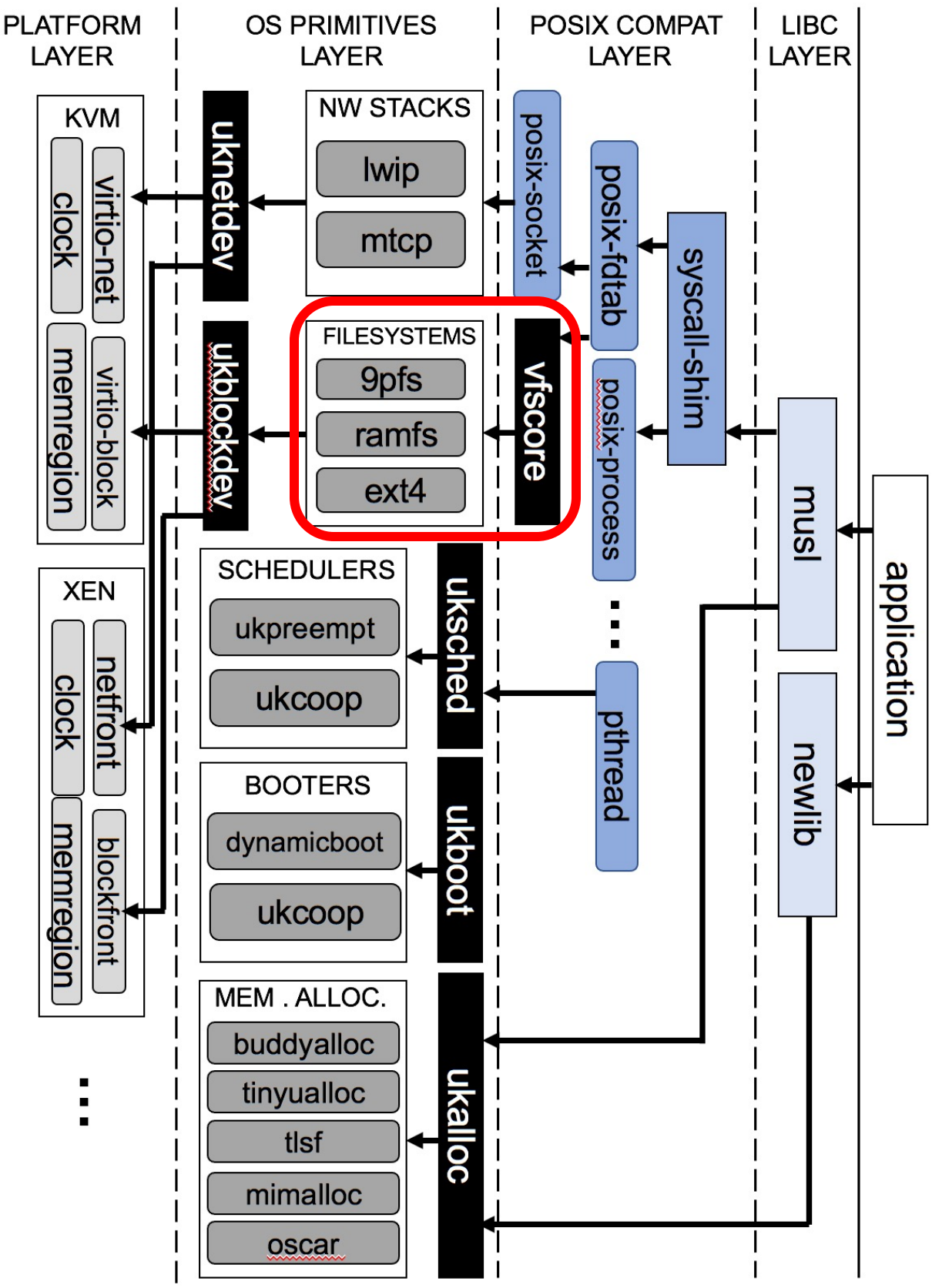


Redis Throughput
Different Allocators

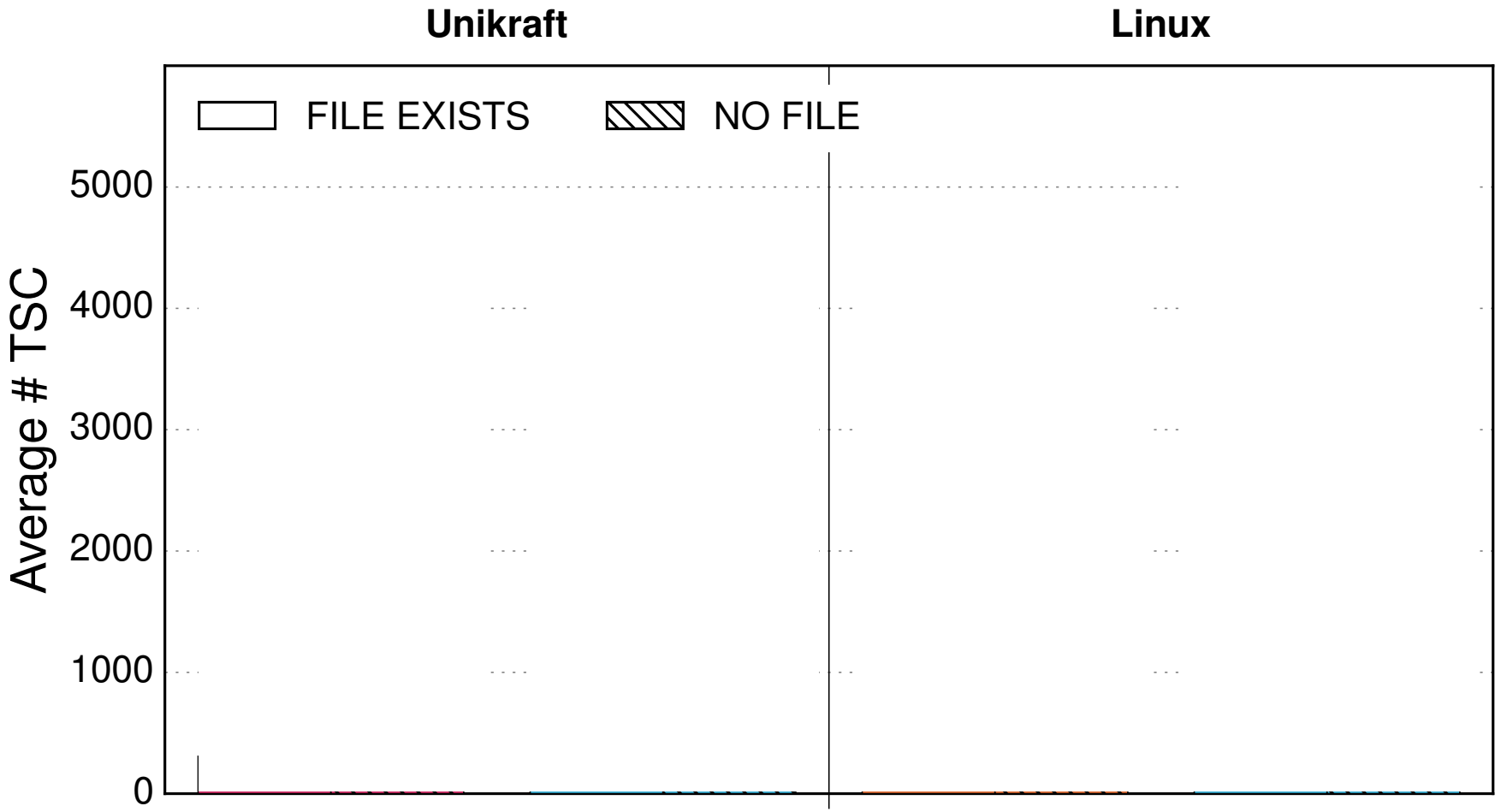


Specialized APIs

Specialization Benefits – Filesystem Performance



Filesystem Specialization – SHFS



on the

aws marketplace

Categories | Delivery Methods | Solutions | AWS IQ | Resources | Your Saved List | Partners | Sell in AWS Marketplace

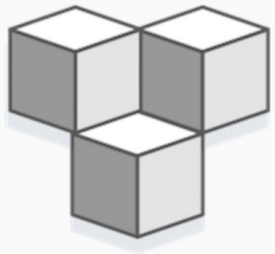
Popular Categories


Operating Systems


Security


