

Project

Unleashing the Power of Unikernels with inkraft

Florian Schmidt, Research Scientist, NEC Laboratories Europe



5G ESSENCE This work has received funding from the European Union's Horizon 2020 research and innovation program under grant agreements no. 675806 ("5G CITY") and 761592 ("5G ESSENCE"). This work reflects only the author's views and the European Commission is not responsible for any use that may be made of the information it contains.

VMs have been around for a long time

• They allow consolidation, isolation, migration, ...

VMs have been around for a long time

• They allow consolidation, isolation, migration, ...

Then containers came and many people LOVED them. Why?

© NEC Corporation 2018

VMs have been around for a long time

- They allow consolidation, isolation, migration, ...
- Then containers came and many people LOVED them. Why?
- "Containers are much faster to bring up than VMs. My VM takes a minute to boot, my container only a second."

VMs have been around for a long time

• They allow consolidation, isolation, migration, ...

Then containers came and many people LOVED them. Why?

- "Containers are much faster to bring up than VMs. My VM takes a minute to boot, my container only a second."
- "Containers are much smaller. My VM takes 10 GB, my container only a few hundred MB."

VMs have been around for a long time

• They allow consolidation, isolation, migration, ...

Then containers came and many people LOVED them. Why?

- "Containers are much faster to bring up than VMs. My VM takes a minute to boot, my container only a second."
- "Containers are much smaller. My VM takes 10 GB, my container only a few hundred MB."
- "Containers are much easier to create and deploy. I just write this Dockerfile and I'm done."

Containers vs Unikernels

- I don't want to bash containers.
 - Containers can be great!
 - For example, I love them for build environments

But VMs do not have to be large, slow, and complicated

Containers vs Unikernels

- I don't want to bash containers.
 - Containers can be great!
 - For example, I love them for build environments

But VMs do not have to be large, slow, and complicated

This is where unikernels come in

- A single binary containing OS and (single) application
- A shared address space, no mode switching
- No isolation within unikernel, but by the hypervisor
- Can be extremely small and blazingly fast







© NEC Corporation 2018



Server: Intel Xeon E5-1630 v3 CPU@3.7GHz (4 cores), 128GB DDR4 RAM, Xen/Linux versions 4.8

Two messages from those results



Unikernels can instantiate as fast as containers

Except when many are colocated

 Speaking of which, what is going wrong there?!

Server: Intel Xeon E5-1630 v3 CPU@3.7GHz (4 cores), 128GB DDR4 RAM, Xen/Linux versions 4.8



Intermission



Anybody remember last year?

Filipe Manco: "Death to the XenStore!" (Xen Developer and Design Summit 2017)



Anybody remember last year?

Filipe Manco: "Death to the XenStore!" (Xen Developer and Design Summit 2017)



There's an implementation now!

Implementation available

- NoXenstore: managing Xen VMs without a central management unit
- Plus assorted other bits and pieces of goodies
- Technology readiness level: research prototype (so don't expect to see this exact implementation hit upstream anytime soon...)

Materials:

- http://sysml.neclab.eu/projects/lightvm/
- https://github.com/sysml/lightvm
- Manco et al., "My VM is Lighter (and Safer) than your Container", SOSP 2017



On a dummy unikernel that does nothing, with 1 net device



On a dummy unikernel that does nothing, with 1 net device



On a dummy unikernel that does nothing, with 1 net device



On a dummy unikernel that does nothing, with 1 net device





OK, back to the main topic



The Downside

So, unikernels are great and everybody should use them now?



Hold your horses...

The big problem is unikernel *development:* Optimized unikernels are manually built



Hold your horses...

The big problem is unikernel *development:* Optimized unikernels are manually built

- Building takes several months or even longer
 - We've done it before, multiple times

Hold your horses...

The big problem is unikernel *development:* Optimized unikernels are manually built

- Building takes several months or even longer
 - We've done it before, multiple times
- Potentially lather, rinse, repeat for each target application

• We've done that too...

Hold your horses...

The big problem is unikernel *development:* Optimized unikernels are manually built

- Building takes several months or even longer
 - We've done it before, multiple times
- Potentially lather, rinse, repeat for each target application

• We've done that too...

