

博士学位论文

**网络数据平面可编程硬件的研究**

学位申请人：乔思祎

指导教师：邹建华教授

合作导师：管哓宏教授

学科名称：控制工程与科学

2020年06月（注：X均为阿拉伯数字，如，2018年06月）

**XXXXXXXXXX XXXXXXXX**（注：此处为论文题目的英文翻译，注意单词的大小写规律，如Structural Design and Rapid Development of Labyrinth Drip Irrigation Emitters）

A dissertation submitted to

Xi’an Jiaotong University

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

By

XXX（注：此处为作者姓名全拼，名在前，姓在后，如Zhengying Wei）

Supervisor: (Associate) Prof. XXX（注：此处为导师姓名全拼，名在前，姓在后，如Bingheng Lu）

Associate Supervisor: Prof. XXX（注：此处为导师姓名全拼，名在前，姓在后）

XXXXXXX （注：此处为学科名称，如Mechanical Engineering）

XXX XXXX（注：此处为英文日期，月在前，年在后，如September 2018）

**博士学位论文答辩委员会**

**XXXXXXXXXXXXXXXXXXXX（论文题目不能超过35个汉字）**

答辩人：XXX

答辩委员会委员：

XXXXXXXXXX大学XXX: \_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_（注：主席）

XXXXXXXXXX大学XXX:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

XXXXXXXXXX大学XXX:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

XXXXXXXXXX大学XXX:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

XXXXXXXXXX大学XXX:\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_

答辩时间：XXXX年XX月XX日

答辩地点：XXXXXX

摘 要

论文摘要由摘要正文、关键词、论文类型、资助申明等部分组成。

博士学位论文摘要正文为1000字(word)左右。

如果论文的主体工作得到了有关基金资助，应在摘要第一页的页脚处标注：本研究得到某某基金（编号：）资助。（五号）

=======================================================

网络通信是支撑当今社会的重要基础设施，当前的发展方向主要集中于建设高性能，高可创新性的网络环境。最近10年，软件定义网络(SDN)和可编程网络(SDN2.0)概念的提出很好的解决了过去网络创新性差、创新难的不足。但随着流量和网络复杂度快速增长，新的网络体系结构也带来了性能和可扩展性两方面的挑战。性能和功能方面:基于CPU 的转发平台性能发展逐步减慢，基于ASIC的智能网卡硬件可编程性差。可扩展性方面:数据平面和控制平面分离的SDN网络架构带来了稳定性不足和效率低的问题。

将问题从网络的三个维度进行分析：

主机侧网络，在服务器网卡层面，基于CPU的智能能网卡的性能难以满足目前虚拟化技术和网络监管细粒度化的发展需求。

交换侧网络，在核心网骨干网层面，基于ASIC的转发平面不足以提供网络网络处理的高灵活性，由于其与成本、性能之间平衡困难，网络工程师的创新空间受到了限制。

流表可扩展性研究，硬件流表是一种高效且昂贵的实现网络转发抽象的核心部件，在软件定义网络时代流表稀缺性更加突出。同时，由于流数目和流量的快速增长，控制平面针对流表的操作导致数据平面和控制平面的大量协议开销，导致网络鲁棒性差，易形成安全隐患。

近年来，现场可编程门阵列（FPGA）器件快速发展，以可编程硬件技术为首的异构架构已经大量融合到网络领域，带来高用户可定制能力的同时也能保证处理性能。

本文主要探索基于可编程硬件的高性能SDN系统，论文从理论抽象分析中提出了体系架构，最后给出了系统实现。本文研究在软件定义网络编程语言中如何将这种可编程硬件抽象层融入整体开发系统，并设计与其配套的控制平面软件和协议，使整体网络系统的软硬件有机结合，增强网络处理性能、灵活性的同时保证安全性。本文将从以下三方面阐述：

1）研究可编程设备加速主机侧网络方法。本文提出利用基于FPGA的智能网卡卸载操作系统层部分网络功能，以达到扩展网络接入层的性能的目的。探讨了不同场景下网络功能的构成，分析并提出一种基于可编程硬件的网络功能定义抽象（Data-Computing，DC抽象）。本文把服务器网络功能任务中可转化为DC抽象的计算密集型功能通过合理转换下放到网卡的FPGA可编程器件中。论文基于可编程网卡设计了一套网络流量捕获，统计分析和回放系统。在满足网络功能不受改变的前提下，证明利用基于FPGA的智能网卡能有效地提升服务器的网络性能、时延和效率（给出真实数据）。

2）研究可编程设备加速网络硬件交换层方法。本文提出一种硬件异构型的可编程网络数据平面架构，将FPGA与ASIC交换芯片有机结合，以增强ASIC报文处理报文的灵活性，同时满足性能需求。论文设计了ASIC面向硬件可编程扩展的接口，将数据包头拆分并通过高速数据互联载体发送给FPGA，利用FPGA可重配特性实现完全可编程的报文处理数据平面；同时，本文基于DC抽象，将网络随路计算（network-centric computing）模式引入可编程网络体系架构；本文通过分析流量模型在FPGA中设计了一种并行化处理单元，在资源消耗可控的前提下大规模提高系统的可扩展性能；另外本文提出了一套基于可编程硬件混合网络架构的软件定义语言编程框架，实现了软件定义需求和可编程硬抽象层分离，以及针对底层数据平面的一种高效自适应的并行单元流分配算法，可以稳定实时地保障系统交换层的高性能。

3）SDN硬件流表可扩展性研究。本文针对不同层面网络设备的控制，进行全局优化、分布式优化。在可编程网卡和交换机组成的网络系统中，数据平面内最重要的资源是流表资源（瓶颈资源），本文从全局视野角度，结合可编程硬件的特性，在全网约束的条件下，对流表资源进行优化，以满足未来可扩展性需求。本文分析不同的流量规模和特征，以及系统多模块直接独特的互联协议，提出一种SDN网络流表空间全局共享机制。实现了在流量大规模扩展的情形下，保证数据平面稳定性，降低系统中关键通信通道失效风险。

