UNSW SYDNEY | Australia's Global University

School of Computer Science & Engineering

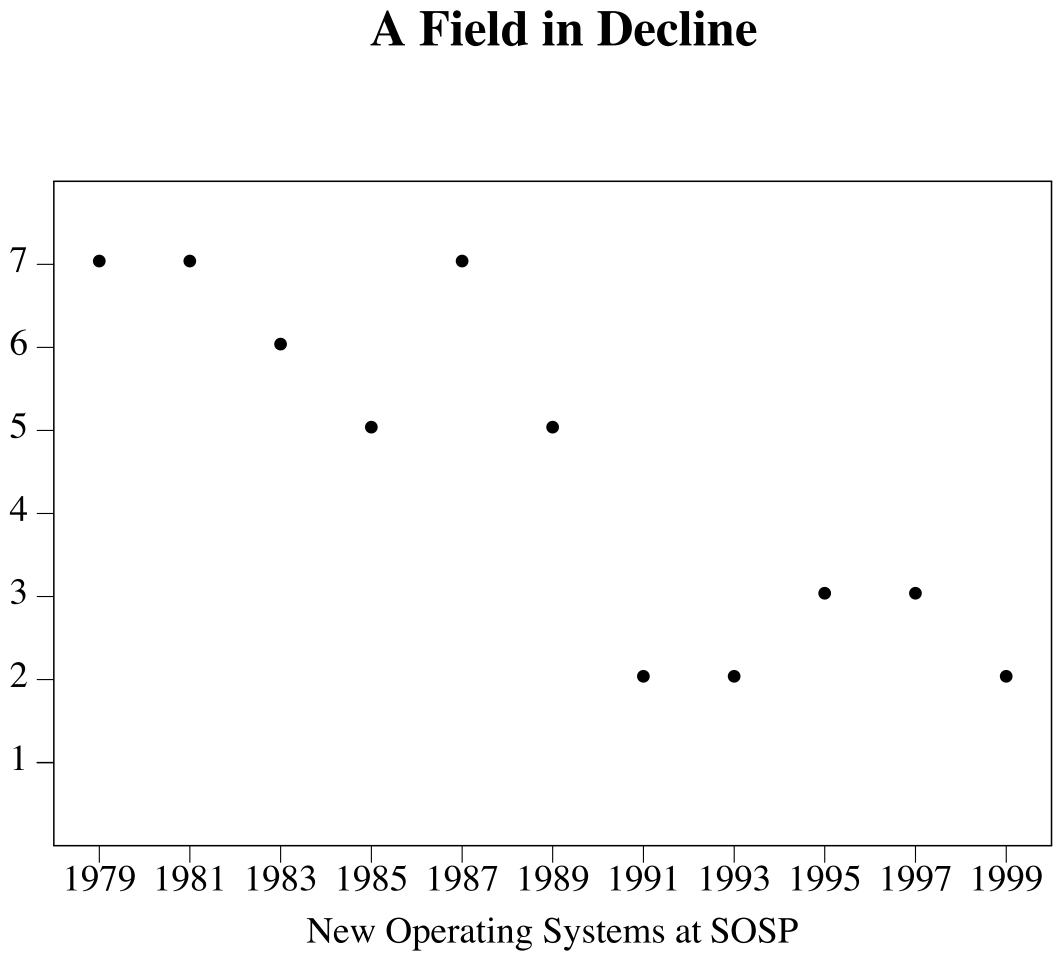
Trustworthy Systems Group

Don't Forget the OS – and the Principles!

