



操作系统设计创新的深度分析及未来展望

In-depth Analysis of Innovations in Operating
System Design and Future Prospects

学 校 浙江大学

学 院 软件学院

专 业 人工智能

姓 名 王伟桥

指导教师 赵新奎

2025 年 10 月 20 日

传统操作系统（如 Unix/Linux）的核心架构已延续数十年，但近年来计算硬件和应用环境发生了巨大变化。现代数据中心服务器拥有数百核心、海量内存，并广泛使用 GPU、TPU 等异构硬件；同时，云计算、大规模机器学习、大数据分析和无服务器（Serverless）等新范式对操作系统提出了更高的性能、弹性和快速调度需求。

现有操作系统在资源利用效率、扩展性、异构支持、安全隔离和可观测性等方面逐渐暴露局限。Linux 等通用 OS 变得庞大且难以适应新需求。这些趋势促使学术界和工业界重新思考操作系统设计。在此背景下，本文对 LegoOS、DBOS 和 FlexOS 三项代表性研究成果进行深入分析。它们分别代表了当前 OS 研究的三个重要方向：硬件层面的资源解耦（LegoOS）、软件架构层面的数据中心 OS 重构（DBOS），以及安全机制层面的 OS 模块化与专用化（FlexOS）。本文将分析它们的核心技术、优劣，并结合 AI 时代趋势展望操作系统未来的演进方向。

LegoOS 旨在突破传统以单机服务器为单位的 OS 设计，以应对“硬件资源解耦”趋势。传统单机单内核模式导致数据中心资源利用不均衡。LegoOS 提出“分离内核（Splitkernel）”架构，将 OS 功能解耦成多个松耦合的管理单元（监视器，monitor），如进程监视器、内存监视器和存储监视器，它们分别驻留在不同类型的硬件组件上。这些监视器像“乐高积木”一样按需组合，通过高速网络（如 RDMA）和显式消息传递协作，不依赖共享内存或硬件缓存一致性。用户实例被抽象为“虚拟节点（vNode）”，其资源可来自多个不同硬件组件。为缓解 CPU 与远程内存分离带来的性能损失，LegoOS 在计算节点设计了“扩展 DRAM 缓存（ExCache）”，利用本地少量 DRAM 充当软缓存。其优势在于：1. 资源池化与弹性：打破了服务器资源孤岛，实现资源池化和动态组合，显著提高内存利用率。2. 独立扩展与容错：硬件组件可独立升级，故障域缩小到组件级别，提高了容错性。3. 架构创新：验证了 CPU 与内存分离这种极端架构的可行性，开创了“分布式多组件操作系统”的新类别。其局限在于：网络延迟开销是资源解耦不可避免的问题。对于内存密集型负载，性能可能显著下降。放弃硬件缓存一致性对应用透明度提出了挑战。

DBOS (Database-Oriented Operating System) 提出一个革命性思想：将分布式关系数据库作为操作系统的内核。其核心主张是“一切系统状态皆表”（Everything is a table）。操作系统中所有状态（如进程、任务队列、文件目录、通信消息等）都被存储在分布式数据库的表格中。操作系统提供的各项服务（如调度、IPC、文件系统）则被实现为对这些数据库表的标准 SQL 事务操作。DBOS 认为管理 OS 状态本质上是一个“大数据管理”问题。其架构底层是微内核，之上是高性能分布式内存数据库（原型使用 VoltDB），OS 服务则作为 SQL 存储过程实现。其优势在于：1. 极高可观测性：由于所有 OS 状态都在数据库表中，管理员可随时通过 SQL 查询和分析系统运行状况。2. 内置高可用与一致性：利用数据库的 ACID 事务和复制机制，OS 服务自动具备强一致性和容错能力。3. 大幅简化代码：借助成熟 DBMS，OS 服务实现复杂度大降。论文称，DBOS 原型用数十行 SQL 实现了 Linux 数万行 C 代码的文件系统功能。4. “时间旅行”调试：可利用数据库快照将系统恢复到过去某一时刻状态，方便调试和容灾。其局限在于：性能开销与实时性是主要挑战，数据库事务难免引

入额外延迟。系统高度依赖底层数据库的性能和可靠性。将所有 OS 服务抽象为关系模型是否对所有场景都合适存疑。

FlexOS 关注操作系统的安全隔离和定制优化。传统 OS（如 Linux 或微内核）的隔离策略在设计时就已固化，难以同时满足应用对高性能和强安全的多样化需求。FlexOS 提出“可定制隔离”，允许开发者在编译时为 OS 的不同组件选择不同的隔离机制。FlexOS 基于 Library OS (LibOS) 模型，将 OS 内部进一步拆分为文件系统、网络栈等细粒度组件。开发者可指定组件间采用何种隔离：从同一地址空间内的轻隔离（软件检查），到使用硬件页表 (MMU) 的强隔离，乃至不同特权级或虚拟机隔离。通过配置，FlexOS 可生成从高性能“unikernel”（无隔离）到高安全“microkernel”（完全隔离）的多种 OS 形态。其优势在于：1. 灵活的安全-性能权衡：开发者可根据应用场景（如对可信组件减少隔离以提速，对不可信模块加强隔离以保安）定制 OS。2. 适配新硬件：能灵活利用 Intel MPK、ARM TrustZone 等新型硬件隔离特性。3. 性能竞争力：实验表明，在等价隔离配置下，FlexOS 性能不输 Linux，甚至更优。

LegoOS、DBOS 和 FlexOS 在操作系统创新谱系中各有侧重。在关注层面，三者并不相互排斥，反而可能互补。LegoOS 聚焦数据中心硬件架构的变革，解决硬件资源如何跨节点组合管理。DBOS 关注系统软件栈的重构，试图统一集群层次的 OS 服务和数据管理。FlexOS 立足于单机 OS 内部，探索 OS 模块划分与隔离策略的弹性调整。

在性能关注点上，LegoOS 牺牲一定的访问延迟换取资源弹性和更高吞吐量。DBOS 强调其宏观性能与传统方案“相当”，但微观操作上有延迟代价。FlexOS 追求在相同隔离级别下不比 Linux 慢，实现可调的性能/安全平衡。

在安全机制上，FlexOS 核心即灵活隔离，能组合多种软硬件隔离手段，定制化安全策略。LegoOS 将故障和安全隔离的粒度强化到部件级别（如内存节点、存储节点）。DBOS 贡献主要体现在数据完整性和审计方面，所有状态变更可追溯，且易于统一访问控制。

面向 AI 应用的迅猛发展，操作系统需要顺应新的趋势。OS 本身的策略（如调度、预取、缓存）越来越复杂，超出人工调参能力。AI 技术（如强化学习）可用于 OS 自适应优化。未来 OS 可配有“OS 大脑”模块，持续分析系统数据，自主调整参数。DBOS 提供的丰富可查询状态数据，为此类智能优化奠定了坚实基础。

随着生成式 AI 的发展，未来可能出现用自然语言操纵操作系统的场景（如通过对话调整资源或检查问题）。这需要 OS 提供高层抽象接口，与 DBOS 的可查询 OS 状态理念相吻合。

综上，LegoOS、DBOS、FlexOS 的创新展现了可能的路径：一个弹性（LegoOS 的解耦与重构）、智能（DBOS 的数据化与自优化）、安全（FlexOS 的灵活隔离）的数据中心操作系统，成为 AI 时代计算基础的坚实支撑。