此外，为支持本文提出的相关设计概念，本文实现了一套基于FPGA的转发平面设备，包括智能网卡和交换机原型平台。此套平台资源容量大，外设接口丰富，可以满足本文在各类网络架构下实验验证需求。

……

**关 键 词**：网络数据平面；可编程硬件；现场可编程门阵列；XXX；XXX

关键词由3～5个组成。关键词应从《汉语主题词表》中摘选，当《汉语主题词表》的词不足以反映主题时，可由申请人设计关键词，但须加注。每一关键词之间用分号分键词，须在该关键词的右上角标注\*，并在该页的页脚处注明“\*表示非汉语主题词”。

**论文类型**：XXXX

论文类型包括：a.理论研究（Theoretical Research)；b.应用基础(Application Fundamentals)；c.应用研究(Application Research)；d.研究报告(Research Report)；e.设计报告(Design Report)；f.案例分析(Case Study)；g.调研报告(Investigation Report)；h.产品研发(Product Development)；i.工程设计(Engineering Design)；j.工程/项目管理(Engineering/Project Management)；k.其它（Others）。

ABSTRACT

英文摘要撰写要求如下：

（1）用词准确，符合语法；

（2）关键词按相应专业的标准术语写出，尽量从《英语主题词表》中摘选；

（3）如果论文的主体工作得到了有关基金资助，应用英文在摘要第一页的页脚处标注：本研究得到某某基金（编号：）资助；

中文摘要和英文摘要均不要求学位申请人及其指导教师签字。

摘要正文每段开头不空格，每段之间空一行；

The key parts in drip irrigation facilities are emitters. The structural design parameters of emitters can directly affect its performance and the function of the whole drip irrigation system ……

1. Because……

2. Only ……

3. To support ……

**KEY WORDS**: XXX; XXX; XXX; XXX

每个关键词组的第一个字母大写，其余为小写，每一关键词之间用分号分开，最后一个关键词后不打标点符号。例如：Drip irrigation emitter; RP&M; Hydraulics; Labyrinth flow channel

**TYPE OF DISSERTATION**: XXXXXXX

须与中文摘要中的论文类型一致；每个单词第一个字母大写，其余为小写。例如：Applied Research

论文类型包括：a.理论研究（Theoretical Research)；b.应用基础(Application Fundamentals)；c.应用研究(Application Research)；d.研究报告(Research Report)；e.设计报告(Design Report)；f.案例分析(Case Study)；g.调研报告(Investigation Report)；h.产品研发(Product Development)；i.工程设计(Engineering Design)；j.工程/项目管理(Engineering/Project Management)；k.其它（Others）。

# 

目 录

摘 要.............................................................................................................................................................I

ABSTRACT................................................................................................................................................II

[**XXXXXXXXXX XXXXXXXX**（注：此处为论文题目的英文翻译，注意单词的大小写规律，如Structural Design and Rapid Development of Labyrinth Drip Irrigation Emitters） III](#_Toc42935251)

[1 绪论 1](#_Toc42935252)

[1.1 研究的背景 1](#_Toc42935253)

[1.1.1 研究的意义 1](#_Toc42935254)

[1.1.2 技术简介 3](#_Toc42935255)

[1.1.3 国内外应用与研究现状 3](#_Toc42935256)

[1.2 研究内容 4](#_Toc42935257)

[1.3 关键科学问题 4](#_Toc42935258)

[1.4 主要研究成果 4](#_Toc42935259)

[1.5 论文组织结构 4](#_Toc42935260)

[2 数据平面可编程综述 4](#_Toc42935261)

[2.1 标题2 4](#_Toc42935262)

[2.1.1 标题3 4](#_Toc42935263)

[3 XX（标题1） 8](#_Toc42935264)

[3.1 标题2 8](#_Toc42935265)

[3.1.1 标题3 8](#_Toc42935266)

[4 XXX（标题1） 9](#_Toc42935267)

[4.1 标题2 9](#_Toc42935268)

[4.1.1 标题3 9](#_Toc42935269)

[5 XXXX（标题1） 10](#_Toc42935270)

[5.1 标题2 10](#_Toc42935271)

[5.1.1 标题3 10](#_Toc42935272)

[6 XXXXX（标题1） 11](#_Toc42935273)

[6.1 标题2 11](#_Toc42935274)

[6.1.1 标题3 11](#_Toc42935275)

[7 XXXXXX（标题1） 12](#_Toc42935276)

[7.1 标题2 12](#_Toc42935277)

[7.1.1 标题3 12](#_Toc42935278)

[8 XXXXXXX（标题1） 13](#_Toc42935279)

[8.1 标题2 13](#_Toc42935280)

[8.1.1 标题3 13](#_Toc42935281)

[9 XXXXXXXX（标题1） 14](#_Toc42935282)

[9.1 标题2 14](#_Toc42935283)

[9.1.1 标题3 14](#_Toc42935284)

[10 XXXXXXXXX（标题1） 15](#_Toc42935285)

[10.1 标题2 15](#_Toc42935286)

[10.1.1 标题3 15](#_Toc42935287)

[11 XXXXXXXXXX（标题1） 16](#_Toc42935288)

[11.1 标题2 16](#_Toc42935289)

[11.1.1 标题3 16](#_Toc42935290)

[12 XXXXXXXXXXX（标题1） 17](#_Toc42935291)

[12.1 标题2 17](#_Toc42935292)

[12.1.1 标题3 17](#_Toc42935293)

[13 结论与展望 18](#_Toc42935294)

[13.1 标题2 18](#_Toc42935295)

[13.1.1 标题3 18](#_Toc42935296)

[致 谢 19](#_Toc42935297)

[参考文献 20](#_Toc42935298)

[附 录 24](#_Toc42935299)

[攻读学位期间取得的研究成果 25](#_Toc42935300)