Gernot Heiser

gernot@unsw.edu.au
@microkerneldude.bsky.social
https://gernot-heiser.org/

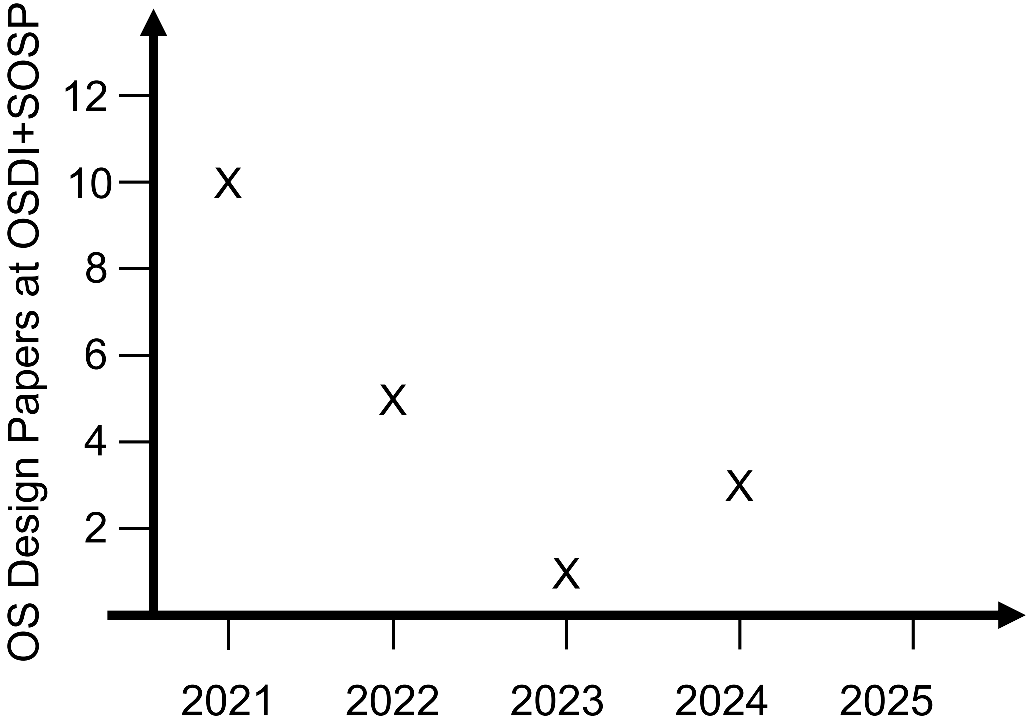
Rob Pike, 2000



Rob Pike: Systems Software Research is Irrelevant, 2000

A Quarter Century Later

Gernot's totally subjective assessment



Reviewing at Top-Tier Conferences

Reviewer writes

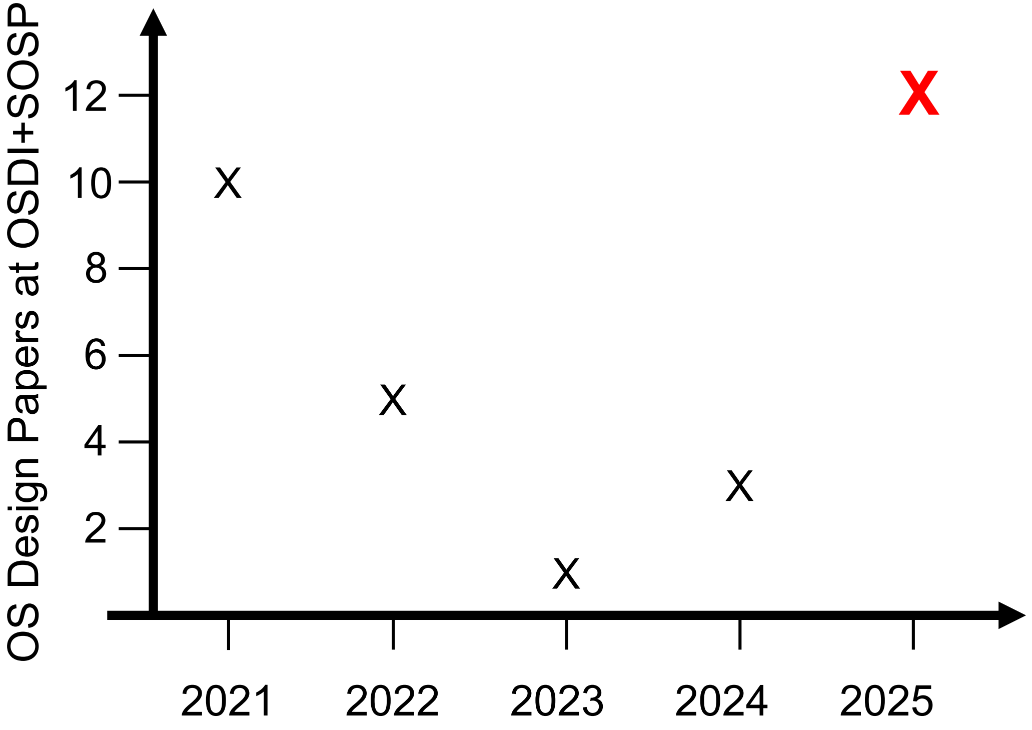
- “[OS] is built atop seL4 and seems to be formally verified (though I did not find how it was done).”
- “not compared with any other embedded OS kernels.”
- “approach well-known and implemented before, incl Mungi and L4Linux.”

