

S32G2 PFE产品简介

目录

1. 软件产品概述	1
2. 软件内容	4
2.1. 高级功能	9
3. 支持的目标	10
4. 质量、符合的标准和测试方法	11
5. 文档信息	12

1. 软件产品概述

面向S32G2硬件（HW）平台的数据包转发引擎（PFE）软件（SW）包括以下软件组件：

- PFE固件；
- 以太网驱动程序；
- 配置API-FCI（快速控制接口）；
- 管理守护程序CMM（Contrack监控模块）。

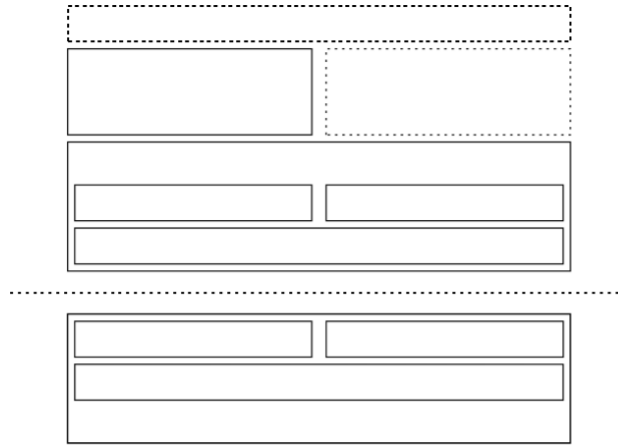


图1. 软件组件

所有组件或只有特定组件需要与目标环境集成，典型目标环境包括目标硬件（HW）平台、操作系统（OS）、操作系统提供的网络组件（网络栈、防火墙）和相关的用户应用程序。

固件

固件是在PFE中运行的软件组件。每个到达PFE（入口/出口）的数据包都由固件处理，固件负责数据包的分类、路由、修改和支持功能。固件受恩智浦的控制，可以根据目标用例的需要进行修改/扩展，以添加与以太网流量处理相关的特定功能。固件以二进制形式交付。

以太网驱动程序

以太网驱动程序通常是操作系统特定的组件，由三个主要功能模块组成：

- 平台驱动程序

包括PFE硬件的初始启动、配置、固件加载和PFE硬件组件的管理。通常，这部分可以被描述为低级PFE驱动程序，为硬件（包括固件）提供接口。

- 数据路径管理

与以太网数据流量管理相关的任务，涉及在PFE和网络栈之间传递数据包，包括主机接口（HIF）驱动程序的执行和与网络栈的连接。

- 控制路径管理

与PFE引擎配置相关的功能，FCI。

驱动程序运行在目标主机操作系统环境中，连接PFE与网络栈。它通过向操作系统开放逻辑接口来提供对物理以太网接口的访问，并以配置和监视PFE硬件和固件资源的形式实现控制路径域。如果主机CPU不需要运行完整驱动程序，但仍需要访问数据路径域（多核环境），驱动程序不仅可以以完整的形式使用，还可以以简单HIF驱动程序的形式使用。

- 主从模式运行时

驱动程序可以运行在同一系统的多个实例中，允许使用专用的主机接口为专用的CPU核共享PFE提供的连接。

配置API

为了让主机应用程序配置PFE功能，PFE 软件提供了名为“快速控制接口 (FCI)” 的配置API。它是运行在驱动程序中，实现形式为用户空间库和内核空间部分的类io-ctl接口。LibFCI提供的像fci_open()、fci_write()、fci_cmd()或fci_query()的函数可用于传递带参数的命令和接收响应。配置API实现与平台驱动程序紧密耦合，因此依赖于操作系统。

CMM

CMM是一种主机应用程序，用于简化一些PFE功能的配置，如L3转发器。它运行在主机操作系统中并监控网络栈。一旦检测到以太网数据流可能是快速转发的（由PFE而不是主机CPU路由），它就会使用LibFCI配置PFE来执行。用户无需编写自己的管理应用程序和配置PFE，只需使用本机操作系统提供的接口配置路由子系统。CMM可以被扩展到自动配置IPsec分流或L2桥接。有关CMM的可用性，请联系恩智浦代表。