Networking


Storage


View all categories


Data Analytics


Dev Ops


Machine Learning


Data Products

Explore solutions by topic area and use case

Business Applications

Find, buy, and deploy SaaS solutions to drive your business forward.

LEARN MORE

Data & analytics

Gain the decision-making power and insights to act in real time.

LEARN MORE

DevOps

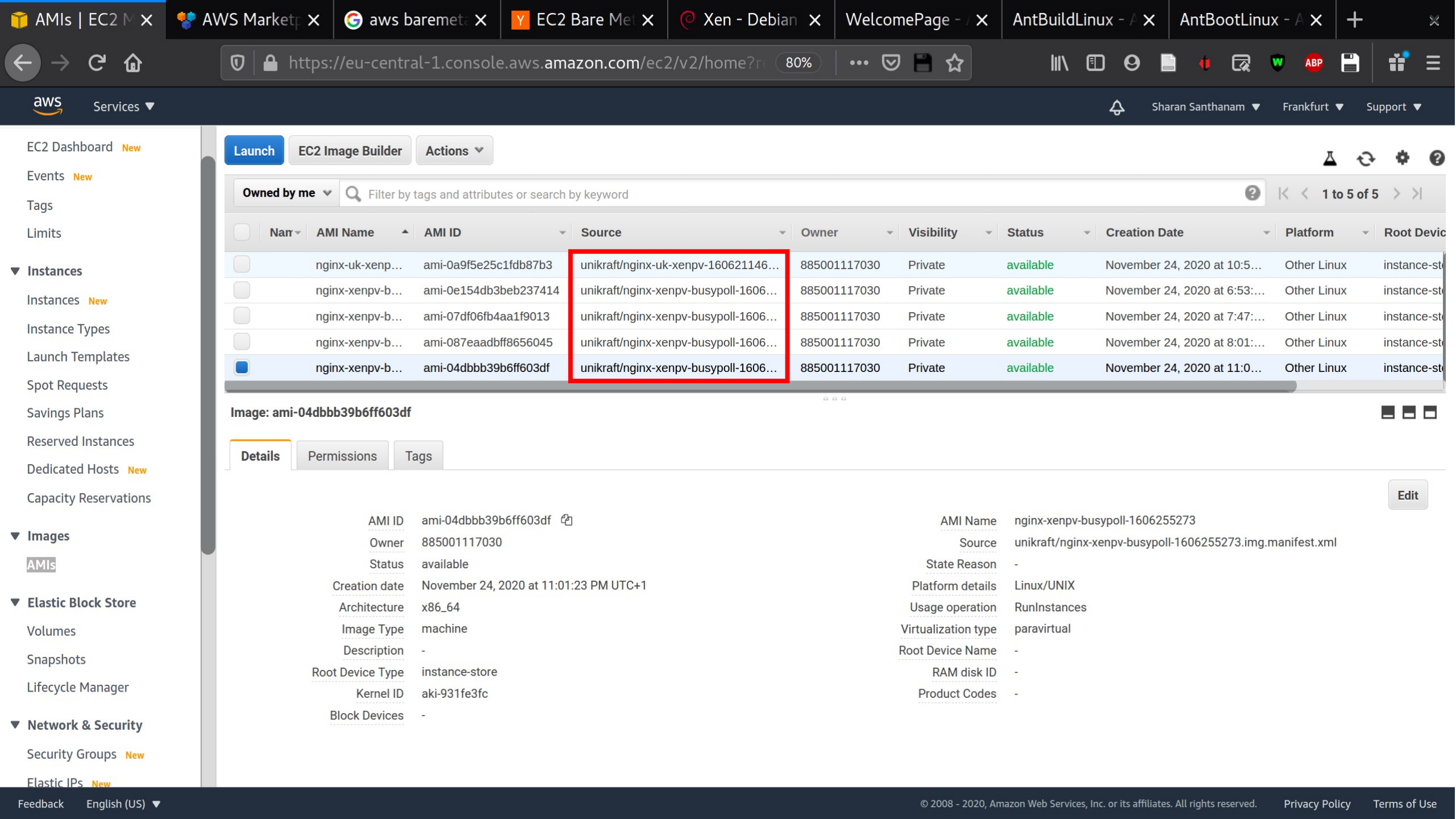
Design and deploy reliable, secure modern applications and infrastructure.