Facts

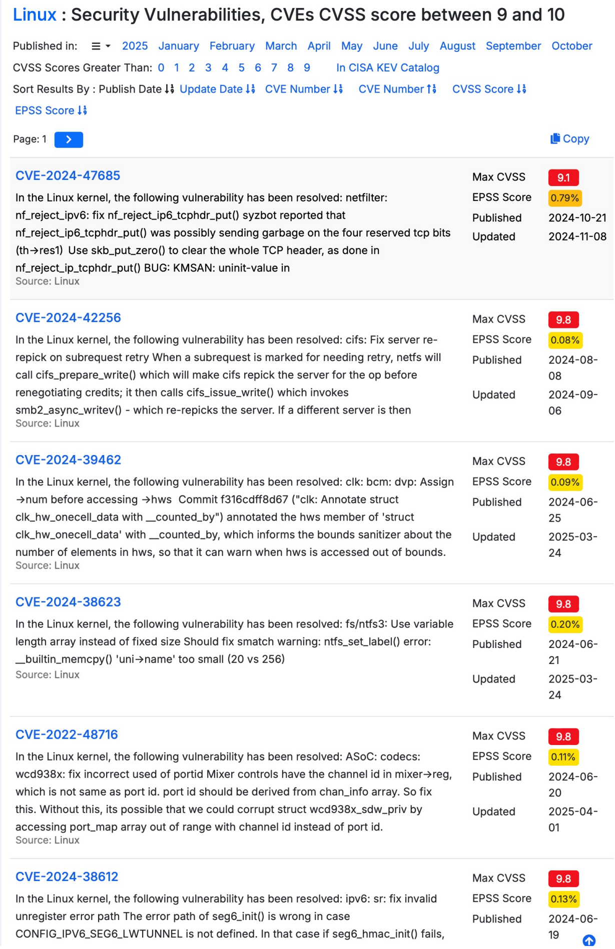
- “for now we leave verification out of scope, but aim for a verification-friendly design” [p 1]
- “We also compare to a commercial microkernel-based operating system, code-named CEOS.[2]” + Fig 4(d).
- Paper is about a highly modular OS, Mungi and L4Linux are the exact opposites: monolithic servers

There's Hope!

Gernot's totally subjective assessment



OSes Are Largely Still Broken!



Linux : Security Vulnerabilities, CVEs CVSS score between 9 and 10

Published in: ☰ 2025 January February March April May June July August September October

CVSS Scores Greater Than: 0 1 2 3 4 5 6 7 8 9 In CISA KEV Catalog

Sort Results By : Publish Date ↓ Update Date ↓ CVE Number ↓ CVE Number ↑ CVSS Score ↓ EPSS Score ↓

Page: 1 Copy

CVE-2024-47685

In the Linux kernel, the following vulnerability has been resolved: netfilter: nf_reject_ipv6: fix nf_reject_ip6_tcphdr_put() syzbot reported that nf_reject_ip6_tcphdr_put() was possibly sending garbage on the four reserved tcp bits (th→res1) Use skb_put_zero() to clear the whole TCP header, as done in nf_reject_ip_tcphdr_put() BUG: KMSAN: uninit-value in

Source: Linux

Max CVSS 9.1
EPSS Score 0.79%
Published 2024-10-21
Updated 2024-11-08

CVE-2024-42256

In the Linux kernel, the following vulnerability has been resolved: cifs: Fix server re-repick on subrequest retry When a subrequest is marked for needing retry, netfs will call cifs_prepare_write() which will make cifs repick the server for the op before renegotiating credits; it then calls cifs_issue_write() which invokes smb2_async_writev() - which re-repicks the server. If a different server is then

Source: Linux

Max CVSS 9.8
EPSS Score 0.08%
Published 2024-08-08
Updated 2024-09-06

CVE-2024-39462

In the Linux kernel, the following vulnerability has been resolved: clk: bcm: dvp: Assign →num before accessing →hws Commit f316cdff8d67 ("clk: Annotate struct clk_hw_onecell_data with __counted_by") annotated the hws member of 'struct clk_hw_onecell_data' with __counted_by, which informs the bounds sanitizer about the number of elements in hws, so that it can warn when hws is accessed out of bounds.