2. 软件内容

本章涵盖了PFE模块和相关软件栈（固件、驱动程序、API）支持的特性的基本功能描述。

数据路径端点

这是固件和以太网驱动程序通过开放给操作系统和对应于PFE的特定物理接口的逻辑接口（PFEn）提供的标准数据流功能。它涵盖了：

- 最初的PFE平台配置（时钟、引脚、内存、固件等）；
- 各个物理/逻辑接口的MAC地址过滤；
- 可选的CRC生成和分流验证（IPv4/IPv6 TCP/UDP/ICMP，FCS）；
- 通过物理接口接收数据包，并通过逻辑接口将其提供给主机网络栈；
- 通过物理接口传输主机网络栈的数据包。

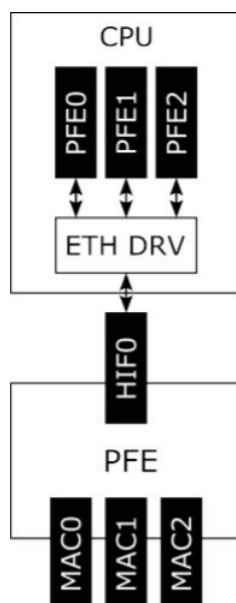


图2. 驱动程序和接口

IPv4/IPv6转发器/路由器

IPv4/IPv6转发器是种专用功能，可以将主机CPU从在两个物理接口之间转发特定IP流量的相关任务中解放出来。通常情况下，入口IP流量被传递到运行TCP/IP栈的主机CPU，该主机CPU负责路由数据包。一旦栈识别出数据包不属于任何本地IP端点，就会在路由表中执行查找，以确定

如何处理这些流量。如果路由表包含与数据包相关的条目（5元组搜索），则栈修改数据包并将其转发到另一个接口，使其到达目标节点。

PFE可以被配置为使用其内部路由表识别不需要进入主机CPU的数据流，并确保将适当的数据包转发到适当的目标接口。实现以下特性以支持转发功能：

- 路由表查找

PFE根据从入口数据包头字段（5元组：SRC/DST IP、SRC/DST端口、协议）解析的信息执行路由表查找。查找操作的输出是一个条目，包含用户编写的指令，指定了要对匹配数据包（目标接口、NAPT配置.....）执行的操作。

- NAPT

如果被指定，PFE根据NAPT设置执行数据包修改。它根据路由表条目中给定的配置的请求更新地址和/或端口。

- TCP连接端监测

PFE实施对快速转发流的监测。如果检测到SYN、FIN或RST数据包，PFE将通知发送到主机应用程序，对其进行适当处理（从本地表中删除数据流、更新防火墙.....）。

- 将数据包转发到目标接口

数据包一旦被路由和修改，就会被转发到目标接口。

L2 网桥（交换机）

L2网桥的功能包括基于MAC地址数据包转发。它提供了将桥接相关的任务从主机CPU转到PFE的可能性，从而分流基于主机的网络栈的负载。L2网桥的功能代表实现以下功能的网络交换设备：

- MAC表和地址学习

L2桥接功能基于确定入口数据包应转发到哪个接口。为此，网络交换设备实现所谓的桥接表（MAC表），搜索该表，为每个进入交换机的数据包获取目标接口。如果接收到数据，接收数据包的接口则被添加-被学习。然后，入口数据包的目标MAC地址被用来查表以确定目标接口。

- 时效性

每条MAC表项在被学习后都会获得默认超时值。这个超时时间会逐渐减少，直到为零。超时值为零的表项将自动从表中删除。每次相应的表项被用来处理数据包时，超时值都会重新设置。

- 端口迁移

如果在交换机的一个接口上看到MAC地址并且该地址表项已被创建，当在另一个接口上看到该MAC地址时表项会自动更新。

- VLAN感知

网桥实现VLAN表。此表用于执行基于VLAN的策略，如入口和出口成员身份。功能包括为每个网桥接口提供可配置VLAN标签和标签取消功能。

网桥利用PFE 硬件加速器执行MAC和VLAN表查找，因此该操作得到了高度优化。主机CPU软件仅负责使用专用API进行正确的网桥配置。

入口QoS

在大流量条件下，网络设备可能会经历所谓的“拥塞状态”。例如，这可能是由内部性能限制（例如有限的处理带宽，如路由大量小数据包或过于复杂/耗时的操作）和设备输入阶段的丢包结果引起。因此，输入接口是入口QoS发生的地方。其主要目的是避免拥塞和控制拥塞。为了支持这一点，入口QoS实现以下机制：