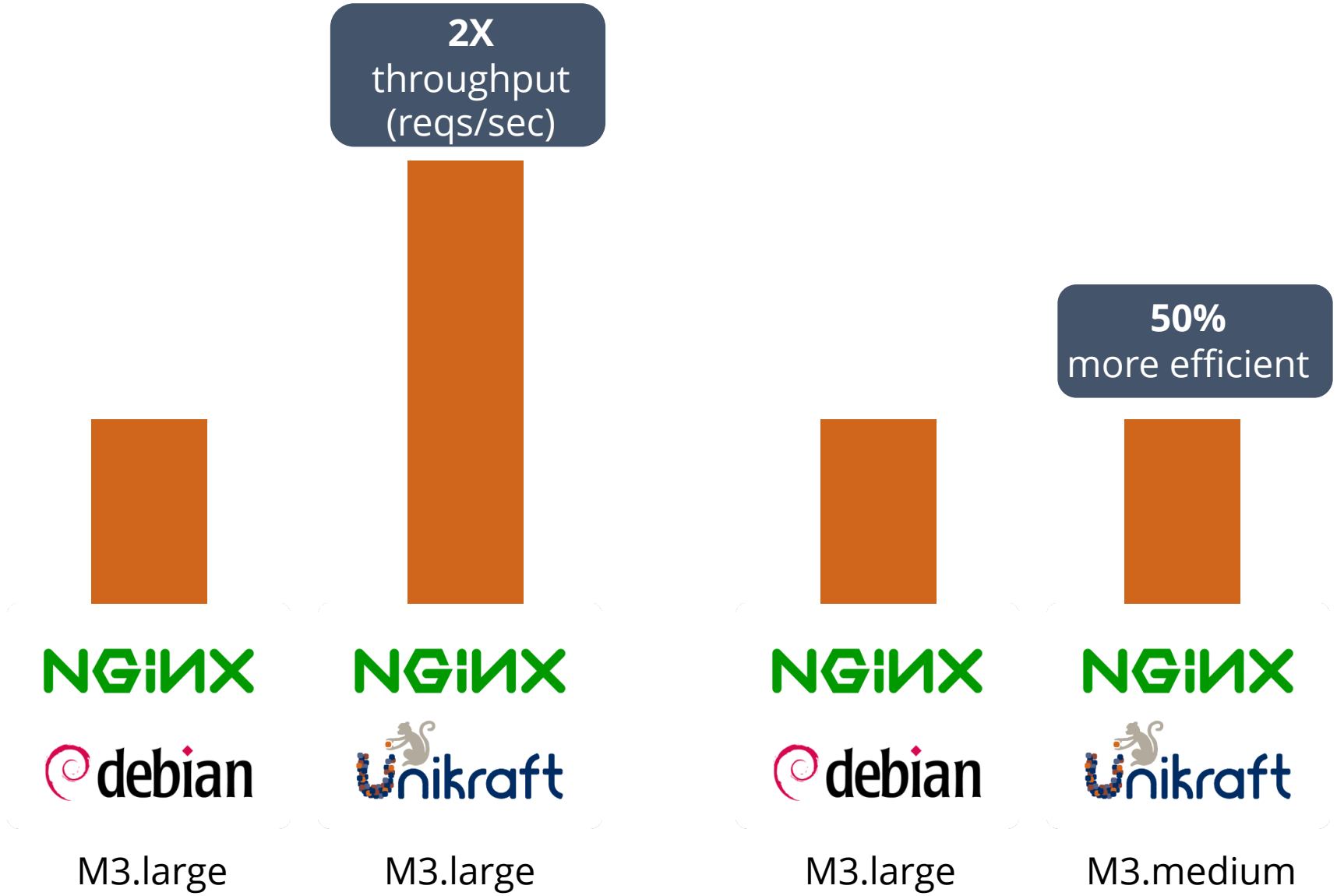
LEARN MORE

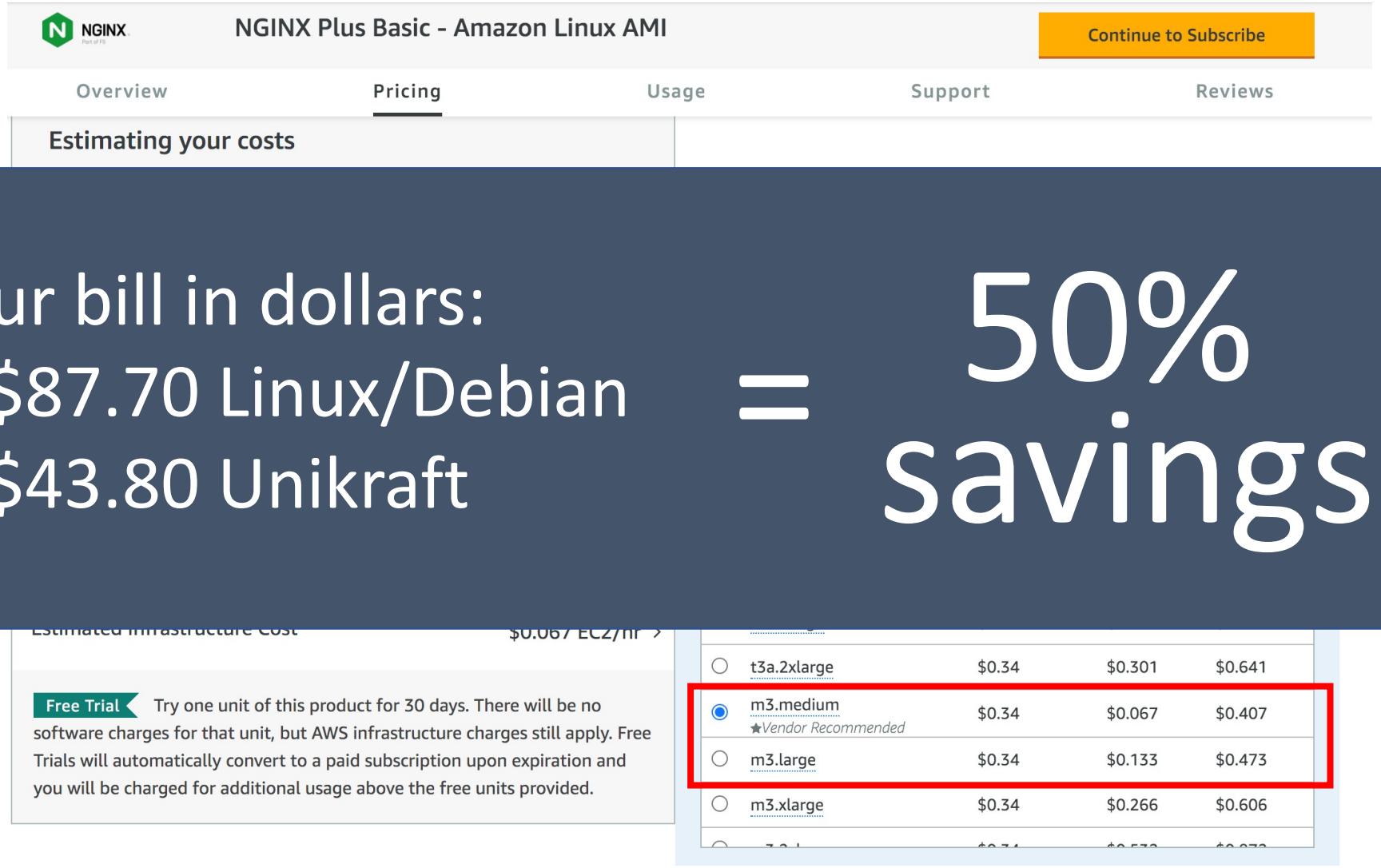
Infrastructure software

Move, modernize, and manage your IT Infrastructure.

LEARN MORE


AMIs | EC2
AWS Market
aws baremet
EC2 Bare Met
Xen - Debian
WelcomePage -
AntBuildLinux -
AntBootLinux -
https://eu-central-1.console.aws.amazon.com/ec2/v2/home?r
80%
aws
Services
Sharan Santhanam
Frankfurt
Support
EC2 Dashboard New
Events New
Tags
Limits
Instances
Instances New
Instance Types
Launch Templates
Spot Requests
Savings Plans
Reserved Instances
Dedicated Hosts New
Capacity Reservations
Images
AMIs
Elastic Block Store
Volumes
Snapshots
Lifecycle Manager
Network & Security
Security Groups New
Elastic IPs New
Launch
EC2 Image Builder
Actions
Owned by me
Filter by tags and attributes or search by keyword
1 to 5 of 5
Nam
AMI Name
AMI ID
Source
Owner
Visibility
Status
Creation Date
Platform
Root Devic
nginx-uk-xenp...
ami-0a9f5e25c1fdb87b3
unikraft/nginx-uk-xenpv-160621146...
885001117030
Private
available
November 24, 2020 at 10:5...
Other Linux
instance-st
nginx-xenpv-b...
ami-0e154db3beb237414
unikraft/nginx-xenpv-busypoll-1606...
885001117030
Private
available
November 24, 2020 at 6:53:...
Other Linux
instance-st
nginx-xenpv-b...
ami-07df06fb4aa1f9013
unikraft/nginx-xenpv-busypoll-1606...
885001117030
Private
available
November 24, 2020 at 7:47:...
Other Linux
instance-st
nginx-xenpv-b...
ami-087eaadbff8656045
unikraft/nginx-xenpv-busypoll-1606...
885001117030
Private
available
November 24, 2020 at 8:01:...
Other Linux
instance-st
nginx-xenpv-b...
ami-04dbbb39b6ff603df
unikraft/nginx-xenpv-busypoll-1606...
