学校编码：10384
学　　号：200215001

厦門大學

博　士　学　位　论　文

基于 eBPF 的异构计算系统下的 serverless 安全容器应用加速技术研究

Capital Reorganization: The Best Choice for A State-Owned Enterprise with Financial Crisis

熊若凡

指导教师姓名：张一鸣教授

专　业　名　称：　计算机技术

论文提交日期：　　年　月

论文答辩日期：　　年　月

学位授予日期：　　年　月

20　年　月

厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下，独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　）课题（组）的研究成果，获得（　　）课题（组）经费或实验室的资助，在（　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

本人声明该学位论文不存在剽窃、抄袭等学术不端行为，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。

声明人　（签名）：

指导教师（签名）：

年　　月　　日

厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1.经厦门大学保密委员会审查核定的涉密学位论文，于　　年　　月　　日解密，解密后适用上述授权。

（　　）2.不涉密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。涉密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

学位论文答辩委员会名单

主席	赵 XX	厦门大学	教授	博士生导师
委员	刘 XX	厦门大学	教授	博士生导师
	杨 XX	中国 XXXX 科学院 XXXXXXX 研究所	研究员	博士生导师
	黄 XX	XXXX 大学	教授	博士生导师
	周 XX	XXXX 大学	副教授	硕士生导师
秘书	吴 XX	厦门大学	助理教授	-

答辩时间：2020 年 5 月 10 日上午

答辩地点：厦门大学海韵园行政 C-505 会议室

摘　要

论文的摘要是对论文研究内容和成果的高度概括。摘要应对论文所研究的问题及其研究目的进行描述，对研究方法和过程进行简单介绍，对研究成果和所得结论进行概括。摘要应具有独立性和自明性，其内容应包含与论文全文同等量的主要信息。使读者即使不阅读全文，通过摘要就能了解论文的总体内容和主要成果。论文摘要的书写应力求精确、简明。切忌写成对论文书写内容进行提要的形式，尤其要避免“第 1 章……；第 2 章……；……”这种或类似的陈述方式

关键词：形状记忆；聚氨酯；织物；合成；应用（一般选 3～8 个单词或专业术语，且中英文关键词必须对应。）

Abstract

An abstract of a dissertation is a summary and extraction of research work and contributions. Included in an abstract should be description of research topic and research objective, brief introduction to methodology and research process, and summarization of conclusion and contributions of the research. An abstract should be characterized by independence and clarity and carry identical information with the dissertation. It should be such that the general idea and major contributions of the dissertation are conveyed without reading the dissertation.

Keywords: shape memory properties; polyurethane; textile; synthesis; application

目 录

Table of Contents

主要缩略词表

缩写	英文全称	中文含义

主要符号表

符号	中文含义

图索引

表索引

第 1 章　绪论

1.1　研究背景及选题意义

1.1.1　研究背景

无服务器计算已成为当今云和数据中心基础架构的新兴范式。它使用一种单点服务或功能作为基本计算单元，该单元以多种方式轻松计算。首先，它可以帮助应用程序开发人员专注于核心逻辑，并将与基础架构相关的任务（例如自动缩放）留在无服务器平台上。其次，它采用了具有细粒度付费模式（例如 1MS [7]）的 pay-as-you-go 模型，以便用户可以节省未使用的计算资源的成本。第三，无服务器计算也使云提供商有益于他们可以更有效地管理资源。Serverless 的核心思想是将基础设施的管理责任交给云服务提供商，用户无需关心服务器的配置和管理，可以专注于业务逻辑的实现。然而，在 Serverless 环境中，如何高效、安全地运行容器化应用仍然是一个亟待解决的关键问题。

容器技术作为现代云计算架构的重要组成部分，广泛应用于 Serverless 平台中。容器通过提供轻量级的隔离和高效的资源利用，为应用的部署和弹性伸缩提供了便利。在 Serverless 平台上，容器化应用需要支持快速启动和高并发运行，这对容器管理系统的性能和可扩展性提出了更高的要求。Linux 内核中的 cgroup（控制组）机制是容器资源管理的基础，它提供了对容器资源（如 CPU、内存、网络带宽等）的隔离和限制。然而，现有 cgroup 机制在面对高并发、大规模容器启动的场景时，仍然存在性能瓶颈和可扩展性问题，亟需进行优化。