- 入口流量的分类：托管、未托管和保留
- 加权随机早期丢弃（WRED）
- 基于入口端口的速率整形器

这些机制允许用户对入口流量进行优先级排序，以低优先级流为代价确保高优先级流量被优先接收。

分类

分类是通过表实现的，用户可以编程数据流参数（L2/L3/L4），并指定数据流所属的流量类型（托管、未托管、保留）。也可以选择将某些数据流归类为“待丢弃”，以识别和终止入口模块内的恶意流，而不必到达下一个处理阶段。

WRED

在分类阶段之后，入口QoS执行WRED算法，在PFE达到拥塞状态时丢弃低优先级的流量。有多个可配置区域，每个区域具有不同阈值和丢弃可能性。被识别的区域越高，越高的丢弃可能性被WRED使用。



图3. WRED区域

整形器

通过WRED的帧被路由到最后一个入口QoS组件：基于端口的速率整形器。此整形器用于完整的端口流量。它实施标准的令牌桶算法来计算允许的字节数。用户可以配置其参数，充分利用特定入口接口能接收的最大数据速率（64 kbps至2.4 Gbps）。

入口QoS还提供统计信息，通知用户特定阶段的丢包数。

出口QoS

出口QoS位于PFE数据包处理管道的出口阶段。它以一组可配置拓扑的队列（8x）、调度器（2x）和整形器（4x）的形式在各个接口上实现。

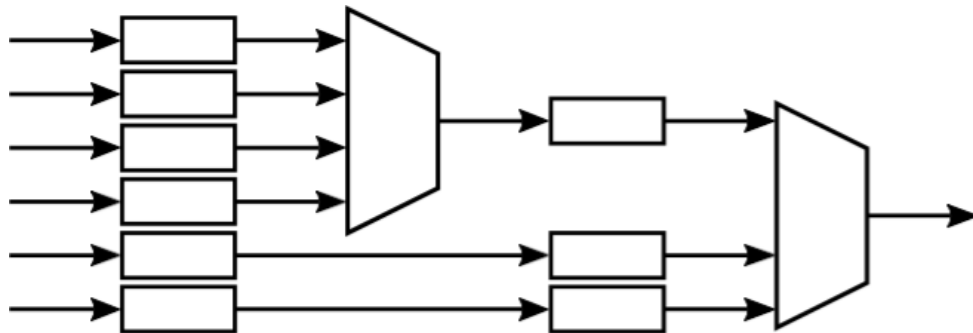


图4. 出口QoS拓扑示例

队列

队列是个实体，代表出口QoS模块的输入FIFO，用于临时保存要传输的数据包。它也可以被理解为一个接口，通过此接口属于某类流量的数据包将被传输。它的大小有限，一旦达到限值，当新数据包被丢弃时，所有将额外数据包入队已满队列的尝试都会以所谓的“尾部丢弃行为”终止。数据包被PFE分类器引擎写入特定队列。

流量类别与目标队列之间的映射由可配置的“优先级映射表”给出，用户可以配置PFE，根据IPv4/IPv6报头中的TOS/TrafficClass值将数据包映射到队列。队列中的数据包传输是使用调度器管理的，各种调度算法能被用来仲裁队列。

调度器

调度器是出口QoS的一个组件，负责多个队列的流量聚合。它的主要任务是从队列中选择接下来要传输的数据包。为此，调度程序实现以下调度规程：

- 简单轮询（RR）
- 差分加权轮询（DWRR）

每个队列都配置了特定的链路速度百分比作为权重。配额和赤字参数用于保证每个队列的承诺带宽和公平机会。

- 优先队列（PQ）

每个队列都被分配了优先级。在访问低优先级队列之前，高优先级队列必须被清空。

分配给单个调度器的最大输入数为8。

整形器

出口QoS执行801.Q基于信用的整形器，其目的是尽可能地分隔帧，减少数据峰值时的突发，并控制最大出口数据速率（以数据速率或包率为单位）。通常，整形器使用可配置的时钟嘀嗒进行操作。每一次嘀嗒，内部信用值就会根据配置的参数进行更新。当输入队列有数据且信用值为正时，队列将获得服务，而信用值将减少。当信用为负时，输入队列不会获得服务。用户可以根据应用的需要配置整形器的位置和特定参数。

灵活的解析器和灵活的路由器

PFE分类器对标准的L2/3/4报头字段进行查找，以做出路由决策。灵活的解析器允许专有字段成为路由决策的一部分。在S32G2上的GMAC的硬件中，实现了类似的方案。对于PFE，这是基于固件的解决方案。

