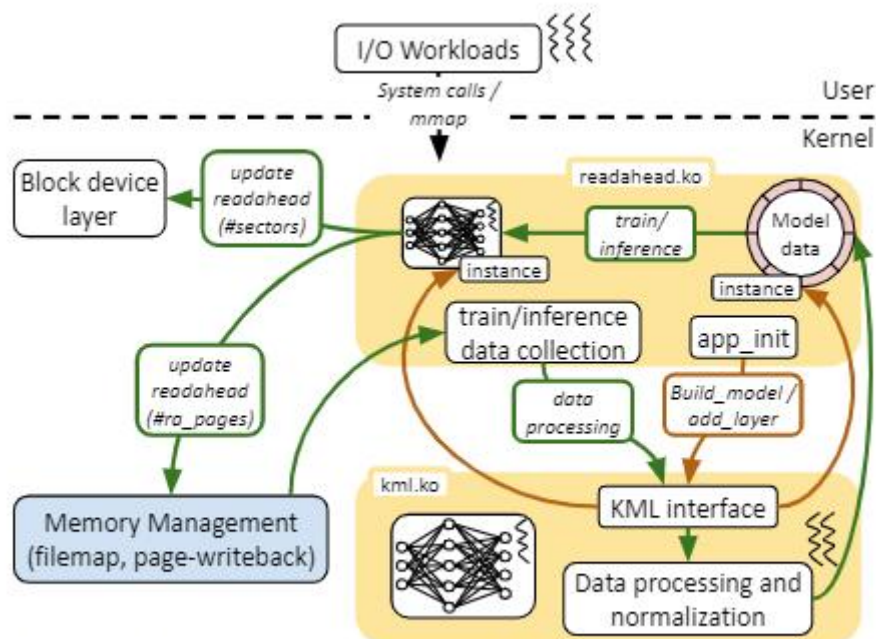


1. Amazon EBS 卷中预制的 SSD 支持 256kIOPS
2. 英特尔与 SAP 合作，为 SAP 工作负载提供支持英特尔 Optane DC 持久内存的 GCP 虚拟机，有益于处理和存储大量数据并运行 SAP HANA 等内存数据库的客户。
3. 打不开网站
4. 各行各业的组织都在生产中使用 eBPF 对内核进行动态编程，以实现高效的联网、可观察性、跟踪和安全性。
5. Graphcore 的官方网站，使用 Graphcore 加速实验。
6. Microsoft Azu 推出面向 SAP HANA 的新产品创新，扩大与 SAP 的 AI 合作。
7. Nvidia 发布了开源的 GPU 内核模块。
8. Nvidia 在 cuDNN5 中优化递归神经网络，支持递归神经网络的四 RNN 模式：ReLU 激活函数、tanh 激活函数、门控循环单元（GRU）和长短期记忆（LSTM）。
9. DPDK 文档
10. TensorFlow 是一种机器学习系统，在异构环境中大规模运行，支持各种应用，专注于深度神经网络的训练和推理。
11. 机器学习（ML）技术能够识别模式、抽象化模式并对新数据进行预测。ML 可以成为优化和调整存储系统的关键组件。本文提出 KML，一种用于存储系统的 ML 框架，可以在小内存占用的情况下提高吞吐量。



**Figure 1: Kernel space training/inference architecture.**

12. KMLib 将 ML 生态系统引入作系统，但功能有限，且不支持加速器。
13. GPU 是许多通用工作负载（如机器学习）的首选平台。这推动了在云等虚拟化环境中对更好的 GPGPU 支持的需求。在 GPGPU 计算中，虚拟机管理程序介导的 API 远程处理是高效的，将设备虚拟化与 GPU ISA 虚拟化分离是保持 GPGPU 加速的原始加速的关键。
14. Linux 通常与高性能集群环境中的并行文件系统一起使用，这些环境中的巨大存储增

长导致需要多错误纠正代码。通常在存储集群上运行的并行文件系统需要存储计算的结果,而无需巨大的等待时间。虽然这种存储集群的前端可以使用标准 PC 构建,但相比之下,几乎不可能使用最终用户硬件构建功能强大的 RAID 后端。文章调查了图形卡对于 Linux 内核中的 RAID 等编码应用的潜力。为此,设计了一个特殊的微驱动程序概念(梭子鱼),可以集成到 Linux 中,而无需更改内核 API。