另一方面，eBPF（扩展的伯克利包过滤器）作为 Linux 内核的一项强大特性，已在网络、安全、性能监控等多个领域得到了广泛应用。eBPF 允许在内核中运行用户定义的程序，极大地扩展了内核的功能。eBPF Verifier 作为 eBPF 框架的一部分，主要用于在编译时静态检查 eBPF 程序的正确性和安全性，确保程序在运行时不会导致系统不稳定或安全漏洞。

此外，随着数据处理能力的提升，异构计算平台（如集成了 DPU 的服务器）在云计算中逐渐得到广泛应用。DPU（Data Processing Unit）作为一种专为数据处理优化的计算单元，可以显著提升网络处理、存储加速等方面的性能。然而，Serverless 平台中使用 DPU 加速计算时，如何确保多租户环境下的安全性，尤其是不同租户间对 DPU 资源的隔离，成为了一个重要问题。为了解决这一问题，利用静态程序

验证技术（如 eBPF verifier）来分析和保证代码的安全性和隔离性，成为了一个有效的解决方案。

1.1.2 选题意义

本研究的主要目标是提升 Serverless 环境中容器应用的安全性和性能，尤其是在异构计算平台下。研究的意义可从以下几个方面进行阐述：

优化 cgroup 机制提升并发启动性能随着 Serverless 架构的广泛应用，容器化的应用需要在高度动态和弹性的环境中快速启动和运行。Linux 内核的 cgroup 机制在容器的资源管理和隔离方面发挥了重要作用，但在面对高并发容器启动时，现有 cgroup 的性能和可扩展性仍然不足。通过对 cgroup 机制进行优化，可以有效提升 Serverless 平台中容器的启动效率，增强平台在高并发环境下的稳定性和扩展能力，满足现代云计算平台对高性能和高并发的需求。

保证 DPU 资源的安全隔离在 Serverless 环境中，DPU 作为一种硬件加速单元，可以显著提升计算效率。然而，DPU 作为共享资源在多租户环境中的使用，容易面临安全性和资源隔离的问题。如何保证不同租户在使用 DPU 时的资源隔离性，防止不同租户之间相互干扰，成为了一个亟待解决的难题。本研究通过引入类似 eBPF verifier 的静态程序验证技术，能够在运行时前对 DPU 访问代码进行验证，确保多租户环境下对 DPU 的安全访问，从而避免内存泄露、数据篡改等安全风险。

推动 Serverless 平台在异构计算环境中的应用发展目前，Serverless 架构的研究主要集中在通用的计算资源和虚拟化技术上，针对异构计算资源（如 DPU、GPU 等）的支持尚处于起步阶段。本研究不仅将研究如何在 Serverless 架构中高效、安全地使用 DPU 加速计算，还将探索如何使 Serverless 平台更好地适应异构计算环境，为异构计算资源的全面融合提供理论依据和技术支持。

促进 Serverless 技术的商业应用和发展随着云计算技术的不断发展，Serverless 架构在互联网企业和各类业务应用中的应用越来越广泛。通过提升 Serverless 平台在高并发、大规模容器管理和异构计算资源安全管理方面的能力，本研究的成果不仅具有重要的学术价值，同时对促进 Serverless 技术在商业应用中的推广和发展具有重要意义。

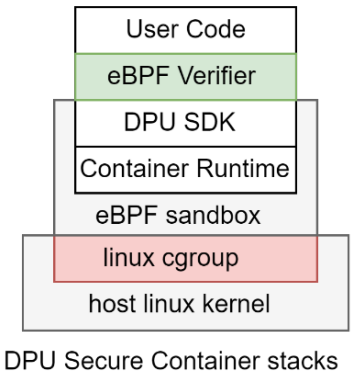

图 1.1 Overview

1.1.3 研究目标和内容

本论文的主要研究目标是通过优化 Linux 内核 cgroup 机制，提高 Serverless 环境下容器的并发启动性能；通过引入静态程序验证技术，保证 Serverless 平台中 DPU 资源的多租户安全隔离。具体内容包括：

优化 cgroup 机制：研究并优化 Linux cgroup 在高并发、大规模容器启动场景中的性能瓶颈，提高其在 Serverless 环境中的可扩展性和响应速度。eBPF 验证技术的应用：通过引入 eBPF verifier 等静态验证技术，确保 Serverless 平台中异构计算资源（如 DPU）的安全隔离，避免不同租户间的资源竞争和安全问题。通过以上研究，本论文旨在为异构计算平台下的 Serverless 架构提供高效、安全的容器应用加速技术，为下一代云计算平台的发展做出贡献。

