

Xilinx SDNet:

定义网络硬件的新方法

作者: Loring Wirbel
高级分析师

2014 年 3 月



www.linleygroup.com

本文介绍了赛灵思的 SDNet 规范环境及其在定义软件定义网络 (SDN) 元素和在实现控制层和数据层可重配置网络元素中的作用。

SDN 前后的网络处理

近期基于 OpenFlow 等协议的软件定义网络大受追捧, 其渊源可以追溯到二十世纪 90 年代末的网络处理器。内联包处理职责与网络节点控制层相分离, 这就导致网络设备 OEM 厂商及其半导体供应商必须重新思考路由器和交换机的工作方式。采用片上控制单元的早期通信处理器随后被侧重于高效包转发和流量管理的器件所取代。

控制层处理器通常基于 MIPS 或 PowerPC 内核等 RISC 架构, 往往侧重于可编程性。设计人员对数据层优化包转发任务越来越熟悉, 也发现可通过硬件来编程和升级许多任务, 包括报头语法解析和位域处理等。一些流量管理和搜索表任务分配给后备处理器, 后备处理器通常采用 ASSO 实现, 大小和功能相对固定, 比说基于三元 CAM 的搜索引擎表大小。

赛灵思通过在 NPU 和软网络功能方面十多年的不懈努力, 现已推出了针对网络的 SDNet 软件定义规范环境。上一个十年中期, 特定领域规范环境作为一个研究项目, 在 Gordon Brebner 博士的带领下正式启动创建。2010 年, 赛灵思成立通信业务部, 将 SDNet 确定为一个商业化环境, 结合该研究项目, 对赛灵思设计环境进行了重写, 这就是后来发布的商用 Vivado 设计套件。

由于 SDNet 支持各种协议、硬件实现细节和性能扩展, 因此其在业内具有独特的灵活性。学术界、网络 OEM 厂商、NPU 厂商和 EDA 厂商层多次试图去定义包描述语言、语法分析语言和高级工具, 以实现一些用于设计软网络元素的功能。但他们都没能提供接近 SDNet 特性集的工具来。

软网络元素的变化特性

在 OpenFlow 兴起之前, 路由器就是路由器, 以太网交换机就是以太网交换机。在网络元素明确定义的年代, ASIC 往往是数据包汇聚和流量管理等特定任务的最佳硬件。而今天, OEM 厂商非常自豪地宣称其网络元素是软的, 硬编码的 ASIC 正在快速消失, 取而代之的是 FPGA、网络处理器和多核处理器的组合体。用于对处理器等器件进行编程的软件工具本身必须具有多样性, 因为其作用对每个不同的芯片和软件厂商来说是不同的, 而且随着时间的推移在可编程性本身性质的变化同时也会发生变化。

在上个十年中期, 学术界一些早期研究项目试图用面向对象的置标风格语法来描述数据包。NetPDL 和 Packet Details Markup Language (数据包详细信息置标语言)

等许多项目只是尝试抽象 OSI 协议栈一至七层的协议行为，并未涉及硬件规范实例化。

Click Modular Router (Click 软件路由器)

首批真正开放式规范标准之一出自麻省理工学院的一个称之为 Click Modular Router (以下简称 Click) 的项目。Click 系统旨在辅助路由器配置并在定义路由器架构情况下对线程进行安排，同时也有助于配置语法分析，能扩展到一般数据包语法分析。赛灵思早在 2003 年就采用了 Click 的一些元素，并在 2004 年 DAC 大会上发表了关于所得硬件的论文。Click 系统由美国加州大学洛杉矶分校(UCLA)和初创企业 Mazu Networks (该公司于 2009 年被 Riverbed 收购)进一步开发，但随后几年该系统并未得到广泛推广。

SDNet 和硬件直联

赛灵思对 SDNet 制定了宏伟目标，在开发过程中就与 Vivado 环境紧密结合。SDNet 最初旨在通过软件描述生成定制硬件组件，但 SDNet 能在包处理模块内生成为可编程元件的固件，由此推动该目标的实现。SDNet 还能集成第三方组件。在所有情况下，要求环境都能执行这些任务，同时又受限于具体的工艺技术或总线速度。