15. 由于 SSD 的 I/O 访问延迟较高,因此仍然是计算机系统的性能瓶颈。改善访问延迟的一种常见方法是预取,在本文中讨论了 SSD 中预取的挑战,解释了为什么以前的方法无法实现高精度,并提出了一种基于神经网络的预取方法,该方法的性能明显优于最先进的方法。
16. 使用基于机器学习的资源感知负载均衡器解决了在运行计算密集型工作负载的服务器中,对有限计算资源的不平衡需求会阻碍执行性能的问题。
17. 文章介绍了 Azure VM 工作负载的详细特征。
18. 文章研究了使用现有性能计数器在硬件中构建恶意软件检测器的可行性。
19. Kleio 是一个混合页面调度器,它将现有的轻量级、基于历史记录的混合内存数据分层方法与基于深度神经网络的新型智能放置决策相结合。Kleio 采用了一种新方法来确定页面的优先级,从而带来最高的性能提升,同时限制由此产生的系统资源开销。
20. 文章从软件的角度描述了使用 OSIRIS 网络适配器的经验,它描述了 OSIRIS 提供的机会,我们能够在主机操作系统(OS)中利用这些机会。
21. 文章描述了使用 rCUDA(一个 GPGPU 虚拟化框架)在 VM 中执行 GPU 加速应用程序,从而在任何虚拟化环境中启用 GPGPU 功能。
22. 文章通过整合动态窗口大小,扩展了对入侵检测的系统调用异常检测建模方法的先前研究。
23. 一堆有关 LAKE-linux 的文件。
24. 一堆有关 LAKE 的文件
25. 介绍了 DGSF,这是一个平台,它使无服务器函数能够通过 CUDA 等通用 API 透明地使用 GPU。
26. DeepPrefetcher,这是一种受深度神经网络启发的新型上下文感知预取方法,可适应任意内存访问模式。
27. LoGV,一种通过利用现代硬件中已经存在的保护机制来虚拟化 GPGPU 的方法。LoGV 支持在 VM 之间共享 GPGPU 以及 VM 迁移,而无需修改主机驱动程序或访客的 CUDA 运行时。
28. 文章概述了新的 AES 指令,以及如何将其用于实现高性能和安全的 AES 处理。还描述了这种架构的一些特殊使用模型。
29. 文章介绍了 GViM,这是一个系统主题,旨在虚拟化和管理由图形过程处理器加速的通用系统的资源。
30. eCryptfs 是一个用于 Linux 的内核原生堆叠加密文件系统,在文件写入或从下级文件系统读取时对其进行加密和解密。文章介绍了 eCryptfs 的使用和安全说明。
31. PacketShader,这是一个高性能软件路由器框架,用于具有图形处理单元(GPU)加速的

一般数据包处理。PacketShader 利用 GPU 的大规模并行处理能力来解决当前软件路由器中的 CPU 瓶颈。

32. 文章介绍了 LinnOS,这是一种操作系统,它利用光神经网络以非常精细的方式推断 SSD 性能,并且帮助并行存储应用程序实现性能可预测性。
33. 文章提出了一种名为 Component Traversal 的新型动态分析方法,它可以尽可能完整地自动执行每个给定 Android 应用程序(应用程序)的代码例程。
34. ghOS 是我们用于将内核调度决策委托给用户空间代码的基础设施,为 Linux 环境中的用户空间进程提供调度策略的通用委托。ghOSt 提供状态封装、通信和作机制,允许在用户空间代理中复杂表达调度策略,同时协助同步。
35. Telekine 使应用程序能够在云中安全地使用 GPU 加速,它基于新颖的 GPU 流抽象,确保通过不受信任的组件执行和交互独立于任何机密数据。
36. 文章探讨了利用时间和空间倍增的几种技术,以提高 GPU 对深度学习推理工作负载的利用率。
37. 文章评估了自 2015 年以来在数据中心部署的一种称为张量处理单元(TPU)的自定义 ASIC,该 ASIC 可加速神经网络(NN)的推理阶段。
38. PACT,即按应用类 Turbo Controller,该系统利用两种新颖的机制来降低功耗,即使在高利用率的数据中心也是如此。
39. 文章探讨了存储内处理对当前固态硬盘(SSD)架构的好处和局限性。
40. 为 64 位英特尔处理器提供了反制模式的 AES 加密的比特实现,同时首次推出 AES-GCM 的恒定实现,其合理速度为 21.99 循环/字节,从而提供了一整套用于身份验证加密的时序分析软件。
41. AMORPHOS 将用户 FPGA 逻辑封装在 *morphable tasks* 或 Morphlets 中。Morphlets 在相互不信任的保护域中提供隔离和保护,从而扩展了软件流程的保证。
42. 文章提出了一种系统调用语言建模方法,用于设计基于异常的主机入侵检测系统。
43. 文章对当前 Linux 设计中的页面管理方案进行了深入分析,将 NUMA 扩展到支持配备 DRAM 和 SCM (Intel 的 DCPMM) 的系统。提出的技术通过释放多层内存层次结构的潜力来显着提高各种工作负载的性能。
44. 文章介绍了 NBA 框架,该框架扩展了 Click 模块化路由器的架构以利用现代硬件,适应不同的硬件配置,并在无需手动优化的情况下达到接近其最高性能。