声明

CONTENTS

ABSTRACT (Chinese)...................................................................................................................................I

ABSTRACT (English)...................................................................................................................................II

1 Preface X

1.1 Drip Irrigation Technology X

1.1.1 Drip Irrigation Systems X

2 Rapid Development of Labyrinth Drip Irrigation Emitters X

2.1 Structural Design of Labyrinth Drip Irrigation Emitters X

2.1.1 Theory X

2.6 Brief Summary X

12 Conclusions and Suggestions X

Acknowledgements X

References X

Appendices（单个附件用Appendix） X

Achievements X

Declarations

（这里的目录没办法自动生成，因为你没有相应的英文标题样式，只好你自己手工添加了，其实就是把中文目录翻译成英文就可以了）

编辑格式：“章节号＋英文标题＋Tab键1次＋页码”，编完以后，套用“CONTENTS”样式。

主要符号表

|  |  |
| --- | --- |
|  | 灌水器流量偏差系数 |
|  | 管道内径/mm |
|  | 灌水器流道当量直径/mm |
|  | 管长/m |
|  | 迷宫流道单元个数/个 |
|  | 灌水器流量/L·h-1 |
|  | 灌水器额定流量/L·h-1 |
| *Re* | 雷诺数 |
|  | 灌水器流量标准偏差 |
|  | 流体的运动粘性系数 |
|  | 流态指数 |

如果论文中使用了大量的物理量符号、标志、缩略词、专门计量单位、自定义名词和术语等，应将全文中常用的这些符号及意义列出。如果上述符号和缩略词使用数量不多，可以不设专门的主要符号表，但在论文中出现时须加以说明。

论文中主要符号应全部采用法定单位，特别要严格执行GB3100～3102—93有关“量和单位”的规定。单位名称的书写，可以采用国际通用符号，也可以用中文名称，但全文应统一，不得两种混用。

缩略词应列出中英文全称。

主要符号表正文统一左缩进一个字符。

符号表排序方法：先按拉丁字母大写、小写排序，再按希腊字母大写、小写排序，如下表所示：

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** | **10** | **11** | **12** | **13** | **14** | **15** | **16** | **17** | **18** | **19** | **20** | **21** | **22** | **23** | **24** | **25** | **26** |
| **Ⅰ** | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z |
| **Ⅱ** | a | b | c | d | e | f | g | h | i | j | k | l | m | n | o | p | q | r | s | t | u | v | w | x | y | z |
| **Ⅲ** | Α | Β | Γ | Δ | Ε | Ζ | Η | Θ | Ι | Κ | Λ | Μ | Ν | Ξ | Ο | Π | Ρ | Σ | Τ | Υ | Φ | Χ | Ψ | Ω |  |  |
| **Ⅳ** | α | β | γ | δ | ε | ζ | η | θ | ι | κ | λ | μ | ν | ξ | ο | π | ρ | σ | τ | υ | φ | χ | ψ | ω |  |  |

**Ⅰ**：拉丁字母大写；**Ⅱ**：拉丁字母小写；**Ⅲ**：希腊字母大写；**Ⅳ**：希腊字母小写。

本部分内容非强制性要求，如果论文中所用符号不多，可以省略《主要符号表》。

在每一章的末尾插入下一章的MathType的章标记（打印前将其字体颜色变为白色，在打印预览中看不见即可）：

# 绪论

绪论部分主要论述论文的选题意义及应用背景[[1]](#footnote-1)、国内外研究现状分析及论文的主要研究内容等。

绪论一般包括：本研究课题的社会背景和学术背景；对社会进步、经济建设、学术发展的理论价值、现实价值和实践价值；国内外研究现状分析与归纳；论文所要解决的理论问题或现实问题；论文的基本思路、运用的主要理论与方法，以及行文结构与章节的安排等。

## 研究的背景

### 研究的意义

现状，矛盾，引出问题，提出需求，

现状：

民生需求-》互联网服务-》互联网流量的特征-》互联网基础设施特征-》

国家产业升级-》新基建-》拉动作用-》

新趋势-》5G IOT IOV 智慧基础设施-》

矛盾：

网络的演进，各个时期所解决的问题

说明研究可编程硬件在数据平面时的重要价值。

为什么SDN，为什么P4，为什么高性能，为什么灵活性，为什么是可编程硬件

统一的结论:需要有一个数据平面的可编程硬件来进行支撑。

21世纪的新20年，网络正以前所未有的速度越来越紧密地参与到民生社会中，为满足民生需求、新基建拉动内需和国家产业升级起到了至关重要的作用。从“百度一下”到网红全民直播带货，从实现“三网通”到发展“新基建”的国家战略，小到优化社会资源效率的办公数字化，大到勾勒出智能交通、智慧城市和万物互联的5G海洋，无一不是构建在网络基础设施的快速发展之上。思科公司预计，到2023年全球家用互联网总带宽将达到5.85Ebps[[2]](#footnote-2)（是现在的3.27倍），移动互联网用户预计达到57亿，其总流量可达11.3Ebps (将达到目前的5倍)，其中5G流量将占据移动互联网总带宽的76.5%（0.6%，2019年）[1-2]。由于深度学习、AI、大数据、云计算、物联网的快速发展，这些新技术将催使新零售、新金融、新医疗、新教育、新制造、云视频和云游戏等行业“云化”，海量的数据会在数据中心内部服务器间网络中及对外网关中交互，这些关键应用将会改变数据中心算力和数据中心内部网络结构特性。

IDC报告称，2019上半年中国公有云服务整体市场（Iaas/PaaS/SaaS）达到54.2亿美元，并预计在未来5年间内以年均复合46%的速度快速增长[3-4]。数据中心内服务器计算力呈现异构化趋势，GPU, AI Chip, FPGA等使用非通用类型指令集和特殊体系架构计算单元已成为目前分布式计算领域的热点话题。现在超大型数据中心一般可容纳数十万台终端服务器，内部网络链接数量多、拓扑规模大、传送海量数据，这使得现有的网络将变的异常复杂。同时，新的数据包类型层出不穷也使得现有网络变得异常脆弱。