885001117030
Private
available
November 24, 2020 at 11:0...
Other Linux
instance-st
Image: ami-04dbbb39b6ff603df
Details
Permissions
Tags
Edit
AMI ID
ami-04dbbb39b6ff603df
Owner
885001117030
Status
available
Creation date
November 24, 2020 at 11:01:23 PM UTC+1
Architecture
x86_64
Image Type
machine
Description
-
Root Device Type
instance-store
Kernel ID
aki-931fe3fc
Block Devices
-
AMI Name
nginx-xenpv-busypoll-1606255273
Source
unikraft/nginx-xenpv-busypoll-1606255273.img.manifest.xml
State Reason
-
Platform details
Linux/UNIX
Usage operation
RunInstances
Virtualization type
paravirtual
Root Device Name
-
RAM disk ID
-
Product Codes
-
Feedback
English (US)
© 2008 - 2020, Amazon Web Services, Inc. or its affiliates. All rights reserved.
Privacy Policy
Terms of Use


2X
throughput
(reqs/sec)
50%
more efficient
NGINX
debian
M3.large
NGINX
Unikraft
M3.large
NGINX
debian
M3.large
NGINX
Unikraft
M3.medium


NGINX
NGINX Plus Basic - Amazon Linux AMI
Continue to Subscribe
Overview
Pricing
Usage
Support
Reviews
Estimating your costs
Free Trial
Try one unit of this product for 30 days. There will be no software charges for that unit, but AWS infrastructure charges still apply. Free Trials will automatically convert to a paid subscription upon expiration and you will be charged for additional usage above the free units provided.
t3a.2xlarge $0.34 $0.301 $0.641
m3.medium $0.34 $0.067 $0.407
★Vendor Recommended
m3.large $0.34 $0.133 $0.473
m3.xlarge $0.34 $0.266 $0.606
Our bill in dollars:
• $87.70 Linux/Debian
• $43.80 Unikraft
= 50% savings

Xen-based

KVM-based

KVM-based

Hyper-V (upcoming)

