

Rockchip Linux DPDK开发指导

文件标识：RK-KF-YF-DPDK

发布版本：V1.0.1

日期：2023-12-19

文件密级：☐绝密 ☐秘密 ☐内部资料 ☒公开

免责声明

本文档按“现状”提供，瑞芯微电子股份有限公司（“本公司”，下同）不对本文档的任何陈述、信息和内容的准确性、可靠性、完整性、适销性、特定目的性和非侵权性提供任何明示或暗示的声明或保证。本文档仅作为使用指导的参考。

由于产品版本升级或其他原因，本文档将可能在未经任何通知的情况下，不定期进行更新或修改。

商标声明

“Rockchip”、“瑞芯微”、“瑞芯”均为本公司的注册商标，归本公司所有。

本文档可能提及的其他所有注册商标或商标，由其各自拥有者所有。

版权所有 © 2023 瑞芯微电子股份有限公司

超越合理使用范畴，非经本公司书面许可，任何单位和个人不得擅自摘抄、复制本文档内容的部分或全部，并不得以任何形式传播。

瑞芯微电子股份有限公司

Rockchip Electronics Co., Ltd.

地址：福建省福州市铜盘路软件园A区18号

网址：www.rock-chips.com

客户服务电话：+86-4007-700-590

客户服务传真：+86-591-83951833

客户服务邮箱：fae@rock-chips.com

前言

概述

Rockchip Linux平台 DPDK 开发移植调试

产品版本

芯片名称	内核版本
RK3588/RK356X	Linux 4.19/5.10

读者对象

本文档（本指南）主要适用于以下工程师：

技术支持工程师

软件开发工程师

硬件开发工程师

修订记录

版本号	作者	修改日期	修改说明
V0.1.0	xy	2022-12	初始版本
V0.2.0	xy	2023-01	增加I320/350 PCIe网卡转发数据
V1.0.0	xy	2023-03	增加GMAC开发指导
V1.0.1	xy	2023-12	更新测试数据且把详细开发指导移动到专门的文档

Rockchip Linux DPDK开发指导

- 1. DPDK简介
- 2. DPDK工作原理
- 3. 平台支持情况
 - 3.1 RK3566
 - 3.2 RK3568
 - 3.3 RK3588
 - 3.4 RK3588S
- 4. DPDK开发指导
- 5. 性能指标
- 6. 支持列表

1. DPDK简介

DPDK全称Intel Data Plane Development Kit，是intel提供的数据平面开发工具集，目前支持为ARM、X86、PowerPC等处理器架构下用户空间高效的数据包处理提供库函数和驱动的支持，它不同于Linux系统以通用性设计为目的，而是专注于网络应用中数据包的高性能处理，目前已经验证可以运行在大多数Linux操作系统上；DPDK使用了BSD License，极大的方便了企业在其基础上来实现自己的协议栈或者应用，有助于更快地部署高性能网络应用程序，并实现比标准基于中断的内核网络堆栈更高效的运算。

除了基本的高性能网络 I/O 之外，DPDK 还提供了许多网络功能，例如LPM/ACL，这些功能用作上层应用程序使用的库，以下是 DPDK 中的基本框架。

2. DPDK工作原理

内核（左）和DPDK（右）的数据包处理

DPDK 使用多种技术来优化数据包吞吐量，但其工作方式（及其性能的关键）基于绕过内核和轮询模式驱动程序（PMD）。

- **Kernel Bypass**：在用户空间内创建从 NIC 到应用程序的路径，换句话说，绕过内核。这消除了为用户空间和内核空间之间移动网络帧时的上下文切换。此外，通过消除 NIC 驱动程序、内核网络堆栈和它们引入的性能损失，可以获得进一步的收益。
- **轮询模式驱动程序（PMD）**：当 CPU 收到帧时，CPU 不会引发 NIC 的中断，而是不断运行 PMD 以轮询 NIC 以获取新数据包。

除此之外，DPDK 还采用其他方法，例如 **CPU affinity**、**CPU isolation**、**大页面内存(huge page memory)**、**缓存行对齐(cache line alignment)**和**批量操作(bulk operations)**，以实现可能的最佳性能。

3. 平台支持情况

RK平台支持GMAC/PCIe两种接口的网卡，GMAC最多支持接两个千兆网卡；PCIe 支持x1/x2/x4类型的网卡，RK3588最多支持外接6个网卡：4个PCIe + 2个GMAC，各个芯片详细资源如下：

3.1 RK3566

资源	模式	支持芯片互联	支持lane拆分	支持DMA	支持MMU
PCIe Gen2 x 1	RC only	否	否	否	否
GMAC 1000M (x2)	RGMII	/	/	/	是

3.2 RK3568

资源	模式	支持芯片互联	支持lane拆分	支持DMA	支持MMU
PCIe Gen2 x 1 lane	RC only	否	否	否	否
PCIe Gen3 x 2 lane	RC/EP	是	1 lane RC+ 1 lane RC	是	否
GMAC 1000M (x2)	RGMII	/	/	/	是

3.3 RK3588

资源	模式	支持lane拆分	可用phy	内部DMA
PCIe Gen3 x 4 lane	RC/EP	是	pcie30phy	是
PCIe Gen2 x 1 lane (x3)	RC only	否	pcie30phy, combphy1_ps	否
GMAC 1000M (x2)	RGMII	/	/	是

3.4 RK3588S

资源	模式	支持lane拆分	可用phy	内部DMA
PCIe Gen3 x 1 lane	RC only	否	combphy2_psu	否
PCIe Gen2 x 1 lane	RC only	否	combphy0_ps	否
GMAC 1000M (x2)	RGMII	/	/	是

4. DPDK开发指导

- **GMAC方案**：参考rk_docs/Rockchip_Developer_Guide_Linux_GMAC_DPDK_CN.pdf
- **PCIE方案**：参考rk_docs/README_PCIE.txt

5. 性能指标

环境 (单核)	L2转发	L3转发
RK3568 GMAC + 64 Byte	580Mbit/s	500Mbit/s
RK3568 GMAC + 1500 Byte	980Mbit/s	980Mbit/s
RK3568 Intel I210/350 PCIe + 64 Byte	500Mbit/s	450Mbit/s
RK3568 Intel I210/350 PCIe + 1500 Byte	940Mbit/s	940Mbit/s
RK3568 RTL8111H PCIe + 64 Byte	300Mbit/s	260Mbit/s
RK3568 RTL8111H PCIe + 1500 Byte	940Mbit/s	940Mbit/s

注意：

1. 对于多核，可以按照85%左右的效率进行转换，比如单核1000Mbps，则8核为 $1000 \times 8 \times 0.85 = 6.8\text{Gbit/s}$ ；
2. N/A表示暂未实际测试数据后续补充，但L2一般都大于L3，它少了一个查找路由表的操作；

6. 支持列表

模块	接口	速率
GMAC+PHY	GMII	1000M
Intel I210	PCIe x1	1000M
Intel I350	PCIe x4	1000M x4
RTL8111H	PCIe x1	1000M