1.2 国内外研究现状

1.3 本文主要内容和章节安排

1.3.1 Serverless 函数并发启动关键技术点

1.3.2 基于 eBPF 的 DPU 静态虚拟化

云计算技术对异构需求越来越高，传统架构存在着处理能力与数据量增长不匹配、资源利用不足、安全风险等问题。DPU 功能 serverless 化变得非常常见。例如英伟达推出的 DOCA 平台框架（DPF）是一个 DPU 程序编排框架，帮助创建 BlueField 加速的云软件平台，使得 DPU 可以在 K8s 环境中被直接使用。同时已有研究人员开始研究异构计算系统（CPU-GPU-DPU）上的 serverless 系统（Serverless

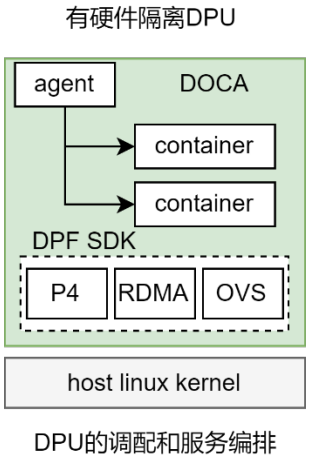

图 1.2 DPU 应用程序的调配和服务编排

Computing on Heterogeneous Computers ASPLOS’22）。

DPU 容器（DPU container）通常是指基于 DPU（数据处理单元）技术，使用容器化技术来部署和管理的计算环境。简单来说，它结合了 DPU 的硬件加速能力和容器的灵活性，通常用于高性能计算、网络加速、存储加速等场景。Nvidia DPU 具有硬件级隔离：NVIDIA BlueField DPU 通过硬件加速引擎和专用处理资源，为每个租户提供独立的计算、存储和网络资源，确保租户之间的资源隔离。在云原生计算平台中，DPU 通过卸载和加速基础设施功能，支持多租户环境下的高性能计算。DOCA 编程框架（DPF）可以实现多租户的 DPU 服务的调配和编排，通过容器化的方式将 Kubernetes 控制平面功能扩展到 DPU，使管理员能够直接在 BlueField DPU 上部署和卸载 NVIDIA DOCA 服务和基于 DOCA 的第三方服务。DPF 配备了专门用于无缝集成的 SDK，提供了一个一致的模块化工具包，例如 OVS，RDMA 开发工具包。

目前国内的 DPU（数据处理单元）在硬件级隔离方面的能力相对有限，要实现完全的硬件级隔离（特别是在安全性、可靠性、性能等方面的完全隔离）仍需要进一步的技术发展和硬件支持，尤其是是在 K8s 等云应用场景下。

1. 虚拟化和资源隔离：许多 DPU 提供了硬件加速的虚拟化支持，尤其是在网络、存储和计算资源的隔离方面。例如，DPUs 如浪潮的“云脉”DPU 或者华为的“昇腾”DPU，已经具备一定的隔离能力，能够在同一物理硬件上支持多个虚拟机或容器的运行，并为其提供网络加速、存储加速等服务。但这种隔离更多是针对资源访问的虚拟化层面的，而非完全的硬件隔离。

2. 安全性与可信计算：要实现完全的硬件级隔离，还需要依赖于更强的安全

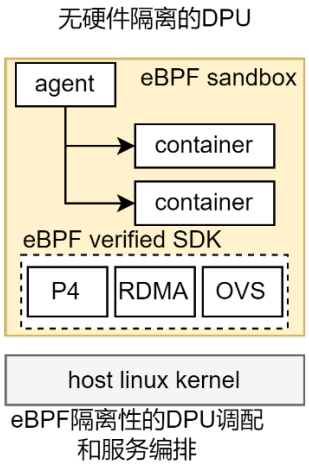

图 1.3 eBPF 隔离的 DPU 调配和服务编排

特性，比如硬件级的加密、完整性保护、物理隔离等机制。目前一些 DPU 已经支持基于硬件的安全功能（如加密引擎），但对于极高要求的安全场景（如政府、金融等领域），硬件级的完全隔离通常需要结合特殊的硬件设计和安全协议，甚至依赖于可信执行环境（TEE）技术。