赛灵思开发人员希望规范软件能够执行无中断升级，这样实时包传输时就能在数据包之间进行底层硬件修改，而且不会中断线路服务。软件还必须能够生成调试和验证测试平台。

在创建这种高层次规范环境过程中，赛灵思希望确保设计人员能够输入底层硬件要实现的特定速度、参数和行为，并独立于高级功能描述，即便设计人员对底层 FPGA 架构一无所知也没有关系。就此而言，与 Vivado 一同开发的纵向领域的软件规范环境能为设计人员提供 Vivado 工具的功能，而这些设计人员对通信和图像处理等纵向应用领域更在行，而不是精于 VHDL 级 FPGA 设计的细节。在创建高层次规范环境（作为芯片设计流程前端）过程中，赛灵思为熟悉软件工程设计方法的系统专家提供了将软件技能应用于硬件优化的方法。

在 2012 年 Vivado 发布之前，赛灵思的集成软件环境 (ISE) 采用了从不同初创企业收购的综合以及布局布线等工具。设计环境的重新架构设计目的在于提供统一数据模型和基于 Tcl 的脚本语言，以满足 FPGA 每一阶段的设计要求。统一数据模型允许在设计过程的不同阶段同时捕获并交叉探询 FPGA 设计的各个部分。综合工具能提供 RTL 级仿真模型或基于 C 语言的算法 IP 核。赛灵思近期采用的 OpenCL 语言将该功能扩展到了并行多核设计。

SDNet 和 OpenFlow

赛灵思 ISE 架构经过重新架构演进为 Vivado 的同时，首批对软件定义网络感兴趣的路由器和交换机生产商启动了 OpenFlow 学术项目。开发 OpenFlow 协议的目的就是将指导数据包交换机行为的控制层元素与转发层硬件交换机本身分开，转发层硬件交换机通常位于服务器相关的网络域控制器中。一个控制器可统一查看所有转发节点，也能推送流程表项目记录关联到这些节点，而节点则对给定流程采取特定的行为。首版 OpenFlow 是由加州大学伯克利分校和斯坦福大学联合开发的，其源自 2006 年斯坦福大学的 Ethane 项目。2011 年，开放网络基金会（ONF）成立，在该基金会的支持下，发布了 OpenFlow 1.2 版。ONF 的创办方主要是数据中心领域的领先企业，包括 Google、Facebook 和 Microsoft 等，不过现在 ONF 的成员也包括思科、Juniper、F5 Networks、Huawei、NEC、IBM、Dell 和 HP 等企业。

对数据中心管理人员来说，为硬件交换机打开转发层接口意味着网络设备采购方不用再受制于封闭式基础架构。路由器集群和纵向层级可即时重配置，随着流量特性的变法而改变拓扑和包转发速率。

对网络设备基于规范的设计环境用户来说，向 OpenFlow 及其它 SDN 控制协议转型的大趋势是 SDNet 理想的试验田。硬件可根据 FPGA 实现进行重配置，甚至能在固件升级进行芯片实例化之后修改。SDNet 在优化控制层和转发层方面的作用只是其网络功能的一个方面而已。

由于 SDNet 能通过固件实现可编程元件的无中断升级，同时还可优化 TCAM 等固定功能元素的吞吐量、表大小和端口配置等，特定的网络节点能实现功能升级且不会中断网络运行。赛灵思看到超越传统 SDN 之外的更广阔的环境，其中 SDNet 用于向数据路径添加智能功能。图 1 给出了两种方法的对比。