传统网络技术已经无法满足当前的网络环境的需求，最近十年来网络技术和架构经历了快速地演进和变革，针对数据中心网络尤为明显。传统网络的互连包含了经典的二三层网络。为增强交换机的扩展能力，二层网络增加了广播，桥接等复杂功能。这种网络架构在小规模应用时可以展现强大的智能性与可扩展性，但当网络规模进一步增加，网络中容易出现的广播风暴、链路收敛等一系列尖锐问题变得难以解决。现代的大型网络设计思想摒除了略显冗余看似小聪明的功能设计，事先规划好网络拓扑层次，完整地保留网络的第三层，从而将网络扁平化。网络拓扑结构演化出可进行大规模扩展的CLOS型架构，为了降低系统复杂性，在各个层次之间的网络设备功能也逐步变得统一透明。网络设备统一化，可降低网络功能开发部署的难度。通常，研究人员需要持续地投入对网络进行测量、监控、容错、提升效能的工作。由于思想的创新和技术的推进，设备厂商不断开发出具备各种高级功能交换芯片。设备、芯片功能强大的同时，复杂的网络功能不断地对于网络的管理层又提出了新的挑战。

为解决设备制造复杂和设备管理复杂的问题，软件定义网络（Software Defined Network，SDN）概念的提出拨开了笼罩在网络体系结构发展道路上的迷雾。SDN将数据平面和控制平面解耦。在数据平面上，对数据包的处理统一做查找-转发（Match-Action）抽象。控制平面复杂建立网络拓扑，控制并下发流表。这样所有的数据包转发行为都又控制平面的软件逻辑完成，数据平面可以支持任意一种网络协议的处理。由于软件具有强大的灵活性以及开发的敏捷性，SDN大大加速了网络创新和智能化进程。数据平面和控制平面的安全通道由OpenFlow协议进行规范，将数据平面统一化、简单化，使得网络交换设备向白盒化方向发展。

大规模网络无论是在底层设备架构还是运维方式上仍不能停止变革的脚步，这为可编程硬件的发展带来了巨大空间。随着云服务概念和大规模机器学习的落地，近年来以云计算为代表的数据中心网络规模指数增长。网络功能虚拟化在数据中心内部是关键一环。虚拟交换机则是主机内各虚拟机之间数据包转发的核心软件。随着众核CPU架构快速发展，服务器内虚拟机布置资源大幅扩张，促使主机出口吞吐量从40GbE向100GbE甚至400GbE演进。不但如此，复杂的网络安全规则、流量监控等模组进一步导致CPU过多地消耗在处理网络功能上面。研究人员可能需要花费大量时间去解决目前网络架构的大规模扩展的方案。（创新变的异常艰难）。传统x86CPU架构适合于处理灵活多变的计算控制任务，对于做重复、常规流式数据处理，通用指令集架构并不能得到最优的效率，厂商往往不得不依靠大量部署server来解决。为缓解主机内CPU消耗过大，目前提出的智能网卡是一种新思路。智能网卡采用FPGA，网络处理器，ARM等器件，或以他们的组合形式形成在网卡端的新的算力集合，这种算力集合对于处理网络流量会有更高的效率。我们可以把转发动作，网络安全规则等功能下放进来，以削减服务器CPU的额外消耗。ASIC具有最好的性能和最高的能量效率，但每次大批量的部署消耗时间长，投入研发资金大。对于运营商来说，设备、仪器等一次性支出都叫做CapEx（Capital Expenditure，资本性支出）。对于目前快速发展的网络环境架构，设备的更新换代周期也在变短，在优化CapEx时已经不能把固定设备投入当做一次性支出。在探索新一代网络架构时，CapEx也会成为重要的参考因素。

随着创新性和需求的进一步发展，让底层硬件拥有灵活的可控制能力才能满足目前行业变革的需求。因此，网络领域提出了编程协议无关（Programming Protocol-Independent Packet Processors，P4）概念。P4协议不但支持SDN网络控制和管理的可编程性，还提出了数据平面可编程的概念。数据包在数据平面内的处理模型遵循解析-查找-匹配的抽象模式。P4规定了一种编程语言[5]，它可以控制数据平面对数据包的任意解析行为，也可以自由配置查找表的数据位宽和多级流表之间的查找流水线[6-7]。这种更高阶的数据平面可编程模型使交换机设备更加白盒化，交换机与任意网络协议解绑，带来了具备灵活性的创新实践。除此之外，端到端大带宽、低时延的网络需求引申出了网络功能硬件卸载、网络随路计算等概念，这进一步增强了对高性能的网络数据平面可编程性的需求。

综上所述，现代网络在向软件定义、数据平面可编程的方向发展。网络架构的变迁的核心是有一套可以映射上层可编程逻辑的硬件数据平面。本文主要探索一种面向网络数据平面的可编程硬件，能够满足快速迭代的网络创新性需求，同时能够提供与目前主流设备相仿的处理性能，以及可扩展性高的全局优化方法。

### 技术简介

软件定义网络的基本设计概念是将数据平面与控制平面分离[10]。其中，网络数据平面是指完成计算机之间通信数据包的匹配、修改、传送、转发的软硬件设备。数据平面的可编程性要求网络管理员拥有对数据平面的各个特性做快速个性化定制。网络的控制平面维护全网视野数据，配置针对流的转发条目。控制平面中的应用程序几乎都由软件构成。当前数据平面的设计思想如图 1‑1所示，主要有软件方法实现，专用硬件实现和新设计的可编程硬件。