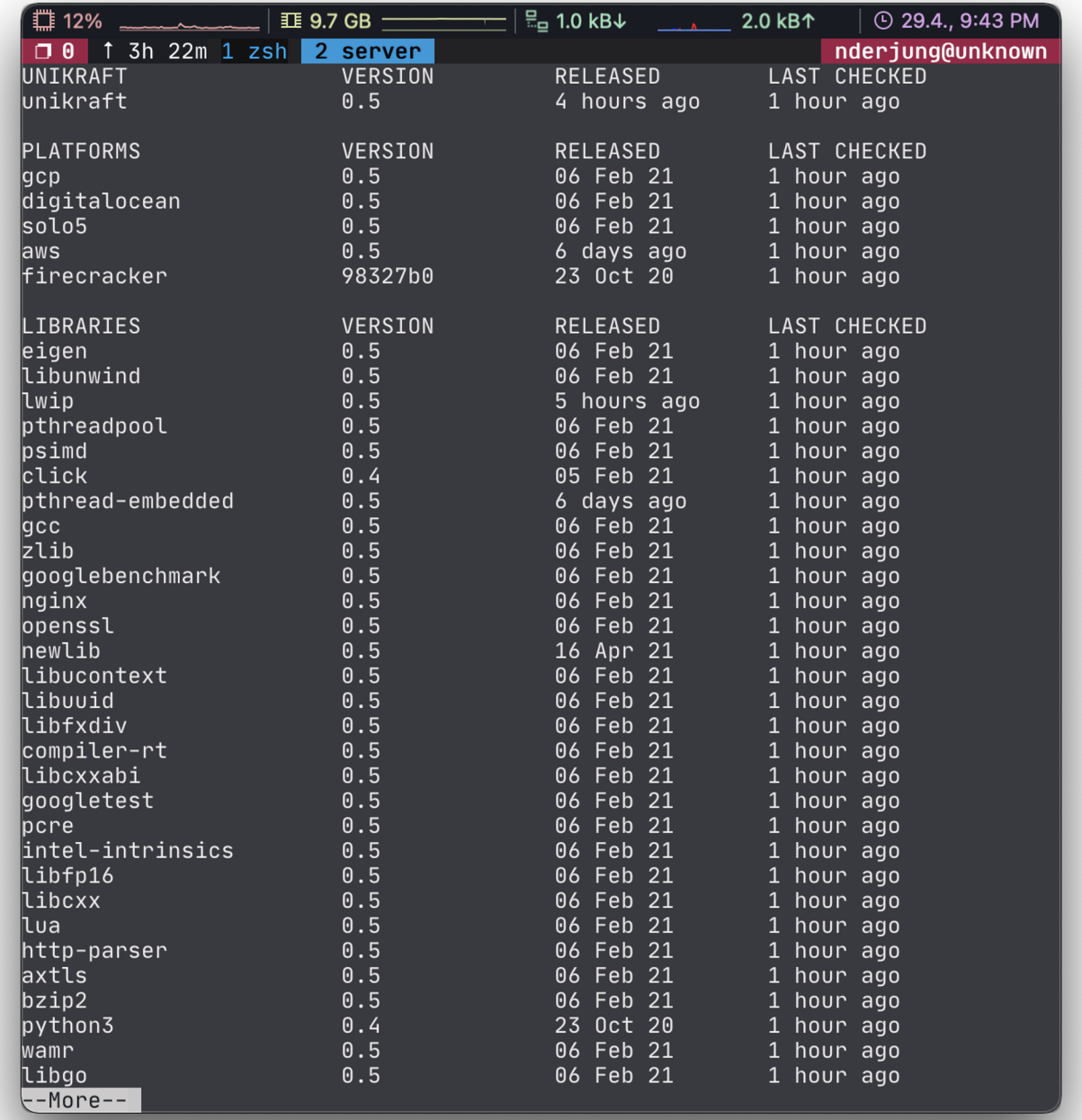
Unikraft on GCP

```
SeaBIOS (version 1.8.2-20190620_103534-google)
Total RAM Size = 0x0000000026600000 = 614 MiB
CPUs found: 1     Max CPUs supported: 1
found virtio-scsi at 0:3
virtio-scsi vendor='Google'  product='PersistentDisk' rev='1' type=0 removable=0
virtio-scsi blksize=512 sectors=2097152 = 1024 MiB
drive 0x000f2770: PCHS=0/0/0 translation=lba LCHS=1024/32/63 s=2097152
Booting from Hard Disk 0...
[    0.914853] ERR:  [libkvmvirtio] virtio_bus.c @ 143  : Failed to find the driver for the virtio device 0x1fc008 (id:8)
[    0.920038] ERR:  [libkvmvirtio] virtio_pci.c @ 391  : Failed to register the virtio device: -14
[    0.923965] ERR:  [libkvmpci] pci_bus.c @ 215  : PCI 00:03.00: Failed to initialize device driver
en0: Added
en0: Interface is up
Welcome to  _ __            _____
 __ _____  (_) /__ _______ _/ _/ /_
/ // / _ \/ /  '_// __/ _ `/ _/ __/
\_,_/_//_/_/_/\_\/_/  \_,_/_/ \__/
       Iapetus 0.3.1~94e6454-custom
Listening on port 8888...
```

Ecosystem - kraft

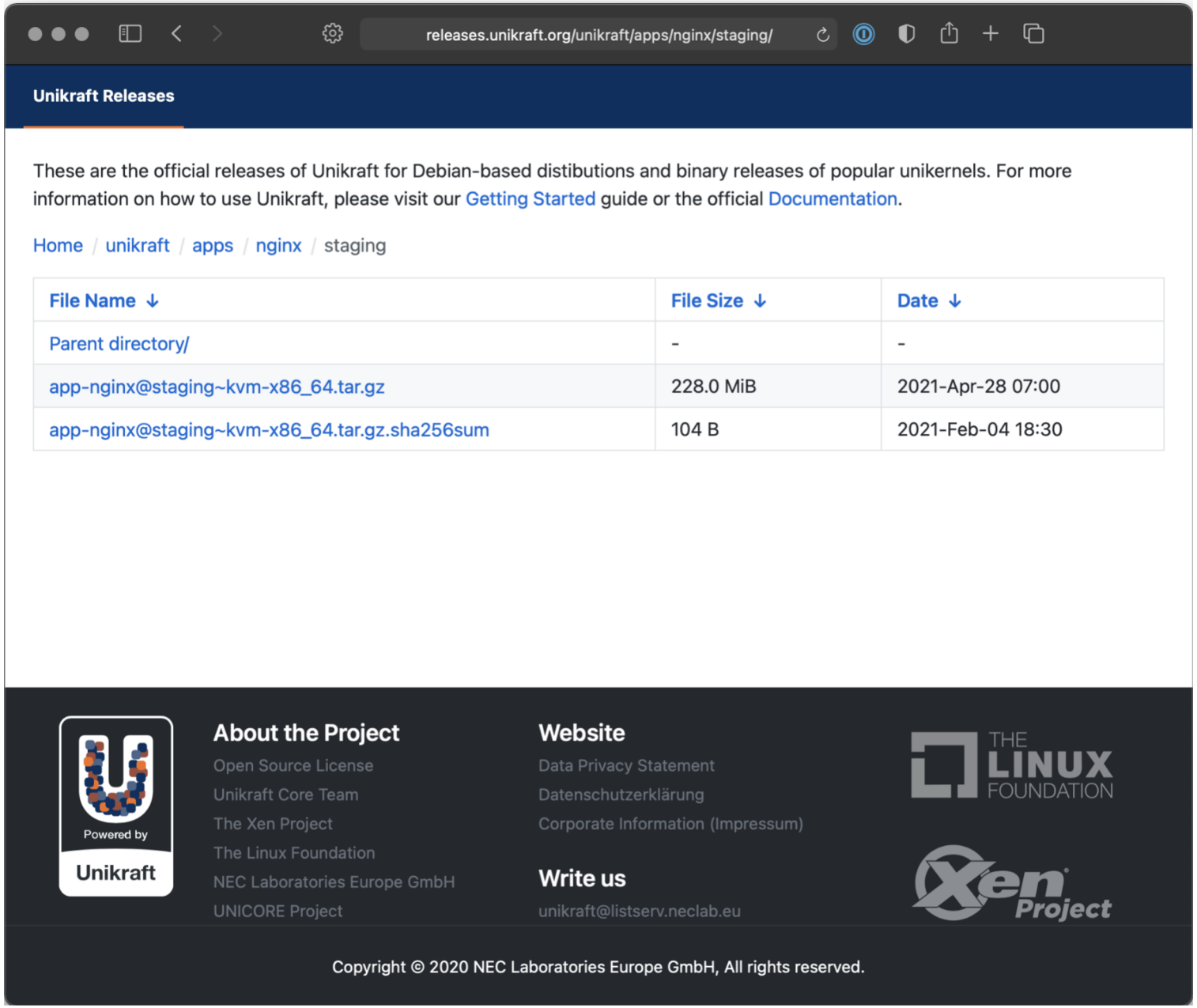
Define, configure, build and run Unikraft unikernels.