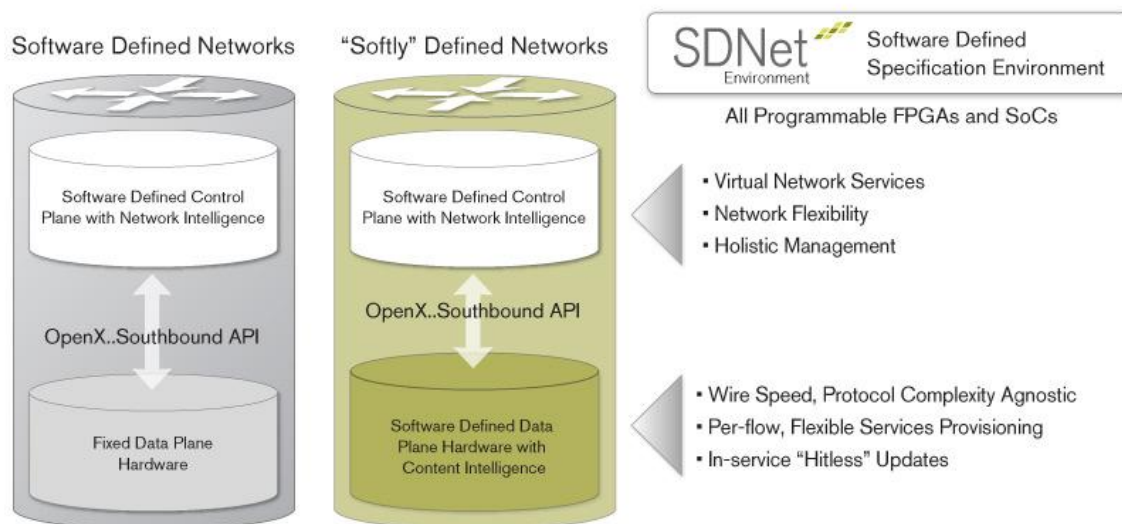


图 1: 在可重配置的数据层中, 更多内容智能可实现更全面的预配置任务和无中断升级。

竞争格局: 苹果和橙子

几乎没有什么工具能执行与 SDNet 内在相同的网络元素配置任务, 这点没什么好吃惊的。ASIC 和 FPGA 设计社区以及通用 EDA 企业对软网络工具没进行过什么研发, 因此在工具套件方面拿不出任何相提并论的东西。NPU 和控制层 CPU 厂商迫切希望提供的工具能够通过使用高级语言协助编程, 但他们在工具扩展以实现多核实用程序优化, 进而提高整体吞吐量、进一步减少 NPU 插座方面没什么动力。同样, 路由器和交换机领域的一些系统 OEM 厂商也支持 OpenFlow 以简化自身 ASIC 设备的操作, 但并没有扩展到网络元素底层硬件修改的领域。

商用芯片方案

10 多年来最高性能包转发引擎供应商之一 EZchip Technologies 公司正在推进传统数据层 NP 系列向新型 NPS 转型, 进而转向运行 Linux 并用 C 语言编程的多线程 CPU。传统上, EZchip 将其 Microcode Development Toolset 视为类似于控制层 CPU 工具的产品, 也提供仿真器、汇编器和子例程库。用于帧生成和数据结构自动生成的其它工具则对数据路径引擎更常见。未来基于 C 语言的 NPS 工具集可能看起来很类似。EZchip 架构通过使用应用程序库进行预先特性描述, 可通过标签交换路由(Label Switched Routing)、二层交换(L2 Switching)、VPLS、防火墙(Firewall)和访问控制列表(Access Control List)等预定义任务简化 NP 功能模块编程。就此而言, 这些库与 FPGA 软 IP 库元素没什么太大差异。

Marvell 为其 Xelerated 处理器提供两款全面的数据层软件套件。Metro Ethernet Application 和 Unified Fiber Access Application 这两款软件包含运行在 NPU 上的应用程序包和运行在主机 CPU 上的控制层 API。控制层 API 的 API 基于硬件适应层, 包括启动脚本、配置模块, 以及用于从控制层访问转发层的预定义消息。这或许是 OEM 客户向市场投放产品的捷径, 不过只能通过 Marvell 提供的源代码进行修改, 处理器本身使用的是汇编码。

Cavium 选择为其 Octeon 系列提供更通用的软件开发套件 (SDK)。该 SDK 包含 Gnu 工具链、仿真器、用于图形分析的 Cavium 自有的 ViewZilla 以及用于深度包分析的正则表达模式编译器。行业软件(vertical application)工具套件为 SSL、TCP、IPsec 以及类似的常见 L3-5 功能提供 C 例程程序。包括一些数据路径元素的通用控制层 CPU 具有高灵活性, 但远未做到包处理优化。