3.DPU 与 CPU 的协同工作：目前的 DPU 在硬件级的完全隔离方面，可能仍然受到 CPU 和内存共享、I/O 访问控制等因素的制约。尤其是在一些复杂的并发计算和高性能计算场景下，虽然 DPU 可以通过硬件支持高效的网络处理和计算加速，但要完全独立于 CPU 的安全和资源调度依然是一个技术难点。

目前国内的这种不具备完全硬件隔离能力的 DPU 无法支持租户之间的资源隔离要求，阻碍了 DPU 容器化的应用。ebpf 借助 ebpf verfier 实现了一个安全的沙箱，可以防止 ebpf 代码以多种方式来损害内核（DOS 攻击、信息窃取攻击等）。Verifier 可以在一定程度上对容器内的代码实现安全性隔离（todo：需要找一些 case 分析）。同时，eBPF verifier 运行在程序加载阶段，不会给函数容器启动添加额外影响。通过 ebpf verifier 实现 ebpf sandbox 可以在软件层面对用户代码实现安全保障作用。加入 ebpf sandbox 后的 DPU 容器化模型如下图所示：

1. 硬件虚拟化与资源分区 SR-IOV（单根 I/O 虚拟化）DPU 的物理网卡通过 SR-IOV 虚拟化为多个虚拟功能（VF），每个租户的虚拟机（VM）或容器独占一个 VF，实现网络流量的直接硬件隔离，减少延迟并提升吞吐量。

硬件资源分片 DPU 的计算核心（如 Arm CPU、加速引擎）和内存资源通过硬件分区分片，为不同租户分配独立的计算单元和内存空间，防止资源争用。

第 2 章　总结与展望

本文采用……。（结论作为学位论文正文的最后部分单独排写，但不加章号。结论是对整个论文主要结果的总结。在结论中应明确指出本研究的创新点，对其应用前景和社会、经济价值等加以预测和评价，并指出今后进一步在本研究方向进行研究工作的展望与设想。结论部分的撰写应简明扼要，突出创新性。）

参考文献

附录 A　关于 XXX 的证明

附录相关内容…

附录 B　Maxwell Equations

因为在柱坐标系下，$\overline{\overline{\mu}}$ 是对角的，所以 Maxwell 方程组中电场 **E** 的旋度

所以 **H** 的各个分量可以写为：

$$H_r = \frac{1}{\mathbf{i}\omega\mu_r}\frac{1}{r}\frac{\partial E_z}{\partial\theta} \tag{B-1a}$$

$$H_\theta = -\frac{1}{\mathbf{i}\omega\mu_\theta}\frac{\partial E_z}{\partial r} \tag{B-1b}$$

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

$$\nabla\times\mathbf{H} = -\mathbf{i}\omega\mathbf{D} \tag{B-2a}$$

$$\left[\frac{1}{r}\frac{\partial}{\partial r}(rH_\theta) - \frac{1}{r}\frac{\partial H_r}{\partial\theta}\right]\hat{\mathbf{z}} = -\mathbf{i}\omega\overline{\overline{\varepsilon}}\mathbf{E} = -\mathbf{i}\omega\varepsilon_z E_z\hat{\mathbf{z}} \tag{B-2b}$$

$$\frac{1}{r}\frac{\partial}{\partial r}(rH_\theta) - \frac{1}{r}\frac{\partial H_r}{\partial\theta} = -\mathbf{i}\omega\varepsilon_z E_z \tag{B-2c}$$

由此我们可以得到关于 E_z 的波函数方程：

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

致　谢

本论文的工作是在导师……。

在学期间完成的相关学术成果

学术论文：

[1] 高凌. 交联型与线形水性聚氨酯的形状记忆性能比较 [J]. 化工进展, 2006, 532－535.（核心期刊）

[2] 杨轶, 张宁欣, 任天令, 等. 硅基铁电微声学器件中薄膜残余应力的研究 [J]. 中国机械工程, 2005, 16(14):1289-1291.

专利：

[3] 任天令, 杨轶, 朱一平, 等. 硅基铁电微声学传感器畴极化区域控制和电极连接的方法: 中国, CN1602118A[P]. 2005-03-30.

[4] Ren T L, Yang Y, Zhu Y P, et al. Piezoelectric micro acoustic sensor based on ferroelectric materials: USA, No.11/215, 102[P]. (美国发明专利申请号.)