- Easily manage multiple libraries from different sources
- Quickly access updates and change between versions
- Automatically download application source dependencies

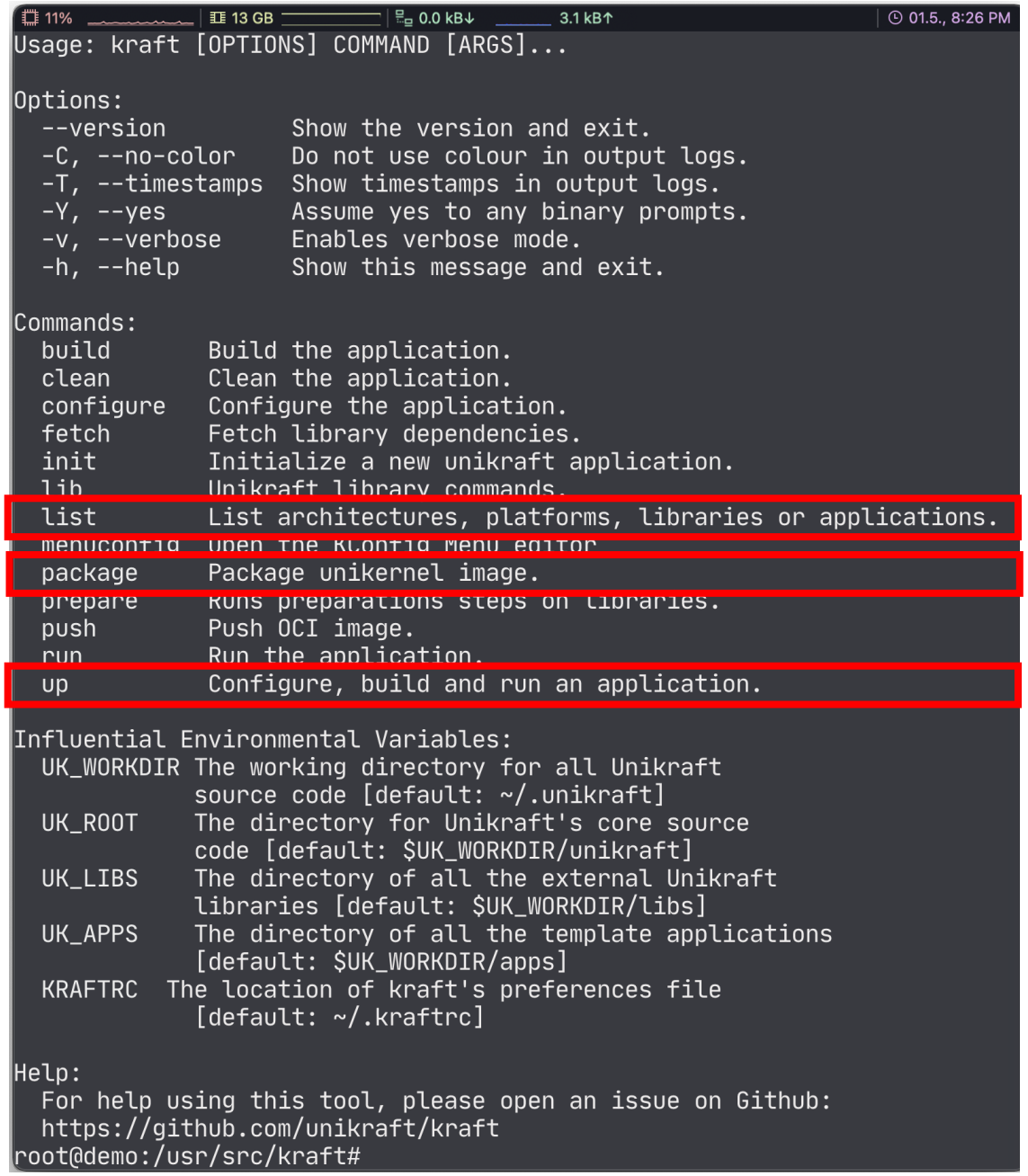


Retrieve stable and latest versions with pre-built source files

(no need to compile everything from scratch)



And Much More!