Freescale 公司在 PowerQUICC 和 QorIQ 系列方面拥有多年的经验, 能够提供丰富的软件工具以优化控制层和数据路径引擎产品组合。不过 Processor Expert、

CodeWarrior 和 VortiQa 系列是相当令人生畏的工具套件，彼此之间不能全面集成，而且不支持所有 SDN 概念。Freescale 的每个 MPU 和 MCU 系列都有独特的 Processor Expert 软件套件，相关性最高的就是为特定任务配置多核处理器的 QorIQ 优化套件。它可链接到 CodeWarrior Development Studio（反映控制层任务的综合编码平台），以实现特定 CPU 架构。构建行业应用的 VortiQa 套件最接近 SDN，而 Freescale 推出了针对 SDN 的专用 VortiQa，其包括 Open Network 指导软件和 Open Network 交换机软件。VortiQa 的 SDN 应用套件和 Freescale 针对 SMB、无线基础设施等的套件一样，都可用于现有通信处理器架构的特性描述。Freescale 的数据路径加速基础架构(Data Path Acceleration Architecture)在不同处理器系列成员中提供许多可配置元素，而 VortiQa 工具则有助于提高协处理器的作用。这是一种全软编译解决方案，类似于赛灵思的 SDNet。

网络设备 OEM 厂商

交换机和路由器生产商一直对 SDN 深感不安，因为 SDN 会将代表其核心业务的成本网络元素实现商品化。不过，思科和 Juniper 等架构领先企业还是认识到 SDN 是大势所趋，正在为他们的客户开放更多编程和配置工具。在某些情况下，他们能帮助客户在中层或后层上分配包转发卡和 I/O 线路卡，但想让他们帮助客户直接实时决定硬件重配置，这些 OEM 还做不到这么好心。

例如，思科通过可扩展网络控制器（XNC）方案，在保留传统控制层的同时，采用 Monitor Manager 混合软件包添加外部控制器，实现向 SDN 平滑过渡，从而为交换机和路由器提供性能增强型配置软件套件。与此同时，其面向核心路由的最复杂的 ASIC 也改进了用户可配置性。在其最新核心路由器从 QuantumFlow 升级到 nPower X1 过程中，思科承诺实现现场无中断软件升级，且不会丢失数据包，同时为实现全面 OpenFlow SDN 提供更好的客户数据路径控制。这意味着客户通过固件对 ASIC 进行编程的能力得以提高，但思科仍未提供 OpenFlow 的全面硬件支持。

Juniper 在 4 年前部署其 Trio 芯片组之后增加了客户配置的灵活性。针对查询、存储器和队列的不同数据路径处理器可在 MX 路由器中一定程度上改变目的，但基本硬件功能不能变。2013 年底，Juniper 在收购 Contrail Systems 的基础上推出了 Contrail SDN 控制器，其可实现企业级数据中心或服务供应商物理网络资源的虚拟化。新型服务供应商套件 Junos Fusion 则能提供更多的资源管理功能，但只是略高于特定路由器或交换机的粒度。

在 2014 年的头几周，竞争对手也陆续跟进，如华为推出其 SoftCOM 虚拟化工具，Alcatel Lucent 则推出了 Cloudband，二者都基于 Session Border Controller 等开放节点的更高可视化。说到底，SDN 重配置涉及网络拓扑而不是底层硬件，只有 OEM 厂商愿意将控制层和转发层间的开放接口保持在服务供应商和数据中心管理人员可接受的粒度水平，才能保护现有平台技术并发挥作用。

EDA 工具和 ASIC 厂商

专用设计工具打造的结构化 ASIC 已被 21 世纪的 ASIC 产业所取代，新的产业更像纯粹的代工厂业务模式，而 EDA 厂商提供的商用工具则成为 OEM 厂商继续生产非 FPGA 定制芯片的主要工具。在此情况下，网络 OEM 生产 ASIC 用的唯一潜在设计方法就来源于 EDA 产业。Cadence、Mentor Graphics、Synopsys 和小型竞争企业的开发人员似乎并未意识到 SDN 的发展趋势，或者对此并不感兴趣。