图 1‑1 软件定义网络结构及其实现方案

在不同场景下，网络对于数据平面的需求千差万别，研究人员一般根据场景的流量大小，处理过程复杂度来思考并选取数据平面的实现方案，本文将在第2 章详细介绍各类数据平面的实现方案的优缺点，并着重于可编程性的分析。目前两种最重要的数据平面是**软件交换机**和**专用硬件交换机**。两者在功能上都是针对数据包做一系列处理，包括匹配、查找、统计、传送、转发和安全校验等等，其中**流表**是实现数据平面核心功能的函数（器件）。数据平面、都包含一个可以与远端控制器沟通的软件代理，这部分功能着重于通信协议的实现以及通道安全性加解密，主要由轻量级通用处理器完成。其二者的主要区别在于处理数据包的性能以及交换容量。数据包处理性能主要看数据吞吐量（字节每秒）和包吞吐量（包每秒），目前软件交换机做高性能的包转发几乎可以达到60G/60Mpps[9]。当数据包处理复杂度增加时，软件交换机的性能会直线下降，几乎与操作步骤数成反比。专用硬件交换机有接口数目多，交换容量大的特点，一般能满足64口乘以每口25Gbps的总交换容量。而且硬件交换机的性能与数据包处理步骤几乎无关，它拥有良好的性能稳定性，低转发时延等特性。虽然在核心网络和高性能网关领域主要使用硬件交换机，但是硬件交换机的功能固定，更换成本高昂，如果需要修改网络或者更改网络功能那么选用专用硬件交换机的场景将无从下手。所以目前在数据中心网络或服务器NFV（网络功能虚拟化）等场景中，软件交换机依然占据很大份额。由于软件交换机的灵活性高，开发人员能够快速迭代部署新功能，由于传统单机CPU通信速率需求不高，软件交换机尚能满足在数据处理时延高、吞吐率低的前提下，提供足够的可编程灵活性。但随着人工智能领域、5G的发展，数据中心网络内通信容量需求快速增长，转发时延需求快速收紧，软件交换机性能瓶颈快速到来，将不得不面对大量无谓堆叠CPU的情形。本文将主要研究在主机侧网络和核心交换网络中使用可编程硬件来大大提高交换机的性能瓶颈。针对控制平面，本文将从单点优化开始用**分布式优化和全局优化**的思想，实现对网络中的瓶颈资源（如流表资源）的可扩展性和安全性提升。

### 国内外应用与研究现状

可编程数据平面的研究现状，软件，ASIC，fpga

Asic 阿里巴巴网关 fpga azure

为增强数据平面的可编程性，工业界学术界互相促进、广泛研究并已经提出了许多方案。分析一下李昊论文的情况。

基于软件的数据平面开发便捷，价格低廉，无需在网络中部署专用设备，是功能快速实现的首选方案。目前在虚拟化的云服务系统中，已经部署了大量基于软件的功能：

a）转发层，华为CE1800V[11]是专为数据中心云计算虚拟化环境部署的一种分布式虚拟交换机。其支持标准Open Flow1.3控制协议，以及Open vSwitch 数据库管理协议（OVSDB），基于英特尔DPDK（Data Plane Development Kit）技术提供每核12Gbps的转发吞吐，比业界平均水平高出20%。

b）流量监管，Activelogic[12]是一个提供安全可靠、流量分类、提高QoE（Quality of Experience）能力的网络管理工具。它基于软件可自动化部署，依靠超大规模性能、人工智能技术以及云计算场景优化的能力，在数据平面解决流量监管的问题。

基于软件的数据平面功能可以依靠堆叠CPU核数来实现大规模的性能扩展，但由于计算复杂度过高、基于指令的图灵机处理效率低下，即使做简单转发功能想达到100Gbps线速也需要占用8个核心以上[9,14]。单纯地依靠软件处理器扩张来增加网络性能，边界收益将越来越小。

基于白盒交换机和P4专用芯片的数据平面在网络性能方面大幅超越基于通用服务器的NFV数据平面。符合SDN架构的白盒交换机将控制平面移交给远端软件层，从而大大提升设备的再开发能力，在某些基础网络转发设备的智能化和可定制化方面给出了比较好的灵活性。数据平面可编程芯片提供了硬件层面上的可编程包头抽取器、可编程流表以及可编程执行器，他们的设计思想是依靠快速查表（TCAM，SRAM）法，或后期编程选取特定的冗余模块（在ASIC芯片内部的空间上堆叠的可编程单元）法，来完成专用电路（ASIC）的直接描述逻辑。

参考文献

[1] 华为2019年报，https://www-file.huawei.com/-/media/corporate/pdf/annual-report

/annual\_report\_2019\_cn.pdf?la=zh

[2] Cisco Annual Internet Report (2018–2023) White Paper https://www.cisco.com/c/en/us/solutions/collateral/executive-perspectives/annual-internet-report/white-paper-c11-741490.html

[3] 《中国公有云服务市场（2019上半年）跟踪》

[4] 中国公有云服务市场 2019Q3 跟踪

[5] p4

[6] MRT

[7] 软件交换机

[8] From Ethane to SDN and Beyond (SDN介绍)

[9]pisces

[10] OpenFlow: Enabling Innovation in Campus Networks

[11] https://carrier.huawei.com/~/media/CNBG/Downloads/Product/Fixed%

20Network/b2b/0920/1800-en.pdf 2018

[12] https://www.sandvine.com/download-ready?dl=2020/Datasheets

/Network%20Optimization/Sandvine\_DS\_ActiveLogic.pdf

[13] The Design and Implementation of Open vSwitch

## 研究内容

## 关键科学问题

## 主要研究成果

统计里介绍存储看《阿里巴巴》240页的内存类型。

## 论文组织结构

# 数据平面可编程综述

## 标题2

### 标题3

#### 标题4

##### 标题5

###### 标题6

###### 标题6

标题7

图、表、公式等一律用阿拉伯数字分章连续编号，如 图1-3、表2-1、（3-2）等。图、表、公式等与正文之间间隔0.5行。

图应有图题，表应有表题，并分别置于图号和表号之后，图号和图题应置于图下方的居中位置，表号和表题应置于表上方的居中位置。引用图或表应在图题或表题右上角标出文献来源。

