

课程报告：

1、论文总结

1.1 LegoOS

LegoOS主要针对传统的单体服务器的资源利用低且存在硬件单点故障等问题而提出Splitkernel 架构，通过将OS的功能部件分散到多个monitor上管理硬件资源，仅通过网络通信来实现分布式服务器，对外仍包装成一个传统服务器。同时在cache、网络协议和内存故障等方面提出了创新。

1.2 DBOS

DBOS主要针对传统的Unix/Linux系统在现代计算场景中表现出来的并行计算和集群资源管理能力等不足而提出的，通过采用以分布式事务型数据库管理系统（DBMS）为核心构建来全新操作系统。DBOS以“分布式事务型 DBMS”为内核，构建四层架构，核心是将所有 OS 状态表示为关系表，通过 SQL 实现 OS 服务，避免冗余并提升可扩展性。

1.3 FlexOS

FlexOS主要针对现代操作系统在设计阶段锁定的安全隔离策略因为依赖固定的硬件 / 软件保护机制在部署后修改需大规模重构存在显著局限性的问题而提出的。FlexOS是基于模块化库操作系统（LibOS）Unikraft 构建的新型 OS，核心是将隔离策略的决策从设计阶段推迟到编译 / 部署阶段，支持用户按需定制安全隔离方案。

2、未来趋势

2.1、操作系统会往分布式的方向发展

随着大模型等需要大量算力资源的方向兴起，现在的操作系统在应对调度和通信方面仍存在短板，一个操作系统能合理利用卡的数量是有限的，当遇到参数相当大的模型时，可能会面临训练或者推理等方面的问题，根据LegoOS到DBOS的发展理念，操作系统以后会往越来越大的分布式集群发展，来应对AI时代的需求。

2.2、操作系统会往智能化方面发展

随着智能体等大模型应用产生，在操作系统方面也有思想启发，现代操作系统对一个应用总会分配最大能用到的内存，无法细粒度调节内存的分配，这在内存利用上是存在不小的浪费的，所以智能体 + 操作系统可能可以相对较好根据实际运行情况，来对资源分配等方面进行细粒度的操作。

3、多模态大模型 + 操作系统

在我的研究领域多模态大模型中，模型的训练与推理过程与底层操作系统存在着紧密且深刻耦合关系。具体而言，多模态大模型的训练和推理主要和操作系统两大核心能力：调度能力与网络通信能力存在关系。前者关乎如何高效地协同CPU、GPU等异构计算资源，以应对模型不同阶段的算力和内存等调度需求；后者则直接决定了分布式训练中海量参数与数据的交换效率。

此外，受到FlexOS等前沿研究的启发，我们认识到操作系统可以为不同的模型或任务提供定制化的安全与资源隔离环境。通过构建细粒度的执行沙箱，不仅能保障多租户场景下的模型数据安全，还能为开发者与用户提供一个更加灵活、便捷的部署与测试平台，从而提升开发与部署的效率。

然而，从另一个角度看，当前多模态大模型在与操作系统的关系中，更多是作为“上层应用”的角色，即利用OS提供的服务，或是被集成于应用中以提升用户交互的便捷性。至于利用AI为操作系统本身赋予“智能”，例如实现智能化的资源调度或故障预测，这一方向的可能性似乎并不局限于多模态模型。理论上，单模态的文本大模型凭借其强大的逻辑推理与代码理解能力，或许已足以胜任此类任务。