Source: Linux

Max CVSS 9.8
EPSS Score 0.09%
Published 2024-06-25
Updated 2025-03-24

CVE-2024-38623

In the Linux kernel, the following vulnerability has been resolved: fs/ntfs3: Use variable length array instead of fixed size Should fix smatch warning: ntfs_set_label() error: __builtin_memcpy() 'uni→name' too small (20 vs 256)

Source: Linux

Max CVSS 9.8
EPSS Score 0.20%
Published 2024-06-21
Updated 2025-03-24

CVE-2022-48716

In the Linux kernel, the following vulnerability has been resolved: ASoC: codecs: wcd938x: fix incorrect used of portid Mixer controls have the channel id in mixer→reg, which is not same as port id. port id should be derived from chan_info array. So fix this. Without this, its possible that we could corrupt struct wcd938x_sdw_priv by accessing port_map array out of range with channel id instead of port id.

Source: Linux

Max CVSS 9.8
EPSS Score 0.11%
Published 2024-06-20
Updated 2025-04-01

CVE-2024-38612

In the Linux kernel, the following vulnerability has been resolved: ipv6: sr: fix invalid unregister error path The error path of seg6_init() is wrong in case CONFIG_IPV6_SEG6_LWTUNNEL is not defined. In that case if seg6_hmac_init() fails,

Max CVSS 9.8
EPSS Score 0.13%
Published 2024-06-19

2024 CrowdStrike-related IT outages

Multiple blue screens of death caused by a faulty software update on baggage carousels at LaGuardia Airport, New York City

Date 19 July 2024; 8 months ago

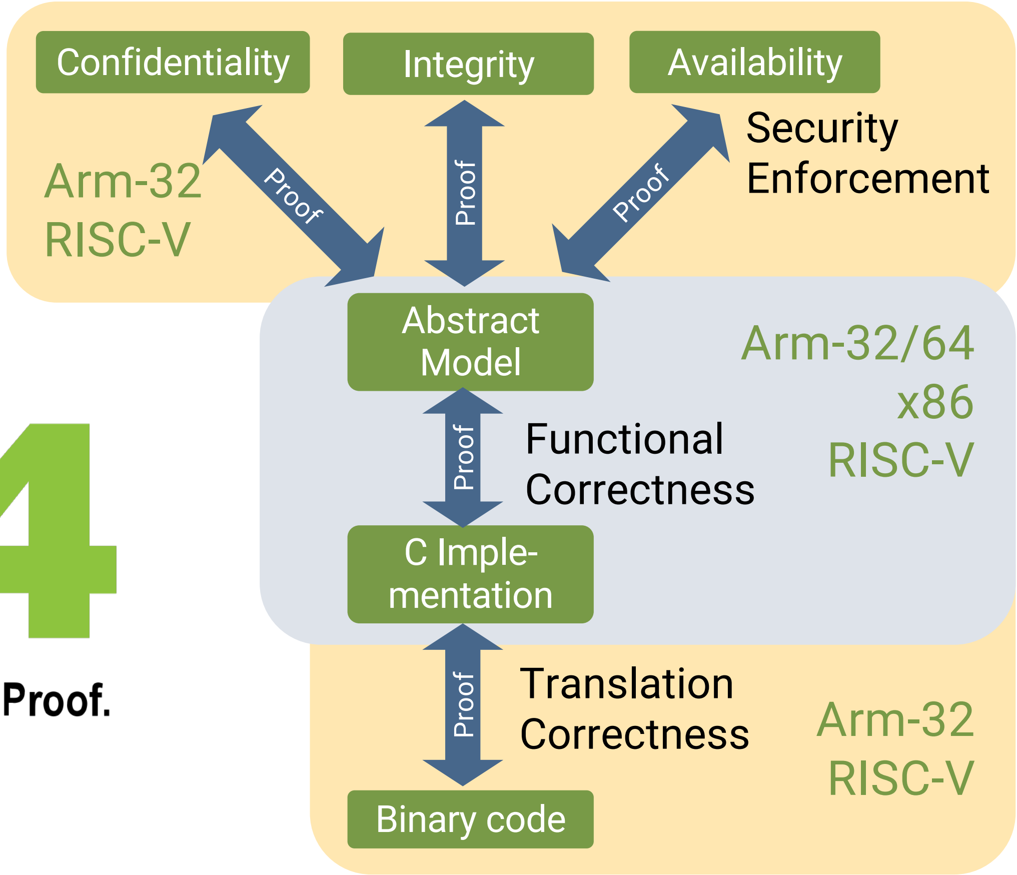
Cyber Attacks That Target Electrical Devices and Equipment: What Engineers Should Know

February 10, 2020 by Ikimi .O

We Have seL4!

