seL4 Microkernel Status Update

Gernot Heiser | gernot.heiser@data61.csiro.au | @GernotHeiser
• FOSDEM, Bruxelles, 2020-02-02

https://trustworthy.systems
What is seL4?
seL4: Assurance and Performance

The world’s **first** operating-system kernel with **provable** security enforcement

**Open Source**

The world’s **only** protected-mode OS with complete, sound timeliness analysis

World’s most advanced mixed-criticality OS

The world’s **fastest** general-purpose microkernel, designed for **real-world** use
World’s Most Secure OS: Arm v7

Confidentiality

Integrity

Availability

Abstract Model

Functional correctness:
C code only behaves as specified

Model enforces security

C Implementation

Translation validation:
Binary retains C-code semantics

Sound worst-case execution time bound

Binary code

Limitations (work in progress):
• Kernel initialisation not yet verified
• MMU & caches modelled abstractly
• Timing channels not ruled out
Military-Strength Security

Unmanned Little Bird (ULB)

DARPA HACMS: Retrofit existing system!

Secure Comms Dongle

Autonomous trucks

Cross-Domain Desktop Compositor
seL4 on RISC-V
**Background: HENSOLD Cyber**

- Munich-based startup
  - Secure RISC-V processor
  - Based on open-source Ariane core (ETH)
  - Supply chain secured through logic encryption
  - Secure OS based on seL4
  - Targets defence, industrial control, critint, automotive

Disclosure: I have an interest in HENSOLDT Cyber
## Performance on RV64

Message-passing round-trip latency in cycles

<table>
<thead>
<tr>
<th>Arch</th>
<th>x86 32b</th>
<th>x86 64b</th>
<th>Arm 32b</th>
<th>Arm 64b</th>
<th>RISC-V 64b</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intra address space</td>
<td>427</td>
<td>565</td>
<td>625</td>
<td>752</td>
<td>690</td>
</tr>
<tr>
<td>Inter address space</td>
<td>752</td>
<td>1041</td>
<td>625</td>
<td>752</td>
<td>1006</td>
</tr>
</tbody>
</table>

- Meltdown-workaround disabled (else much slower!)
- Not yet fully optimised!
- No ASIDS on **HiFive Unleashed**, else inter-AS would be same as intra-AS
- Hypervisor extensions supported in branch, tracking draft spec
Verification: RISC-V Status

Confidentiality

Integrity

Availability

Abstract Model

Functional correctness: RISC-V due Q1’20

Translation validation: RISC-V due Q2’20

C Implementation

Sound WCET bound RISC-V in progress

Binary code

Translation validation: RISC-V due Q2’20

Sound WCET bound RISC-V in progress
Experience with RISC-V Architecture

- Kernel port straightforward:
  - simple and clean RISC architecture
- Verification benefitted from cleanness
  - … but some challenges from less typing in page tables
- Hypervisor (draft) extensions even simpler
  - M (machine) mode makes firmware explicit
    - configures HW, delegates to S (supervisor) mode
    - emulates features not implemented in HW
    - should be verified
  - Extensibility of ISA could be a concern
    - could undermine portability
  - Formal ISA spec is great!
Mixed-Criticality Scheduling
(FOSDEM’19 Refresher)
**Mixed Criticality: Critical + Untrusted**

**NW driver must preempt control loop**
- … to avoid packet loss
- Driver must run at high prio
- Driver must be trusted not to monopolise CPU

- **Critical:**
  - Control loop
  - Sensor readings
  - Runs every 100 ms for few milliseconds

- **Untrusted:**
  - NW driver
  - NW interrupts
  - Runs frequently but for short time (order of µs)
MCS Challenge: Sharing

Vehicle control must see consistent state

Critical

Vehicle Control → Shared Data

Updates

Shared Data ← Navigation

Less critical
Sharing Through **Resource Server**

- **Control** $P_1$
- **Navig.** $P_2$
- **Communication endpoint (port)**
- **Server** $P_S$

- Single-threaded, guarantees atomicity
- Who pays for server time?
- Implements *immediate priority ceiling protocol* (IPCP) if $P_S = \max(P_1, P_2)$
Solution: Time Capabilities

Classical thread attributes

- Priority
- Time slice

New thread attributes

- Priority
- Scheduling context capability

Scheduling context object

- T: period
- C: budget (≤ T)

Limits CPU access!

Not runnable if null

Capability for time

Enables reasoning about time and temporal isolation for mixed-criticality systems

C = 2
T = 3

C = 250
T = 1000
Time Caps (MCS) Kernel Verification

- **Mainline Arm v7**
  - Spec
  - Proof
  - C
  - Binary

- **MCS Arm v7**
  - Spec
  - Proof
  - C
  - Binary

- **MCS RISC-V**
  - Spec
  - Proof
  - C
  - Binary

- **Mainline RISC-V**
  - Spec
  - Proof
  - C
  - Binary

- **New Mainline**

- **Q1'20**
  - Q4'20
  - Merge
  - Q4'20
  - Merge
  - Q2'20
Community/Ecosystem
<table>
<thead>
<tr>
<th>Security Standing Committee</th>
<th>Privileged Spec Tech Committee</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Invited me on</td>
<td>• Hypervisor-extension feedback well received</td>
</tr>
<tr>
<td>• Very receptive and supportive</td>
<td>- Easy engagement</td>
</tr>
<tr>
<td>• Committed to making RISC-V “most secure architecture”</td>
<td>- Constructive proposal from TC chair addressing our issues</td>
</tr>
<tr>
<td>• Facilitated engagement with Privspec TC (now Standing Committee)</td>
<td>• Time-protection slow to get traction</td>
</tr>
<tr>
<td></td>
<td>- Now good engagement, hopefully progress soon</td>
</tr>
</tbody>
</table>

• Open but skeptical
• They need to manage conflicting ideas
• Keen to get “most secure arch” recognition
We Are Creating the **seL4 Foundation**!

**Aims:**

- Provide a neutral entity for coordinating & enhancing seL4 ecosystem
- Grow adoption of seL4
- Improve (organisational and individual) community participation & cooperation
  - Developers
  - Adopters
- Develop / standardise seL4 system
  - kernel & proofs
  - libraries, services, tools
- Protect and promote the seL4 brand
  - prevent reputational damage from using modified seL4 (verification invalidated)
- Provide platform for pooling funds for critical “big-ticket” items (verification)
Foundation Structure

seL4 Foundation

seL4 Board

seL4 Fund Charter

seL4 Directed Fund $$

LF Projects LLC

seL4 Series LLC

seL4 TM

https://sel4.systems

seL4 Technical Charter

Technical Project

Contributor
Membership and Governance

- **Trustworthy Systems**: 3 directors
- **Premium Members**: US$ 100k/a - 1 director each
- **Members**: US$ 3–30k/a - 1 director
- **Associate Members**: US$ 0

Initial Board:
- June Andronick, TS
- Gernot Heiser, TS
- Gerwin Klein, TS
- John Launchbury, Galois (ex DARPA)
- Sascha Kegreiß, HENSOLDT Cyber
- Daniel Potts, Ghost Locomotion

Note: members must be financial members of Linux Foundation!
Foundation Status

• Legal docs (fund charter & technical charter) approved by Linux Foundation
• Trademark ready for transfer to Foundation
• Initial board appointed
• Interim web site shows structure, “Principles” and legal docs
• Hopefully days away from being able to set up members
  - Mail foundation@sel4.systems if you’re interested in joining!
  - Will make announcement on seL4.systems mailing lists

https://sel4.systems/Foundation