

ReL4：高性能异步微内核设计与实现

廖东海

**** 年 * 月

ReL4:

高性能异步微内核设计与实现

廖东海

北京理工大学

中图分类号： TQ028.1

UDC分类号： 540

- ★ 特别类型
- ☐ 交叉研究方向
- ☐ 政府项目留学生

ReL4： 高性能异步微内核设计与实现

作 者 姓 名	廖东海
学 院 名 称	计算机学院
指 导 教 师	陆慧梅
答辩委员会主席	***
申 请 学 位	工学硕士
学 科 / 类 别	计算机科学与技术
学位授予单位	北京理工大学
论文答辩日期	**** 年 * 月

ReL4: Design and Implementation of High-performance Asynchronous Microkernel

Candidate Name:	<u>Liao Donghai</u>
School or Department:	<u>****</u>
Faculty Mentor:	<u>Lu Huimei</u>
Chair, Thesis Committee:	<u>***</u>
Degree Applied:	<u>****</u>
Major:	<u>****</u>
Degree by:	<u>Beijing Institute of Technology</u>
The Date of Defence:	<u>*, ****</u>

研究成果声明

本人郑重声明：所提交的学位论文是我本人在指导教师的指导下独立完成的研究成果。文中所撰写内容符合以下学术规范（请勾选）：

☐ 论文综述遵循“适当引用”的规范，全部引用的内容不超过 50%。

☐ 论文中的研究数据及结果不存在篡改、剽窃、抄袭、伪造等学术不端行为，并愿意承担因学术不端行为所带来的一切后果和法律责任。

☐ 文中依法引用他人的成果，均已做出明确标注或得到许可。

☐ 论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果。

☐ 与本人一同工作的合作者对此研究工作所做的任何贡献均已在学位论文中作了明确的说明并表示了谢意。

特此声明。

签 名：

日期：

关于学位论文使用权的说明

本人完全了解北京理工大学有关保管、使用学位论文的规定，其中包括：

① 学校有权保管、并向有关部门送交学位论文的原件与复印件；

② 学校可以采用影印、缩印或其它复制手段复制并保存学位论文；

③ 学校可允许学位论文被查阅或借阅；

④ 学校可以学术交流为目的，复制赠送和交换学位论文；

⑤ 学校可以公布学位论文的全部或部分内容（保密学位论文在解密后遵守此规定）。

签 名：

日期：

导师签名：

日期：

摘要

随着科技的发展，微内核广泛应用于工控系统、嵌入式系统等领域。相比于宏内核，微内核将内存管理、设备驱动、文件系统等与核心功能分离，运行在用户空间，这种隔离机制使得单个服务的故障不会直接影响到内核和其他服务，从而提升了系统的整体稳定性；微内核通过精简核心功能，减少攻击面，从而提升内核安全性；此外，模块化的设计使得系统更易于维护和升级。在微内核系统中，应用程序通过进程间通信 (IPC) 而非系统调用来请求服务，这或许能够满足早期性能不敏感的软件需求，然而在对软件性能有着更高要求的今天，频繁的 IPC 造成的特权级切换将产生巨大开销，成为了系统的性能瓶颈。

以 seL4 为代表的现代微内核将同步 IPC 作为主要通信手段，通过组合系统调用、fast-path 等优化手段提升 IPC 性能，同时更好地利用硬件并行性，简化并发模型，seL4 在内核中支持异步的通知机制，这些机制在一定程度上减少了 IPC 的开销，提升了系统性能，然而它们在设计上仍有三点不足：1) 在支持同步 IPC 的情况下冗余地支持了异步通知机制，这违反了内核最小化原则；2) 通知机制依赖内核的转发，会造成大量的特权级切换；3) 系统调用和同步 IPC 会导致无关请求顺序执行，无法充分利用硬件资源。为了解决以上三点缺陷，本文设计并实现了 ReL4——一套异步化的高性能微内核架构。为了保证内核最小化原则，ReL4 仅支持异步的通知机制，IPC 在用户态通过通知机制和共享缓冲区实现。而为了避免通知机造成的大量特权级切换，ReL4 基于特殊硬件（用户态中断和 TAIC），在兼容 seL4 接口的基础上，设计了无需内核转发的 U-notification 机制；最后，为了避免无关 IPC 和系统调用请求的顺序执行，ReL4 通过异步运行时来实现异步 IPC 和异步系统调用，更充分利用硬件资源，提升系统并发度。本文在 FPGA 上实现了 ReL4 的原型系统，并与 seL4 进行了全面的对比测试，测试结果表明，ReL4 在减少特权级切换、提升系统并发性方面有着显著优势。