That's not an effective way of doing things



Unicore







Motivation

- Core principle: Support a wide range of use cases
- Provide common code base for unikernel development
- Simplify building and optimizing



Motivation

- Core principle: Support a wide range of use cases
- Provide common code base for unikernel development
- Simplify building and optimizing





Motivation

- Core principle: Support a wide range of use cases
- Provide common code base for unikernel development

Simplify building and optimizing



Our Approach

- Decompose OS functionality into libraries
- Unikraft's two components:
 - Library Pool
 - Build Tool





Motivation

- Core principle: Support a wide range of use cases
- Provide common code base for unikernel development
- Simplify building and optimizing



Our Approach

- Decompose OS functionality into libraries
- Unikraft's two components:
 - Library Pool
 - Build Tool

Started as an internal project at NEC Labs in early 2017 Made public early on

- Discussed ideas at Xen Summit 2017
- •Accepted as an incubator project in October 2017
- •First public code release in December 2017

Decompose OS into a set of libraries ("Everything is a library")

Recompose them to meet the needs of particular applications
Recompose them to meet the needs of particular applications

Application(s)					
network stack					
profiling	filesystem				
memory	timers				
allocator	scheduler				
drivers					



Recompose them to meet the needs of particular applications

Application(s)						
network stack	Cite events and					
profiling	filesystem					
memory	timers					
allocator	scheduler					
drivers						

Recompose them to meet the needs of particular applications





Recompose them to meet the needs of particular applications





myapp

(1) Select / build / port Application











Two Library Types

Built-in: functionality specific to Unikraft, live in the main unikraft repo

- ukboot
- ukschedpreempt

•...

Two Library Types

Built-in: functionality specific to Unikraft, live in the main unikraft repo

- ukboot
- ukschedpreempt

External: software projects external to Unikraft, have their own unikraft-lib repos

- Iwip
- micropython
- ...

48

• ...



Example System

Micropython Unikernel for Xen on x86_64



app_my_python.o	libmicropython.o
liblwip.o	

app_my_python.o	libmicropython.o
liblwip.o	libvfscore.o

app_my_python.o	libmicropython.o
liblwip.o	libvfscore.o
libschedcoop.o	

app_my_python.o	libmicropython.o
liblwip.o	libvfscore.o
libschedcoop.o	liballocbbuddy.o

app_my_python.o	libmicropython.o
liblwip.o	libvfscore.o
libschedcoop.o	liballocbbuddy.o
libxenplat.o	

app_my_python.o	libmicropython.o
liblwip.o	libvfscore.o
libschedcoop.o	liballocbbuddy.o
libxenplat.o	libx86_64arch.o

app_my_python.o	libmicropython.o
liblwip.o	libvfscore.o
libschedcoop.o	liballocbbuddy.o
libxenplat.o	libx86_64arch.o



Kconfig/Makefile based

Kconfig/Makefile based make menuconfig

Kconfig/Makefile based make menuconfig

• Choose options in the menu that you want for your application



Kconfig/Makefile based make menuconfig

- Choose options in the menu that you want for your application
- Choose your target platform(s) (currently: Xen, KVM, Linux) and architectures

Kconfig/Makefile based

make menuconfig

- Choose options in the menu that you want for your application
- Choose your target platform(s) (currently: Xen, KVM, Linux) and architectures

Save config and make

Kconfig/Makefile based

make menuconfig

- Choose options in the menu that you want for your application
- Choose your target platform(s) (currently: Xen, KVM, Linux) and architectures Save config and make

1/1 - + [* □ * Tilix: Default Q = _ □ *	Open ▼ IFI .config Save	e = _ - ×
config - Unicore/0.1-alpha~cd730c8-custom Configuration	1#	
Unicore/0.1-alpha-cd730c8-custon Configuration Arrow keys navigate the menu. <enter> selects submenus> (or empty submenus>). Highlighted letters are hotkeys. Pressing <y> selectes a feature, while <n> will exclude a feature. Press <esc> to exit, <? > for Help, for Search. Legend: [*] feature is selected [] feature is Architecture (x86_64)> Architecture options> Platforn support> Libraries> Build options> (unicore) Image name Select> < Exit > < Help > < Save > < Load ></esc></n></y></enter>	<pre>2 # Automatically generated file; D0 NOT EDIT. 3 # Unicore/0.1-alpha~cd730c8-custom Configuration 4 # 5 ARCH_X86_64=y 6 # ARCH_X86_32 is not set 7 # ARCH_ARM_32 is not set 8 9 # 10 # Architecture options 11 # 12 MARCH_GENERIC=y 13 # MARCH_OOREIC=y 13 # MARCH_CORE2 is not set 15 # MARCH_CORE17 is not set 15 # MARCH_CORE17 is not set 16 # MARCH_COREI7 is not set 17 # MARCH_COREI7AVX is not set 18 # MARCH_K8 is not set 19 # MARCH_K8 is not set 20 # MARCH_K8SE3 is not set 21 # MARCH_ANDFAM10 is not set 22 # MARCH_BTVER1 is not set 23 # MARCH_BDVER1 is not set 24 # MARCH_BDVER2 is not set</pre>	
	Plain Text 🔻 Tab Width: 8 👻 🛛 Ln	1, Col 1 🔻 INS

Æ confin



A Baseline Example...