Towards Seamless Integration & Deployment

KUBERNETES

Deploy extremely efficient Unikraft images seamlessly on your Kubernetes cluster

ongoing

PROMETHEUS

Monitor your Unikraft instances through a standard monitoring platform

ongoing

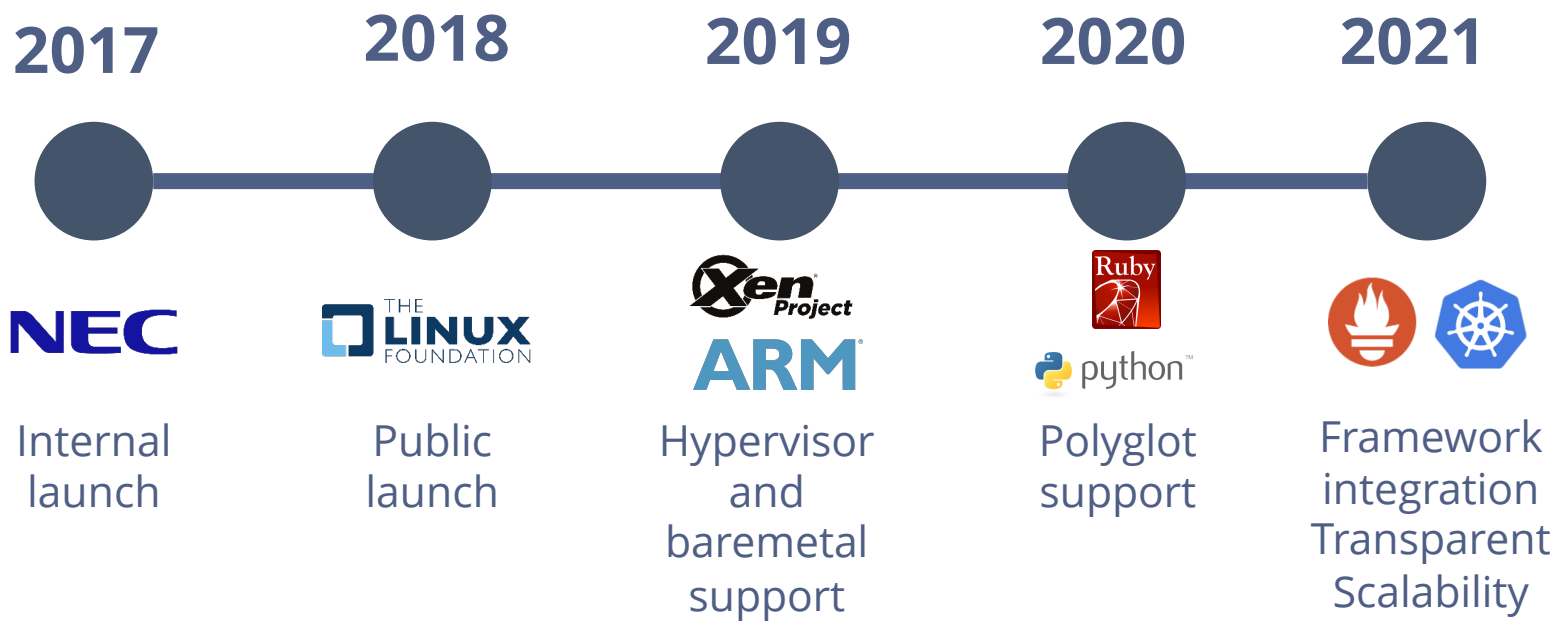
BOSH

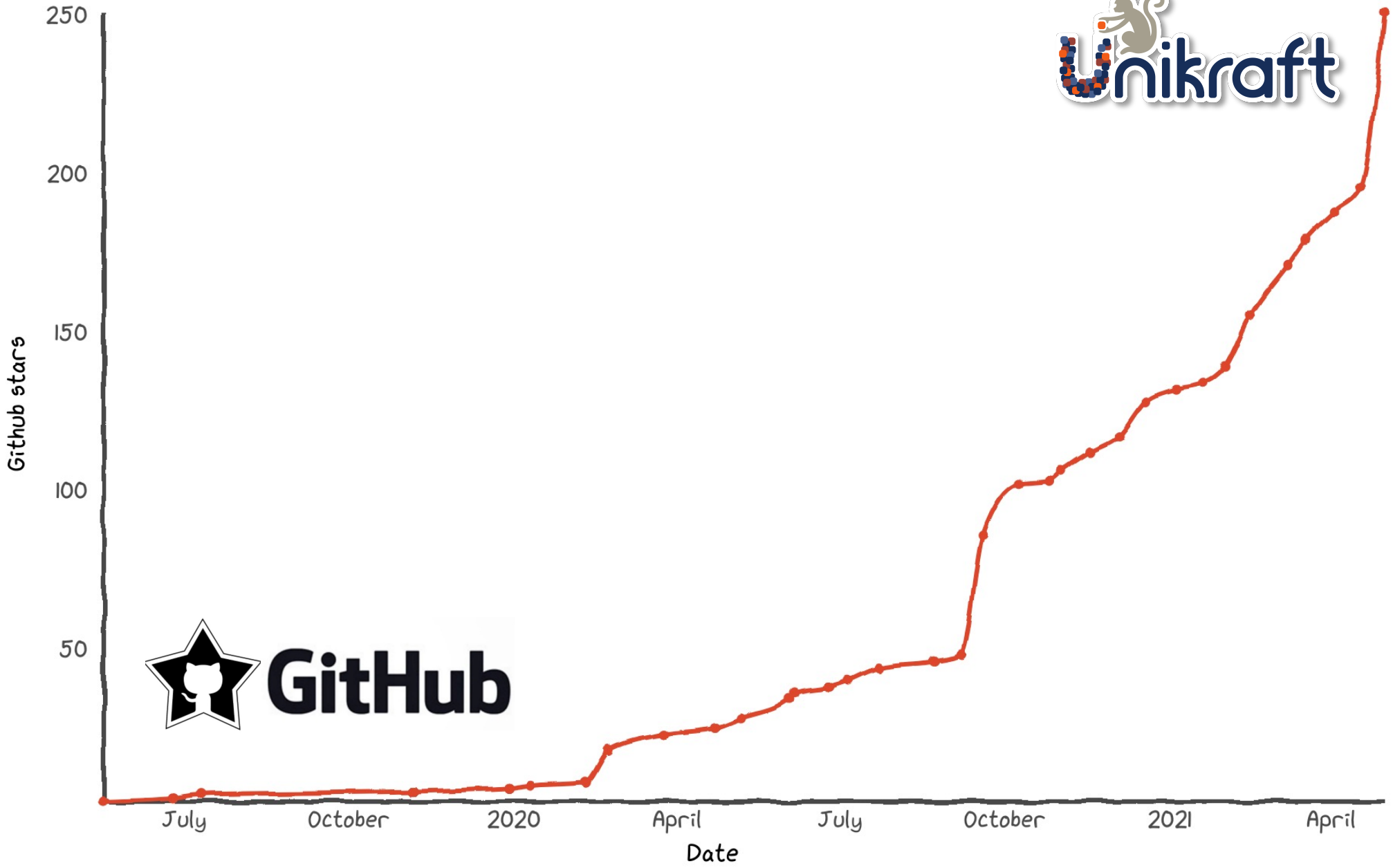
Perform your Unikraft builds directly via your Cloud Foundry infrastructure

planned

Open Source Project

Project History




Unikraft
GitHub
250
200
150
100
50
Github stars
July
October
2020
April
July
October
2021
April
Date