关键词：微内核；异步；进程间通信；用户态中断

Abstract

With the advancement of technology, microkernels have been widely applied in industrial control systems, embedded systems, and other domains. Compared to monolithic kernels, microkernels separate functionalities such as memory management, device drivers, and file systems from the core kernel, running them in user space. This isolation mechanism ensures that failures in individual services do not directly impact the kernel or other services, thereby enhancing the overall stability of the system. By streamlining core functionalities, microkernels reduce the attack surface, thereby improving kernel security. Additionally, the modular design makes the system easier to maintain and upgrade. In microkernel systems, applications request services through inter-process communication (IPC) rather than system calls, which might have sufficed for early software with less performance sensitivity. However, in today's context where higher software performance is demanded, frequent IPC-induced privilege-level switches incur significant overhead, becoming a performance bottleneck for the system.

Modern microkernels, represented by seL4, employ synchronous IPC as the primary communication mechanism. They optimize IPC performance through a combination of system calls, fast-path optimizations, and better utilization of hardware parallelism, while simplifying concurrency models. seL4 supports asynchronous notification mechanisms within the kernel, which to some extent reduce IPC overhead and improve system performance. However, their design still has three shortcomings: 1) redundantly supporting asynchronous notification mechanisms alongside synchronous IPC violates the principle of minimality in kernel design; 2) the notification mechanism relies on kernel forwarding, resulting in numerous privilege-level switches; and 3) system calls and synchronous IPC cause unrelated requests to execute sequentially, failing to fully utilize hardware resources. To address these three limitations, this paper designs and implements ReL4—a high-performance, asynchronous microkernel architecture. To adhere to the principle of minimality, ReL4 supports only asynchronous notification mechanisms, with IPC implemented in user space through notifications and shared buffers. To avoid the excessive privilege-level switches caused by notification mechanisms, ReL4 introduces U-notification, a kernel-forwarding-free mechanism based on

user-level interrupts, while maintaining compatibility with seL4's interface. Finally, to prevent the sequential execution of unrelated IPC and system call requests, ReL4 employs an asynchronous runtime to implement asynchronous IPC and asynchronous system calls, better leveraging hardware resources and improving system concurrency. This paper implements a prototype of ReL4 on an FPGA and conducts comprehensive comparative tests with seL4. The results demonstrate that ReL4 exhibits significant advantages in reducing privilege-level switches and enhancing system concurrency.

Key Words: microkernel; asynchronous; inter-process communication; user-mode interrupt

目录

附录 A	***	1
附录 B	Maxwell Equations	2
攻读学位期间发表论文与研究成果清单		3
致谢		4
作者简介		5

插图

表格

主要符号对照表

BIT	北京理工大学的英文缩写
\LaTeX	一个很棒的排版系统
$\text{\LaTeX 2}_{\epsilon}$	一个很棒的排版系统的最新稳定版
$\text{X}_{\text{\LaTeX}}$	\LaTeX 的好兄弟，事实上他有很多个兄弟，但是这个兄弟对各种语言的支持能力都很强
ctex	成套的中文 \LaTeX 解决方案，由一帮天才们开发
H_2SO_4	硫酸
$e^{\pi i} + 1 = 0$	一个集自然界五大常数一体的炫酷方程
$2\text{H}_2 + \text{O}_2 \longrightarrow 2\text{H}_2\text{O}$	一个昂贵的生成生命之源的方程式