Why didn't it solve the problem?

Security. Performance. Proof.



Confidentiality

Integrity

Availability

Arm-32
RISC-V

Proof

Proof

Proof

Security Enforcement

Abstract Model

Arm-32/64
x86
RISC-V

Proof

Functional Correctness

C Implementation

Proof

Translation Correctness

Arm-32
RISC-V

Binary code

The Assembly Language of OS

seL4 is a pure microkernel:

- Small: 10 kLOC
- Only fundamental, policy-free mechanisms
- No application-oriented services/abstractions
- **BYO file system, memory manager, device drivers**

Good design on seL4 requires deep (and rare) expertise

LionsOS

Need an seL4-based OS that is:

- well-designed
- easy to use
- verified

Principled OS Design

Radical simplicity:

- Fine-grained modularity, strict separation of concerns
- Event-driven programming model strictly sequential modules
- Static architecture
- Use-case-specific policies

Helps development **and** correctness!

Concurrency by distributing modules across cores

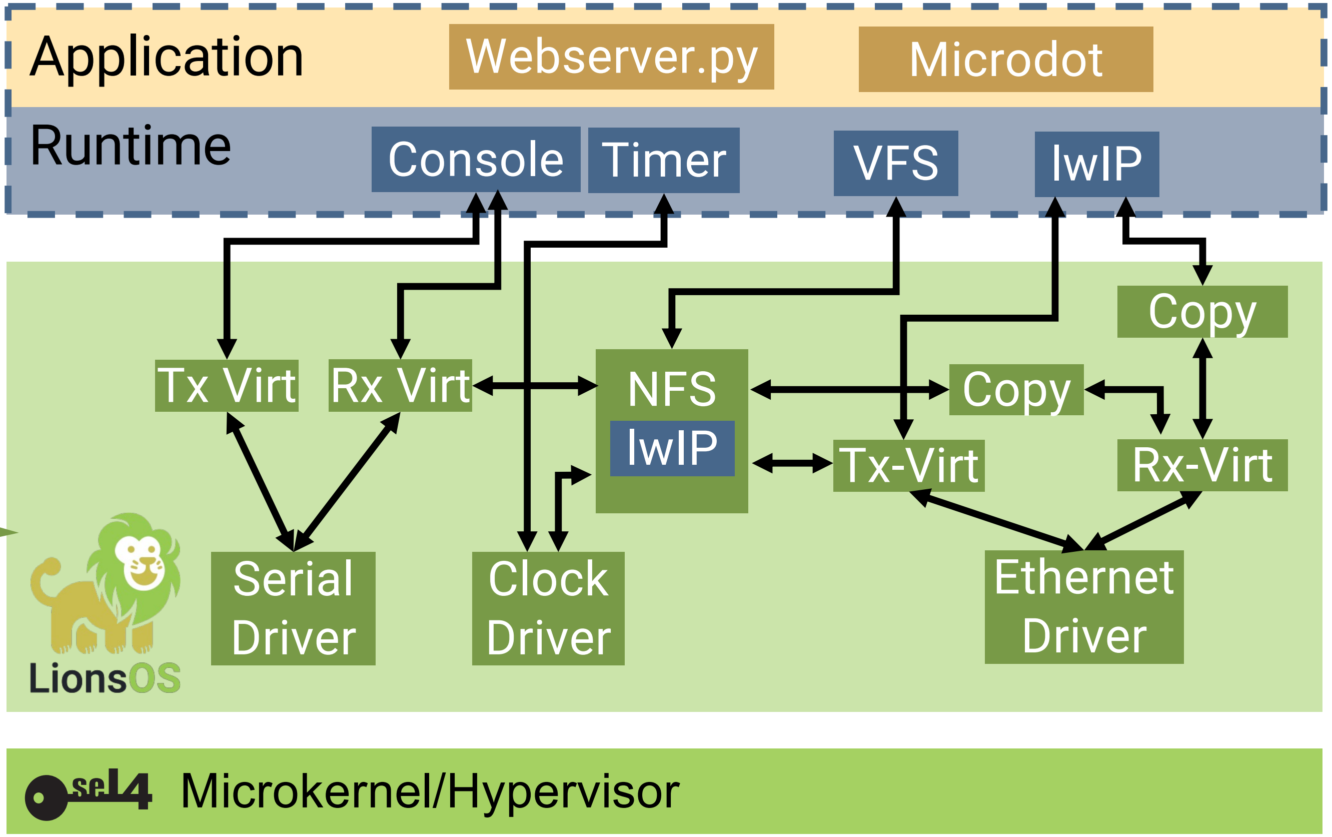
Matches embedded space – little dynamic resource management