Unikraft: Fast, Specialized Unikernels the Easy Way

Simon Kuenzer
NEC Laboratories Europe GmbH

Vlad-Andrei Bădoiu*
University Politehnica of Bucharest

Hugo Lefeuvre*
The University of Manchester

Sharan Santhanam*
NEC Laboratories Europe GmbH

Alexander Jung*
Lancaster University

Gaulthier Gain*
University of Liège

Cyril Soldani*
University of Liège

Costin Lupu
University Politehnica of Bucharest

Ștefan Teodorescu
University Politehnica of Bucharest

Costi Răducanu
University Politehnica of Bucharest

Cristian Banu
University Politehnica of Bucharest

Laurent Mathy
University of Liège

Răzvan Deaconescu
University Politehnica of Bucharest

Costin Raiciu
University Politehnica of Bucharest

Felipe Huici
NEC Laboratories Europe GmbH

Abstract

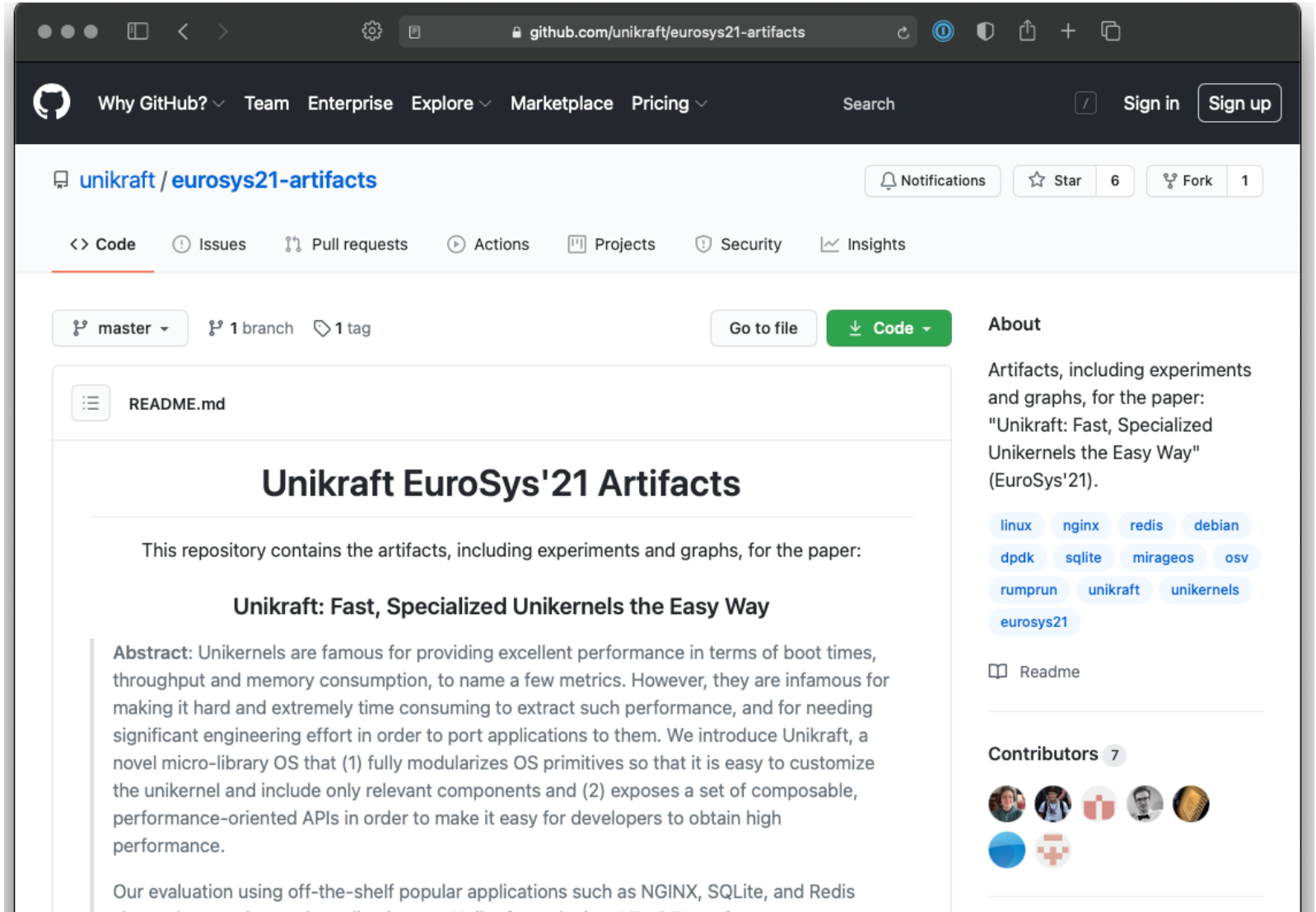
Unikernels are famous for providing excellent performance in terms of boot times, throughput and memory consumption, to name a few metrics. However, they are infamous for making it hard and extremely time consuming to extract such performance, and for needing significant engineering effort in order to port applications to them. We introduce Unikraft, a novel micro-library OS that (1) fully modularizes OS primitives so that it is easy to customize the unikernel

65], or providing efficient container environments [62, 76], to give some examples. Even in the hardware domain, and especially with the demise of Moore's law, manufacturers are increasingly leaning towards hardware specialization to achieve ever better performance; the machine learning field is a primary exponent of this [30, 32, 34].

In the virtualization domain, unikernels are the golden standard for specialization, showing impressive results in terms of throughput, memory consumption, and boot times,

All our experiments are open and reproducible:

github.com/unikraft/eurosys21-artifacts



High performance POSIX unikernels are now a reality!

Info: https://unikraft.org/
Code: https://github.com/unikraft
Reproduce: https://github.com/unikraft/eurosys21-artifacts

felipe.huici@neclab.eu