若图或表中有附注，采用英文小写字母顺序编号，附注写在图或表的下方。

**图：**

（1）插图须紧跟文述。在正文中，一般应先见图号及图的内容后再见图，一般情况下不能提前见图，特殊情况须延后的插图不应跨节；

（2）提供照片应大小适宜，主题明确，层次清楚，金相照片一定要有比例尺；

（3）图应具有“自明性”，即只看图、图题和图例，不阅读正文，就可理解图意。

通常使用的函数图采用简化形式，称为简写函数图，例如图 1‑1。

图中的标目是说明坐标轴物理意义的项目，它是由物理量的符号或名称和相应的单位组成。物理量的符号由斜体字母标注，单位的符号使用正体字母标注，量与单位间用斜线隔开。例如：*I*/A，*ρ*/kg·m-3 ，*F*/N，*υ*/m·s-1 等等。

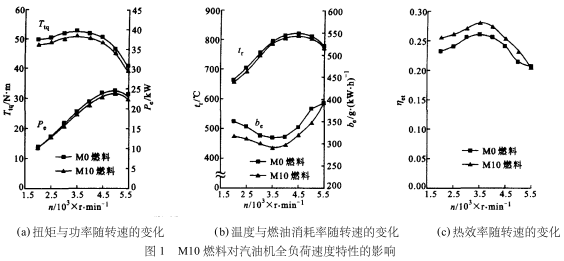
（4）图中用字为五号，如排列过密，用五号字有困难时，可小于五号字，但不得小于七号字。



图 2‑1 2005年相对2001年，5所大学SCI-e文献总数增幅图

（5）图的大小一般为宽6.67 cm×高5.00cm。特殊情况下，也可宽9.00 cm×高6.75cm，或宽13.5 cm×高9.00cm。总之，一篇论文中，同类图片的大小应该一致，编排美观、整齐。

（6）一幅图如有若干幅分图，均应编分图号，用(a)，(b)，(c),...... 按顺序编排，且各分图的分题注直接列在各自分图的正下方，总题注列在所有分图的下方正中，如下图所示：



**表：**

（1） 如某个表需要转页接排，在随后的各页上应重复表的编号。编号后跟表题（可省略）和“（续）”，如表1（续），续表均应重复表头和关于单位的陈述。

表格的设计应紧跟文述。表的编排一般是内容和测试项目由左至右横读，数据依序竖读，应有自明性。若为大表或作为工具使用的表格，可作为附表在附录中给出，论文中的表格参数应标明量和单位的符号；

（2）表中各物理量及量纲均按国际标准(SI) 及国家规定的法定符号和法定计量单位标注；

（3）一律使用三线表，与文字齐宽，顶线和底线线粗磅，中线线粗1磅。例如表1-1；

（4）使用他人表格须注明出处。

（5）表中用字为五号字体。如排列过密，用五号字有困难时，可小于五号字，但不小于七号。

（6）表格必须通栏，即表格宽度与正文版面平齐，如下表所示。

表‑1文献类型和标志代码

|  |  |  |  |
| --- | --- | --- | --- |
| 文献类型 | 标志代码 | 文献类型 | 标志代码 |
| 普通图书 | M | 会议录 | C |
| 汇编 | G | 报纸 | N |
| 期刊 | J | 学位论文 | D |
| 报告 | R | 标准 | S |
| 专利 | P | 数据库 | DB |
| 计算机程序 | CP | 电子公告 | EB |

在三线表中可以加辅助线，以适应较复杂表格的需要，如表1‑2所示。

表‑2方弯管内流动最大速度比较

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 项目 | 层流 | | 紊流 | |
| 0°截面 | 90°截面 | 0°截面 | 90°截面 |
| 理论值*Vmax*/m·s-1 | 0.04 | 0.03 | 1.30 | 1.25 |
| 计算值*Vmax*/m·s-1 | 0.04 | 0.03 | 1.26 | 1.21 |
| 误差/% | 0.00 | 3.12 | 3.07 | 3.20 |