Use-case diversity by replacing components

Underneath https://sel4.systems/

Web-server OS:

- 10 modules
- 3 libraries
- 3.5k SLOC trusted code

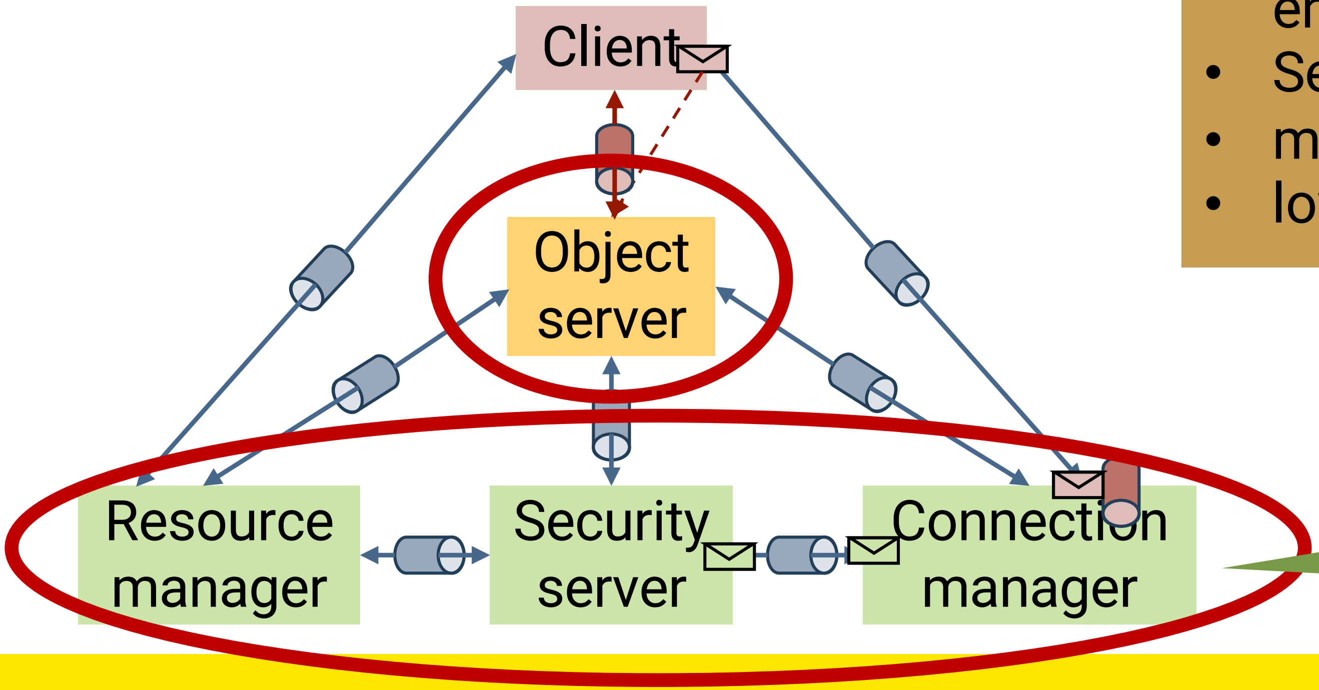


Take-Aways: Principled Design Works!

- Multiple deployed systems
- Ease of use:
  - Takes few hours to get started
  - 2nd-year students write performant device drivers
- Performance is great – beats Linux hands-down
- End-to-end verification in progress

Can We Go Further?

Aim: General-purpose OS that **provably** enforces a general security policy



Trusted core servers

Requires:

- mandatory security-policy enforcement
- Security-policy diversity
- minimal TCB
- low-overhead enforcement

Summary

- There are plenty of unsolved OS problems left
- Addressing them properly is possible
- … but requires principled designs
- But it won’t happen with an “it’s all been done in the ‘70s” attitude