Ipv4/Ipv6多播路由器

基本固件路由选项不支持在没有主机内核处理的情况下在多个端口发送多播/广播帧。此功能允许更新的路由表定义支持到多个队列/端口的帧复制、VLAN标签插入、TTL限制递减和MAC源插入。

IPsec

基本固件包括IPsec支持通过PFE内的实用PE将受保护的数据包分流到HSE。基本固件不使用HSE提供的额外的IPsec协议加速功能（即IPsec ESP封装和解封装）。固件修改利用了此功能并提高了解决方案的性能。

Can2Ethernet

在没有主机干预的情况下，路由CAN帧到以太网，反之亦然。所有这些运行在M0+LLCE核上。

2.1. 高级功能

IDPS

与Argus PFE 固件API协作，为客户库创建了入侵检测和保护系统扩展，支持在固件级对数据包流的特定客户分类。它提高了客户应用程序的性能并增加了灵活性。有关IDPS的更多详细信息，请与恩智浦的代表联系。

其他统计计数器

通过加速器在特定点进行数据包流分析。

3. 支持的目标

有关特定版本的详细信息，请参阅发行说明文档。

- o 支持的目标：S32G2处理器

4. 质量、符合的标准和测试方法

S32G2 PFE软件产品是根据恩智浦软件开发流程开发的，符合汽车SPICE、IATF16949和ISO 9001标准。

5. 文档信息

表1. 测试样本修订记录

版本号	日期	实质性变更
1	2021年10月	初版发布

How to Reach Us:

Home Page:
nxp.com

Web Support:
nxp.com/support

Information in this document is provided solely to enable system and software implementers to use NXP products. There are no express or implied copyright licenses granted hereunder to design or fabricate any integrated circuits based on the information in this document. NXP reserves the right to make changes without further notice to any products herein.

NXP makes no warranty, representation, or guarantee regarding the suitability of its products for any particular purpose, nor does NXP assume any liability arising out of the application or use of any product or circuit, and specifically disclaims any and all liability, including without limitation consequential or incidental damages. "Typical" parameters that may be provided in NXP data sheets and/or specifications can and do vary in different applications, and actual performance may vary over time. All operating parameters, including "typicals," must be validated for each customer application by customer's technical experts. NXP does not convey any license under its patent rights nor the rights of others. NXP sells products pursuant to standard terms and conditions of sale, which can be found at the following address: nxp.com/SalesTermsandConditions.

NXP, the NXP logo, NXP SECURE CONNECTIONS FOR A SMARTER WORLD, COOLFLUX, EMBRACE, GREENCHIP, HITAG, I2C BUS, ICODE, JCOP, LIFE VIBES, MIFARE, MIFARE CLASSIC, MIFARE DESFire, MIFARE PLUS, MIFARE FLEX, MANTIS, MIFARE ULTRALIGHT, MIFARE4MOBILE, MIGLO, NTAG, ROADLINK, SMARTLX, SMARTMX, STARPLUG, TOPFET, TRENCHMOS, UCODE, Freescale, the Freescale logo, AltiVec, C 5, CodeTEST, CodeWarrior, ColdFire, ColdFire+, C Ware, the Energy Efficient Solutions logo, Kinetis, Layerscape, MagniV, mobileGT, PEG, PowerQUICC, Processor Expert, QoriQ, QoriQ Converge, Ready Play, SafeAssure, the SafeAssure logo, StarCore, Symphony, VortiQa, Vybrid, Airfast, BeeKit, BeeStack, CoreNet, Flexis, MXC, Platform in a Package, QUICC Engine, SMARTMOS, Tower, TurboLink, and UMEMS are trademarks of NXP B.V. All other product or service names are the property of their respective owners. ARM, AMBA, ARM Powered, Artisan, Cortex, Jazelle, Keil, SecurCore, Thumb, TrustZone, and μ Vision are registered trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. ARM7, ARM9, ARM11, big.LITTLE, CoreLink, CoreSight, DesignStart, Mali, mbed, NEON, POP, Sensinode, Socrates, ULINK and Versatile are trademarks of ARM Limited (or its subsidiaries) in the EU and/or elsewhere. All rights reserved. Oracle and Java are registered trademarks of Oracle and/or its affiliates. The Power Architecture and Power.org word marks and the Power and Power.org logos and related marks are trademarks and service marks licensed by Power.org.

© 2021 NXP B.V.

Document Number: 1.6
Rev. 1.6
01/2022