Xen PV x86_64 binary

A Baseline Example...

Xen PV x86_64 binary

unikraft_xen-x86_64.o



A Baseline Example...

Xen PV x86_64 binary

unikraft_xen-x86_64.o



Boots and prints messages to debug console (with min. 208kB RAM)

1 /	1 👻	+	Ct	무			Tilix: I	Default		۹	:		8	×
(d9) (d9)	Info): C D: [libx@ libx@	enplat] enplat]	setup.c @ setup.c @	174 189	: Ente	ering fro start in	m Xen (x86, fo: 0x15600	PV). 00				
(d9)	Info	o: Ē	libxe	enplat	setup.c @	190	: 5	shared_in	fo: 0x2000					
(d9)	Info	D: [libxe	enplat]	setup.c @	191	: hype	ercall_pa	ge: 0x3000					
(d9)	Info	D: [libxe	enplat]	setup.c @	154		start_p	fn: 15e					
(d9)	Info	D: [libxe	enplat]	setup.c @	155		max_p	fn: 20000					
(d9)	Info): C	libxe	enplat]	MM.C @ 160) : (Mapping	д мемогу	range 0x15e	≥000 -	0×200	00000		
(d9)	Кегі	n: W	elcor	me to				_						
(d9)	Кегі	n:		(_) /		_/ _/ /	/_						
(d9)	Кегі	n: /	11	/ _ \/`	/ '_///	′_`.	/ _/	_/						
(d9)	Кегі	n: \	_,_/_	_//_/_/	_/_\/_/ \	/	_/ \/	/						
(d9)	Кегі	n: Ö			Titar	10.2	~de72ed	de						
(d9)	Info): [libuk	kboot]	boot.c @ 72	2 :	Callir	ng main(2	, ['unikraf	⁼t', '	consol	e=hvc	0'])	
(d9)	Кегі	n: w	eak r	main()	called. Syr	1bol	was not	t ⁻ replace	d!					
(d9)	ERR	: [libu	kboot]	boot.c @ 19	95 :	weak r	main() ca	lled. Symbo	ol was	not r	eplace	ed!	
(d9)	Info	D: [libu	<pre>kboot]</pre>	boot.c @ 82	2 :	main r	returned	-22, haltir	ng sys	tem			
ð 🔒	evel:	1	root	> ~ >	workspace)	> uni	kraft)	> unikraf	t					

Building a Unikraft Hello World App



Clone the main Unikraft repo

git clone git://xenbits.xen.org/unikraft/unikraft.git

Clone any external library repos

git clone git://xenbits.xen.org/unikraft/libs/newlib.git

Create repo for the actual application

Clone the main Unikraft repo

git clone git://xenbits.xen.org/unikraft/unikraft.git

Clone any external library repos

git clone git://xenbits.xen.org/unikraft/libs/newlib.git

Create repo for the actual application



Clone the main Unikraft repo

git clone git://xenbits.xen.org/unikraft/unikraft.git

Clone any external library repos

git clone git://xenbits.xen.org/unikraft/libs/newlib.git

Create repo for the actual application



Clone the main Unikraft repo

git clone git://xenbits.xen.org/unikraft/unikraft.git

Clone any external library repos

git clone git://xenbits.xen.org/unikraft/libs/newlib.git

Create repo for the actual application


Repo Structure

Clone the main Unikraft repo

git clone git://xenbits.xen.org/unikraft/unikraft.git

Clone any external library repos

git clone git://xenbits.xen.org/unikraft/libs/newlib.git

Create repo for the actual application



Makefile: specify where the main Unikraft repo is, as well as repos for external libraries

```
UK_ROOT ?= $(PWD)/../../unikraft
UK_LIBS ?= $(PWD)/../../unikraft-libs
LIBS := $(UK_LIBS)/newlib
all:
    @make -C $(UK_ROOT) A=$(PWD) L=$(LIBS)
$(MAKECMDGOALS):
    @make -C $(UK ROOT) A=$(PWD) L=$(LIBS) $(MAKECMDGOALS)
```

Hello World – Four Required Files (I)

Makefile: specify where the main Unikraft repo is, as well as repos for external libraries

Hello World – Four Required Files (I)

Makefile: specify where the main Unikraft repo is, as well as repos for external libraries