**公式：**

（1）公式应另起一行，居中编排，较长的公式尽可能在等号后换行，或者在“+”、“-”等符号后换行。公式中分数线的横线，长短要分清，主要的横线应与等号取平。

（2）公式后应注明编号，公式号应置于小括号中，如公式（2-3）。写在右边行末，中间不加虚线；

（3）公式下面的“式中：”两字左起顶格编排，后接符号及其解释；解释顺序为先左后右，先上后下；解释与解释之间用“；”隔开。

（4）公式中各物理量及量纲均按国际标准（SI）及国家规定的法定符号和法定计量单位标注，禁止使用已废弃的符号和计量单位。

范例：



式中：*q* —— 灌水器流量/L·h-1；*kd* —— 流量系数；*H* —— 工作压力/ｍ；*x* —— 流态指数。

（此处，“式中：”改为顶格输出）

中，（公式引用）………………



# XX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 3‑1 XXXXXX

# XXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 4‑1 XXXXXX

# XXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 5‑1 XXXXXX

# XXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 6‑1 XXXXXX

# XXXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 7‑1 XXXXXX

# XXXXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 8‑1 XXXXXX

# XXXXXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 9‑1 XXXXXX

# XXXXXXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 10‑1 XXXXXX

# XXXXXXXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 11‑1 XXXXXX

# XXXXXXXXXXX（标题1）

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 12‑1 XXXXXX

# 结论与展望

结论部分着重总结出论文的创新点或新见解及研究展望或建议。

## 标题2

### 标题3

公式按章重新编号：



公式说明，…………（公式在正文中的引用）

图题注：

图 13‑1 XXXXXX

# 致 谢

致谢中主要感谢导师和对论文工作有直接贡献和帮助的人士和单位。

一般致谢的内容有：

（一）对指导或协助指导完成论文的导师；

（二）对国家科学基金、资助研究工作的奖学金基金、合同单位、资助或支持的企业、组织或个人；

（三）对协助完成研究工作和提供便利条件的组织或个人；

（四）对在研究工作中提出建议和提供帮助的人；

（五）对给予转载和引用权的资料、图片、文献、研究思想和设想的所有者；

（六）对其他应感谢的组织和个人。

致谢言语应谦虚诚恳，实事求是。字数不超过1000汉字

用于双盲评审的论文，此页内容全部隐去。

# 参考文献

（此上两空行不能删除，是为EndNote的参考文献列表所预留）

文后著录的参考文献务必实事求是。论文中引用过的文献必须著录，未引用的文献不得出现。应遵循学术道德规范，避免涉嫌抄袭、剽窃等学术不端行为。

参考文献一般应是作者亲自考察过的对学位论文有参考价值的文献，除特殊情况外，一般不应间接引用。

参考文献应有权威性，要注意引用最新的文献。

参考文献的数量：

博士学位论文，一般应在80篇以上，其中，期刊文献60篇以上，国外文献30篇以上，均以近5年的文献为主。

对于申请专业学位的学位论文，参考文献的数量可参照执行。

参考文献的著录格式应符合国家标准GB/T 7714-2015《文后参考文献著录规则》。参考文献中每条项目应齐全。

1. **顺序编码制**

文献中的作者不超过三位时全部列出，超过三位时，一般只列前三位，中文的后面加 “等”字，英文的后面加 “et al”，作者姓名之间用逗号分开。

外国人名一般采用姓在前，名在后的著录法，姓全写且第一个字母大写，名简写成单个大写字母且不加标点，姓和名之间空1格，如：“MetcalfSW”。也可采用名在前，姓在后的著录法，姓全写且第一个字母大写，名简写成单个大写字母且不加标点，名和姓之间空1格，如：“SW Metcalf”。

中文人名的英文表达方式：

简写时，采用姓在前，名在后的著录法，姓全写且第一个字母大写，名简写成单个大写字母且不加标点，如，“钱学森”，简写为“Qian XS”。

全拼时，名在前，姓在后的著录法，名的第一个字母大写，名连写，名后空1格写姓，姓的第一个字母大写。如，“钱学森”，写为“XuesenQian”。

文后参考文献著录格式范例样板，采用五号。

具体要求如下：

A 专著（包括普通图书［M］、论文集和会议录［C］、科技报告［R］、学位论文［D］、标准［S］）

主要责任者．文献题名［文献类型标志］．其他责任者．版本项(第１版不标注) ．出版地：出版者，出版年：引文页码．获取和访问路径．

B 专著中的析出文献

析出文献主要责任者．析出文献题名[文献类型标志]．析出文献其他责任者//专著主要责任者．专著题名：其他题名信息. 版本项(第１版不标注) ．出版地：出版者，出版年：析出文献的起止页码．获取和访问路径．

C连续出版物

主要责任者．题名:其他题名信息［文献类型标志］．年，卷（期）－年，卷（期）.出版地：出版者，出版年．获取和访问路径．

D连续出版物中的析出文献（包括期刊中析出的文献[J]、报纸中析出的文献[N].）

析出文献主要责任者．析出文献题名［文献类型标志］．连续出版物题名：其他题名信息，年，卷（期）：页码．获取和访问路径．

E专利文献

专利发明者/专利申请者或所有者．专利题名: 专利国别,专利号［文献类型标志］.公告日期或公开日期. 获取和访问路径．

F电子文献（包括专著或连续出版物中析出的电子文献）

主要责任者．题名：其他题名信息[文献类型标志/载体类型标志]．出版地：出版者，出版年（更新或修改日期）［引用日期］．获取和访问路径．

表2-2 文献类型和标志代码

|  |  |  |  |
| --- | --- | --- | --- |
| 文献类型 | 标志代码 | 文献类型 | 标志代码 |
| 普通图书 | M | 会议录 | C |
| 汇编 | G | 报纸 | N |
| 期刊 | J | 学位论文 | D |
| 报告 | R | 标准 | S |
| 专利 | P | 数据库 | DB |
| 计算机程序 | CP | 电-子公告 | EB |

表2-3 电子文献载体和标志代码

|  |  |  |  |
| --- | --- | --- | --- |
| 载体类型 | 标志代码 | 载体类型 | 标志代码 |
| 磁带（magnetic tape） | MT | 磁盘（disk） | DK |
| 光盘（CD-ROM） | CD | 联机网络（online） | OL |

样例：

1. 刘国钧，郑如斯．中国书的故事［M］．北京：中国青年出版社，1979：110-115．
2. 昂温 G．外国出版史［M］．陈生铮译．北京：中国书籍出版社，1988．
3. 辛希孟．信息技术与信息服务国际研讨会论文集：A集［C］．北京：中国社会科学出版社，1979．
4. 冯西桥．核反应堆压力容器的LBB分析［R］．北京：核能技术设计研究院，1997．
5. 张和生．地质力学系统理论［D］．太原：太原理工大学，1998．
6. 全国文献工作标准化技术委员会第七分委员会．GB/T 5795-1986．中国标准书号［S］．北京：中国标准出版社，1986．
7. 罗云．安全科学理论体系的发展及趋势探讨［M］//白春华，何学秋，吴宗之．21世纪安全科学与技术的发展趋势．北京：科学出版社，2000：1-5．
8. 钟文发．非线性规划在可燃毒物配置中的应用［C］//赵玮．运筹学的理论与应用：中国运筹学会第五届大会论文集．西安：西安电子科技大学出版社，1996：468－471．
9. 高义民，张凤华，邢建东等．颗粒增强不锈钢基复合材料冲蚀磨损性能研究［J］．西安交通大学学报，2001，35(7)：727-730．
10. Papworth A, Fox P, Zeng GT, et al. Ability of aluminum alloy to wet alumina fibres by addition of bismuth[J]. Mater Sci & Technol,1999,15(4):419-428.
11. 丁文祥．数字革命与竞争国际化［N］．中国青年报，2000－11－20(15)．
12. 姜锡洲．3一种温热外敷药制备方案：中国，881056078［P］．1989-07-26．
13. Koseki A,Momose H,Kawahito M,et alComplier:US,828402［P/OL］2002-05-25 [2002-05-28].http://FF&p.
14. Online Computer Library Center, Inc. History of OCLC[EB/OL].[2000-01-08]. http://www. clc.org/ about/history/default.htm.
15. 江向东．互联网环境下的信息处理与图书管理系统解决方案［J/OL］．情报学报，1999，18(2)：4[2000-01-18]．http://www.chinainfo.gov.cn/periodical/qbxb．
16. Scitor C. Project scheduler[CP/DK].Sunnyvale,Calif.:Scitor Corp, 1983.
17. Metcalf SW. The Tort Hall air emission study[C/OL]//The International Congress on Hazardous Waste, MarquisHotel, Atlanta,Georgia,June 5-8,1995: impact on human and ecological health[1998-09-22]. <http://atsdrl>.atsdr.cdc.gov:8080/cong95. html.