附录 A ***

附录相关内容…

附录 B Maxwell Equations

因为在柱坐标系下， $\bar{\mu}$ 是对角的，所以 Maxwell 方程组中电场 \mathbf{E} 的旋度
所以 \mathbf{H} 的各个分量可以写为：

$$H_r = \frac{1}{\mathbf{i}\omega\mu_r} \frac{1}{r} \frac{\partial E_z}{\partial \theta} \quad (\text{B-1a})$$

$$H_\theta = -\frac{1}{\mathbf{i}\omega\mu_\theta} \frac{\partial E_z}{\partial r} \quad (\text{B-1b})$$

同样地，在柱坐标系下， $\bar{\epsilon}$ 是对角的，所以 Maxwell 方程组中磁场 \mathbf{H} 的旋度

$$\nabla \times \mathbf{H} = -\mathbf{i}\omega\mathbf{D} \quad (\text{B-2a})$$

$$\left[\frac{1}{r} \frac{\partial}{\partial r} (r H_\theta) - \frac{1}{r} \frac{\partial H_r}{\partial \theta} \right] \hat{\mathbf{z}} = -\mathbf{i}\omega\bar{\epsilon}\mathbf{E} = -\mathbf{i}\omega\epsilon_z E_z \hat{\mathbf{z}} \quad (\text{B-2b})$$

$$\frac{1}{r} \frac{\partial}{\partial r} (r H_\theta) - \frac{1}{r} \frac{\partial H_r}{\partial \theta} = -\mathbf{i}\omega\epsilon_z E_z \quad (\text{B-2c})$$

由此我们可以得到关于 E_z 的波函数方程：

$$\frac{1}{\mu_\theta\epsilon_z} \frac{1}{r} \frac{\partial}{\partial r} \left(r \frac{\partial E_z}{\partial r} \right) + \frac{1}{\mu_r\epsilon_z} \frac{1}{r^2} \frac{\partial^2 E_z}{\partial \theta^2} + \omega^2 E_z = 0 \quad (\text{B-3})$$

攻读学位期间发表论文与研究成果清单

(二) 发表的学术论文

- [1] XXX, XXX. Static Oxidation Model of Al-Mg/C Dissipation Thermal Protection Materials[J]. Rare Metal Materials and Engineering, 2010, 39(Suppl. 1): 520-524. (SCI收录, IDS 号为 669JS, IF=0.16)
- [2] XXX, XXX. 精密超声振动切削单晶铜的计算机仿真研究 [J]. 系统仿真学报, 2007, 19 (4) : 738-741, 753. (EI 收录号: 20071310514841)
- [3] XXX, XXX. 局部多孔质气体静压轴向轴承静态特性的数值求解 [J]. 摩擦学学报, 2007 (1) : 68-72. (EI 收录号: 20071510544816)
- [4] XXX, XXX. 硬脆光学晶体材料超精密切削理论研究综述 [J]. 机械工程学报, 2003, 39 (8) : 15-22. (EI 收录号: 2004088028875)
- [5] XXX, XXX. 基于遗传算法的超精密切削加工表面粗糙度预测模型的参数辨识以及切削参数优化 [J]. 机械工程学报, 2005, 41 (11): 158-162. (EI 收录号: 2006039650087)
- [6] XXX, XXX. Discrete Sliding Mode Control with Fuzzy Adaptive Reaching Law on 6-PEES Parallel Robot[C]. Intelligent System Design and Applications, Jinan, 2006: 649-652. (EI 收录号: 20073210746529)

(二) 申请及已获得的专利 (无专利时此项不必列出)

- [1] XXX, XXX. 一种温热外敷药制备方案: 中国, 88105607.3[P]. 1989-07-26.

(三) 参与的科研项目及获奖情况

- [1] XXX, XXX. XX 气体静压轴承技术研究, XX 省自然科学基金项目. 课题编号: XXXX.
- [2] XXX, XXX. XX 静载下预应力混凝土房屋结构设计统一理论. 黑龙江省科学技术二等奖, 2007.

致谢

本论文的工作是在导师……。

作者简介

本人…。