Hello World – Four Required Files (I)

Makefile: specify where the main Unikraft repo is, as well as repos for external libraries

```
UK_ROOT ?= $(PWD)/../../unikraft → path to unikraft repo
UK_LIBS ?= $(PWD)/../../unikraft-libs → path to external libs
LIBS := $(UK_LIBS)/newlib → external libs needed
all:
@make -C $(UK_ROOT) A=$(PWD) L=$(LIBS)
$(MAKECMDGOALS):
@make -C $(UK_ROOT) A=$(PWD) L=$(LIBS) $(MAKECMDGOALS)
```

Makefile.uk: specifies the sources to build for the application

```
$(eval $(call addlib,apphelloworld))
```

APPHELLOWORLD_SRCS-y += \$(APPHELLOWORLD_BASE)/main.c

Makefile.uk: specifies the sources to build for the application

APPHELLOWORLD_SRCS-y += \$(APPHELLOWORLD_BASE)/main.c

Makefile.uk: specifies the sources to build for the application



Config.uk: to populate Unikraft's menu with application-specific option

```
### Invisible option for dependencies
config APPHELLOWORLD DEPENDENCIES
       bool
        default y
        select LIBNOLIBC if !HAVE LIBC
### App configuration
config APPHELLOWORLD PRINTARGS
        bool "Print arguments"
        default y
        help
          Prints argument list (argv) to stdout
```

Hello World – Four Required Files (IV)

main.c: source file to provide (at least) a main() function

```
#include <stdio.h>
/* Import user configuration: */
#include <uk/config.h>
int main(int argc, char *argv[])
{
      printf("Hello world!\n");
#if CONFIG APPHELLOWORLD PRINTARGS
       int i;
      printf("Arguments:s");
       for (i=0; i<argc; ++i)</pre>
              printf(" \"%s\"", argv[i]);
      printf("\n");
#endif
```

Hello World – Four Required Files (IV)

main.c: source file to provide (at least) a main() function

```
#include <stdio.h>
/* Import user configuration: */
#include <uk/config.h>
int main(int argc, char *argv[])
                                                 Unikernel entry point
                                                 after boot
{
      printf("Hello world!\n");
#if CONFIG APPHELLOWORLD PRINTARGS
       int i;
      printf("Arguments:s");
       for (i=0; i<argc; ++i)</pre>
              printf(" \"%s\"", argv[i]);
      printf("\n");
#endif
```

Hello World – Four Required Files (IV)

main.c: source file to provide (at least) a main() function

```
#include <stdio.h>
/* Import user configuration: */
#include <uk/config.h>
int main(int argc, char *argv[])
                                                Unikernel entry point
                                                after boot
{
      printf("Hello world!\n");
                                         ——— defined by Config.uk
#if CONFIG APPHELLOWORLD PRINTARGS 🔶
      int i;
      printf("Arguments:s");
      for (i=0; i<argc; ++i)</pre>
             printf(" \"%s\"", argv[i]);
      printf("\n");
#endif
```

Porting an External Library



Write Makefile.uk and add the external library source files to it

- The library's original Makefile can serve as a template
- I'll show how to in a second

Write Makefile.uk and add the external library source files to it

- The library's original Makefile can serve as a template
- I'll show how to in a second

Write Config.uk with library-specific options

Write Makefile.uk and add the external library source files to it

- The library's original Makefile can serve as a template
- I'll show how to in a second

Write Config.uk with library-specific options

- Isn't strictly required
- But at least a simple "library on/off" option for kbuild is a good idea



Write Makefile.uk and add the external library source files to it

- The library's original Makefile can serve as a template
- I'll show how to in a second

Write Config.uk with library-specific options

- Isn't strictly required
- But at least a simple "library on/off" option for kbuild is a good idea

Sometimes a bit of glue code is needed

• E.g., in newlib to link POSIX thread creation to Unikraft's thread library

Write Makefile.uk and add the external library source files to it

- The library's original Makefile can serve as a template
- I'll show how to in a second

Write Config.uk with library-specific options

- Isn't strictly required
- But at least a simple "library on/off" option for kbuild is a good idea

Sometimes a bit of glue code is needed

- E.g., in newlib to link POSIX thread creation to Unikraft's thread library
- In this early phase: implement core unikraft functionality that is required to support your library, for example:
 - File descriptors/sockets
 - Threading support

Library registration

\$(eval \$(call addlib_s,libaxtls,\$(LIBAXTLS)))



Library registration

\$(eval \$(call addlib_s,libaxtls,\$(LIBAXTLS)))

