Caps$^2$: L4Re and CHERI

Adam Lackorzynski, Hermann Härtig

Apr 23$^{rd}$ 2016
Very Brief L4Re Intro

• L4 microkernel system
• OS framework – Security, RT, Virtualization
• Developed since 1996
• Small TCB, component based, isolating
• Capability-based object system: Kernel & User
• SMP – ARM, x86, MIPS – VT, SVM, HYP, TZ, VZ
• Paravirtualization (e.g. L^4Linux)
• Libc, C & C++, pthreads, POSIX, client/server framework, platform management
• Many (ARM) platforms
• Commercially supported & open source
L4Re Architecture Overview

- **L4Re Runtime Environment**
  - Device management
  - Crypto Service
  - Secure GUI

- **L4Re Microkernel & Hypervisor**
  - Direct Devices
  - Shared Devices
  - Multi-Core CPU, Memory

**Isolated Domains**
- Realtime App
- Secure App

**User**

**Kernel**

**Hardware**
• Everything is an object
  • All system calls run on an object
  • „one system call“ – send message to object
• Reference to object: capability
• Uniform interfaces
  • For kernel and user-level objects
• Pass access rights to physical memory pages to other domains
  • Flex-page: Size-aligned region description
• Revoke access rights
• Recursively: Stored in kernel
L4Re Object Capabilities

- References to kernel objects
  - Local naming (per protection domain)
  - Access control
- Kept in per task capability table, managed by user-level
- Pointer sized
- Can be passed to other protection domains
- User-level objects:
  - Via Kernel-provided communication object
- IPC sends to == invokes a capability
Communication

Sender Thread

UTCB
MRs
BRs
TCRs

capability

IPC GATE
LABEL

Receiver Thread

UTCB
MRs
BRs
TCRs

LABEL
TAG
p i w f

LABEL
TAG
p i w f
Resolving Faults

- Page faults & other exceptions
- Each thread has handler caps
  - Upon faults, the kernel invokes those caps
  - In good case: Handler replies with a resolving IPC containing a memory mapping / new CPU state
- Memory mappings can also be made explicit
- Memory access granularity: 4k (typically)
• Pass Cheri-Cap(s) via L4 messages to transfer fine-granular access rights
• Cheri-Cap faults as L4 exceptions
Byte-level Sharing

Server / Sender

Page

Partially Copied Page

With page-size granularity

Client / Receiver

Page

Server / Sender

Page

Client / Receiver

Page

With CHERI
System-wide CHERI Caps

Fine-grained memory access:
• No more copy-in/copy-out from data structures
• Sharing densely packed data
• Fine-grained access to MMIO

Applications:
• Device drivers (but could (should!)) be fixed by HW
• Database: KVS, ...
• ...