**2. 著者-出版年制**

1. 正文引用的文献采用著者-出版年制时，各篇文献的标注内容由著者姓氏与出版年构成，并置于“（ ）”内，倘若只标注著者姓氏无法识别该人名时，可标著者姓名，例如中国人、韩国人、日本人用汉字书写姓名。集体著者著述的文献可标注机关团体名称。倘若正文中已提及著者姓名，则在其后的“（ ）”内只著录出版年。
2. 正文中引用多著者文献时，对欧美著者只需标注第一个著者的姓，其后附“et al.”“等”之间留适当空隙。

C. 在参考文献表中著录同一著者在同一年出版的多篇文献时，出版年后应用小写字母a, b, c ...区别。

D. 多次引用同一著者的同一文献，在正文中标注著者与出版年，并在“（ ）”外以角标的形式著录引文页码。

样例：

BAKER S K, JACKSON M E. 1995. The future of resource sharing [M].

New York: The Haworth Press.

尼葛洛庞帝．1996. 数字化生存［M］．胡永，范海燕，译. 海口：海南出版社．

杨宗英．1996. 电子图书馆的现实模型［M］．中国图书馆学报(2): 24-29．

刘斌．2014. 力学［M］．合肥：中国科学技术大学出版社．

参考文献里面标点符号：英文文献用半角,中文文献用全角。

# 附 录

附录编号依次编为附录A，附录B。附录标题各占一行，按一级标题编排。每一个附录一般应另起一页编排，如果有多个较短的附录，也可接排。附录中的图表公式另行编排序号，与正文分开，编号前加“附录A-”字样。

本部分内容非强制性要求，如果论文中没有附录，可以省略《附录》。

# 攻读学位期间取得的研究成果

1）已发表或已录用的学术论文、已出版的专著/译著、已获授权的专利按参考文献格式列出。

2）科研获奖，列出格式为：

获奖人（排名情况）．项目名称．奖项名称及等级，发奖机构，获奖时间．

3）与学位论文相关的其它成果参照参考文献格式列出。

4）全部研究成果连续编号编排。

样例：

1. Wei ZY, Tang YP, Zhao WH, et al. Rapid development technique for drip irrigation emitters[J]. RP Journal,UK., 2003, 9(2):104~110 (SCI: 000350930600051; EI: 03187452127).
2. 魏正英,唐一平,卢秉恒.滴灌管内嵌管状滴头的快速制造方法研究[J].农业工程学报, 2001,17(2):55~58 (EI:01226526279,01416684777).

用于双盲评审的论文，只列出已发表的学术论文的篇名、发表刊物名称。

学位论文独创性声明（1）

本人声明：所呈交的学位论文系在导师指导下本人独立完成的研究成果。文中依法引用他人的成果，均已做出明确标注或得到许可。论文内容未包含法律意义上已属于他人的任何形式的研究成果，也不包含本人已用于其他学位申请的论文或成果。

本人如违反上述声明，愿意承担以下责任和后果：

1．交回学校授予的学位证书；

2．学校可在相关媒体上对作者本人的行为进行通报；

3．本人按照学校规定的方式，对因不当取得学位给学校造成的名誉损害，进行公开道歉。

4．本人负责因论文成果不实产生的法律纠纷。

论文作者（签名）： 日期： 年 月 日

学位论文独创性声明（2）

本人声明：研究生所提交的本篇学位论文已经本人审阅，确系在本人指导下由该生独立完成的研究成果。

本人如违反上述声明，愿意承担以下责任和后果：

1．学校可在相关媒体上对本人的失察行为进行通报；

2．本人按照学校规定的方式，对因失察给学校造成的名誉损害，进行公开道歉。

3．本人接受学校按照有关规定做出的任何处理。

指导教师（签名）：日期： 年 月 日

学位论文知识产权权属声明

我们声明，我们提交的学位论文及相关的职务作品，知识产权归属学校。学校享有以任何方式发表、复制、公开阅览、借阅以及申请专利等权利。学位论文作者离校后，或学位论文导师因故离校后，发表或使用学位论文或与该论文直接相关的学术论文或成果时，署名单位仍然为西安交通大学。

论文作者（签名）： 日期： 年 月 日

指导教师（签名）： 日期： 年 月 日

(本声明的版权归西安交通大学所有，未经许可，任何单位及任何个人不得擅自使用)

1. 脚注是对文中有关内容的解释、说明或补充，使用上角标（序号①、②…）标注，脚注可用小号字（一般小五号宋体）列在相应正文同一页最下部并与正文部分用细线（版面宽度的1/4长）隔开。（删除脚注的方法：直接删除正文中的脚注编号即可） [↑](#footnote-ref-1)
2. 1 Ebps =Tbps =bps [↑](#footnote-ref-2)