Synopsys 的片上系统 DesignWare 最接近满足软网络要求，但这更多是出于巧合而不是刻意而为。该公司仅仅是在最近向其 DesignWare for Data Center SoC 包添加了 40G 以太网控制器 IP，尚未探讨用 SDN 实现网络功能虚拟化。Synopsys 和 Cadence 则试图通过第三方 IP 供应商推出通用以太网和 OTN IP，而 Mentor 则致力于系统级设计，而且近年来已经偏离了网络领域。没有任何一家 EDA 公司将 SDN 作为关键目标在未来发布。

这或许反映出基于 ASIC 或 EDA 的定制芯片今后在网络设备领域所发挥作用的不确定性。各大网络 OEM 厂商用定制数据路径开发自己的 NPU，而初创企业则正在从 ASIC 转向 FPGA 和商用 NPU。

FPGA 设计工具

Altera 和 Tabula 这两家公司是赛灵思的竞争对手，他们正在向网络领域推出大型复杂器件，不过这两家公司的设计套件几乎没有显示出要转向 SDN 支持的迹象。Altera 通过添加 OpenCL 开发工具（高级并行处理方法），增强了其 Quartus II 设计环境 ARM 部分的功能。赛灵思和 Altera 均使用 OpenCL。有一点需要记住的是，OpenCL 是通用的，而不是专为网络数据流定制的。迄今为止，Altera Quartus II 环境中的各款工具，包括 Qsys 和 DSP Builder，都能协助设计编译，但并没有将现场无中断网络升级作为总目标。

Tabula 的 Stylus 编译器提供综合和布局功能，可将不同的逻辑层整合到独特的 ABAX2 多路复用框架 Spacetime 架构中，但独特的 FPGA 架构对新手来说编程不是一件容易的事。去年，Tabula 发布了以太网语法分析评估套件，以提升网络 ABAX2 架构的采用率。最新 IP 软核包含调度程序和 400G 桥接器，但所有都基于采用 Stylus 设计的固定的二级以太网功能。

赛灵思 SDnet 反映了未来网络发展趋势

SDNet 环境体现出赛灵思未来针对其它垂直领域的发展计划。总体说来，赛灵思 SDnet 环境会让可编程性的含义发生巨变，重新带回能即时修改的可重配置架构概念。SDNet 功能随着 OpenFlow 模型得到支持而不断发展，而 SDNet 在 OpenFlow SDN 领域之外乃至赛灵思 7 系列和 UltraScale 架构之外的影响力会非常大。SDNet