Library registration

\$(eval \$(call addlib s,libaxtls,\$(LIBAXTLS)))

Source Download

Nothing here: sources are small and included directly in the uk library repo

— Register library with unikraft build system

Library registration

\$(eval \$(call addlib_s,libaxtls,\$(LIBAXTLS)))

Source Download

\$(eval \$(call fetch,libnewlibc,\$(LIBNEWLIB_URL)))

\$(eval \$(call patch,libnewlibc,\$(LIBNEWLIB_PATCHDIR),newlib-\$(LIBNEWLIB_VERSION)))

— Register library with unikraft build system

— Download and patch library code



Library registration

\$(eval \$(call addlib_s,libaxtls,\$(LIBAXTLS)))

Source Download

\$(eval \$(call fetch,libnewlibc,\$(LIBNEWLIB_URL)))

\$(eval \$(call patch,libnewlibc,\$(LIBNEWLIB_PATCHDIR),newlib-\$(LIBNEWLIB_VERSION)))

Library includes

CINCLUDES-y += -I\$(LIBAXTLS_BASE)/include \

-I\$(LIBAXTLS_BASE)/crypto \

-I\$(LIBAXTLS_BASE)/ssl

Download and patch library code

The library's original include directories

Library registration \$(eval \$(call addlib s,libaxtls,\$(LIBAXTLS))) # Source Download \$(eval \$(call fetch,libnewlibc,\$(LIBNEWLIB URL))) \$(eval \$(call patch,libnewlibc,\$(LIBNEWLIB PATCHDIR),newlib-\$(LIBNEWLIB VERSION))) ****************************** # Library includes CINCLUDES-y += -I\$(LIBAXTLS BASE)/include \ -I\$(LIBAXTLS BASE)/crypto \ -I\$(LIBAXTLS BASE)/ssl # sources LIBAXTLS SRCS-y += \$(LIBAXTLS BASE)/crypto/aes.c LIBAXTLS SRCS-y += \$(LIBAXTLS BASE)/crypto/bigint.c LIBAXTLS SRCS-y += \$(LIBAXTLS BASE)/crypto/sha512.c © NEC Corporation 2018

— Register library with unikraft build system

— Download and patch library code

— The library's original include directories



Unikraft 0.2 Titan

Current Status



Available Libraries

Core Libraries

- libfdt
 - Flat device tree parser
- libnolibc
 - A tiny libc replacement
- libukalloc
 - Memory allocator abstraction
- libukallocbbuddy
 - Binary buddy allocator
- libukargparse
 - Argument parser library
- libukboot
 - Unikraft bootstrapping
- libukdebug
 - Debug and kernel printing
 - Assertions, hexdump
- libuksched
 - Scheduler abstraction
- libukschedcoop
 - Cooperative scheduler

libukbus

- abstraction for device buses, e.g., PCI
- libuklock
 - mutexes and semaphores
- libukmpi
 - message-passing interface
- libukswrand
 - pseudo-RNG interface
- libuktimeconv
 - time calculation/conversion
- libvfscore
 - basic file descriptor management / mapping / handling

External Libraries

- Iibnewlib
 - libc originally aimed at embedded devices
- Iiblwip
 - lightweight TCP/IP stack

Architecture Libraries

- libarmmath
 - 64bit arithmetic on ARMv7

Platform Libraries

- libxenplat
- Xen (PV)
- x86_64, ARMv7
- libkvmplat
 - QEMU/kvm
 - x86_64
- liblinuxu
 - Linux userspace
 - x86_64, ARMv7

Coming soon: in the pipeline

Core Libraries

- libukschedpreempt
 - Pre-emptive scheduler

External Libraries

- Ibclick
 - Click modular router (e.g., for NFV)
- Iibaxtls
 - TLS support aimed at embedded devices
- Ibmicropython
 - Python implemented for microcontrollers

Platform Libraries

- libxenplat
 - ARM64 support(?)
 - netfront support
- libkvmplat
 - ARM64 support
 - virtio-net support
- liblinuxu
 - tap device based networking support



The road ahead

Project got accepted as incubator project in December Initially, mostly internal contributors

- Currently contributions in the pipeline from contributors from
 - Romania (netfront, scheduling; from University Politehnica Bucharest)
 - Israel (bare-metal support)
 - China (ARM64 support; from ARM)

We welcome additional contributors!

Resources:

- Code: xenbits.org/unikraft/*
- On-line documentation: unikraft.org
- IRC: #unikraft @ freenode
- Mailing list: shared with minios: minios-devel@lists.xen.org

Project DEVELOPER AND DESIGN SUMMIT

TTTTT