可能成为一个模型, 使得硬件能够通过固件重编程和修改 IP 核配置针对不同任务进行重新架构设计。图 2 给出了从 SDNet 规范和编译到线卡实例化的典型流程。

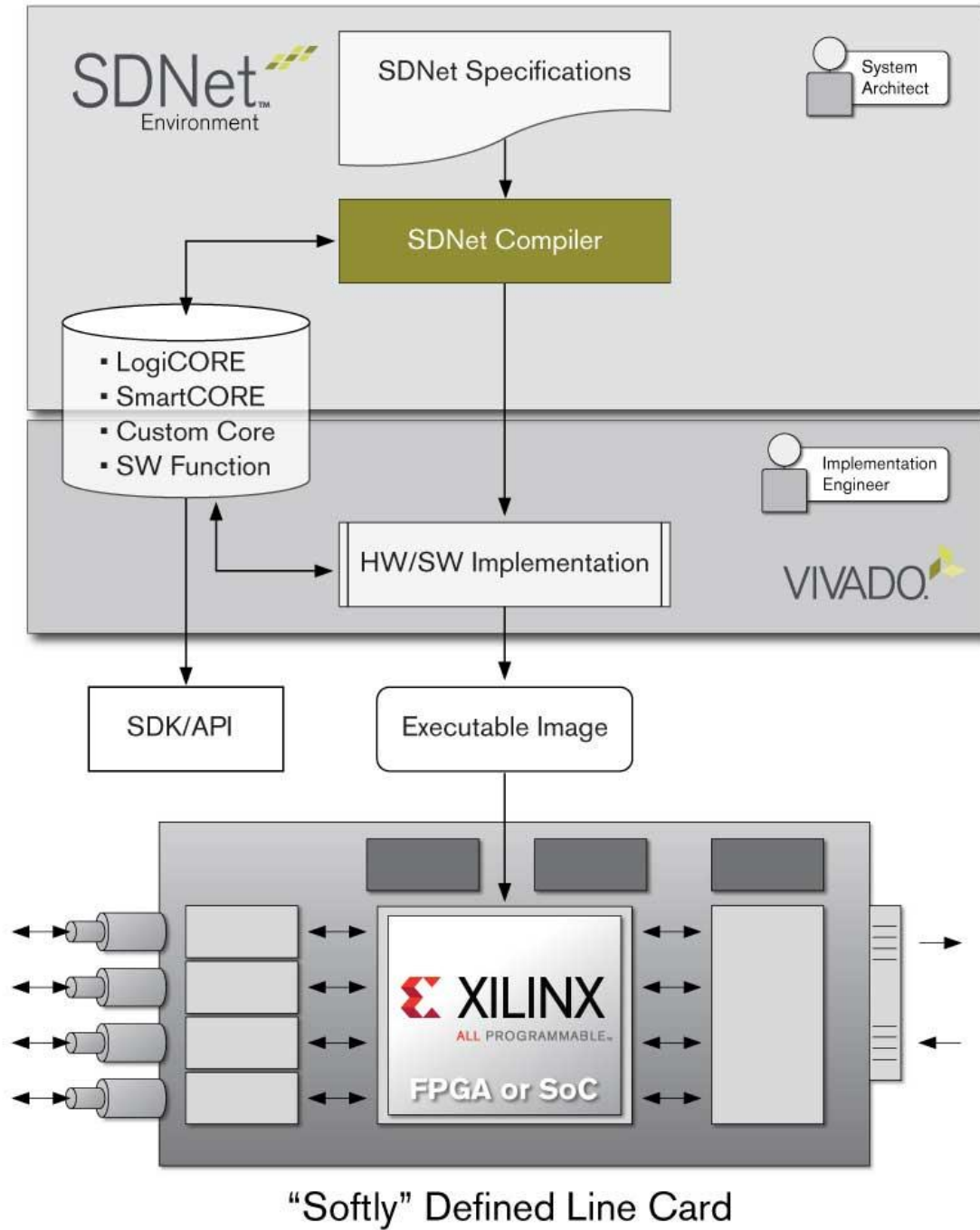


图 2: SDNet 编译创建的可执行镜像能实现在赛灵思 FPGA 或 SoC 中。

SDNet 的专用引擎（如 QoS 策略引擎和搜索引擎）互联层级结构功能提供了一个模型，使得仅了解高级语法的设计人员也能创建出动态可再编程架构，而且能链接数据层引擎和协处理器。赛灵思采用的设计流程从功能规范到编译再到 VHDL，显然是业界通用的。其他开发商可能也会尝试模仿赛灵思的规范方法。赛灵思已暗示可能将 SDNet 某些元素向标准化组织开放，而且会采取与特定赛灵思 7 系列和 UltraScale FPGA 实例化不相关的形式。

目前，SDNet 没有直接竞争对手，因为 SDNet 完全改变了硬件实现 SDN 功能的方法，可对特定的任务进行重配置。显然，它在 OpenFlow 网络中会发挥作用，不过赛灵思定义的工具提供的功能远远超出了通常所认为的 SDN 领域。◆

关于作者

Loring Wirbel 是研究机构 Linley Group 的资深分析师。Linley Group 可针对网络芯片产业提供最全面的分析。我们不仅分析商业战略，也分析所有已宣布推出的产品的内部技术。我们的周刊和深度报告涵盖以太网芯片、网络处理器、多核嵌入式处理器以及无线基站处理器等专题。如需了解更多信息，敬请访问网站：www.linleygroup.com。

本文涉及的商标名称采用文章编辑风格，未加商标符号。这些商标均为其各自所有者的财产。

本文得到赛灵思支持，但所有观点和分析均属于作者本人。