



中国计算机学会文集

China Computer Federation
Proceedings
CCFP 0019

2012 中国计算机科学技术 发展报告

中国计算机学会 主编

图书在版编目 (CIP) 数据

2012 中国计算机科学技术发展报告 / 中国计算机学会主编 . —北京：机械工业出版社，2013. 10
(中国计算机学会文集)

ISBN 978-7-111-44306-3

I. 中… II. 中… III. 计算机科学—发展—研究报告—中国—2012 IV. TP3-12

中国版本图书馆 CIP 数据核字 (2013) 第 235661 号

版权所有 · 侵权必究

封底无防伪标均为盗版

本书法律顾问 北京市展达律师事务所

本书是由中国计算机学会学术工作委员会组织编写的具有权威性的计算机科学技术年度发展报告，总结了 2012 年度计算机科学技术发展的热点问题和现状，展望了未来的发展趋势。今年的发展报告组织并选择了多媒体数据索引与检索技术研究进展、高性能计算机系统冷却技术、SDN 技术发展、计算机辅助设计与图形学研究进展与趋势、嵌入式系统抗恶劣环境技术的研究进展、普适计算的研究进展与趋势、微处理器芯片可靠性设计研究进展、软件测试技术研究进展、计算机存储技术的研究进展与趋势、计算系统虚拟化的研究新进展与趋势、社会化协同计算研究进展报告、多模态物理过程仿真研究进展与趋势 12 篇报告。这些报告分别由各个研究方向从事一线研究的科研人员撰写。这些报告基本上反映了我国计算机科学和技术工作者当前的研究进展，对学术研究有重要参考价值。

本书可作为广大计算机科学技术人员了解当前计算机科学技术发展动态的一个渠道，适合本领域决策人员和科研人员参考，并可作为高等院校计算机相关专业硕士生、博士生的参考书。

机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码 100037）

责任编辑：王春华

中国电影出版社印刷厂印刷

2013 年 10 月第 1 版第 1 次印刷

185mm × 260mm · 24.25 印张

标准书号：ISBN 978-7-111-44306-3

定 价：85.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88378991 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294 88379649 68995259

读者信箱：hzjsj@hzbook.com

前　　言

计算机科学已经成为支撑现代社会经济发展、社会进步和科技创新的最重要信息基础科学，计算机科学技术的水平是衡量一个国家基本国力和经济竞争力的重要标志。中国计算机学会自 2004 年以来每年都组织编写“中国计算机科学技术发展报告”，在国内产生了很大的影响，受到广大学会会员和计算机科技人员的广泛好评。

2012 年是中国发展进程中具有重大而深远意义的一年，中国共产党第十八次全国代表大会成功召开，顺利实现了中央领导机构新老交替。2012 年，国际金融危机深层次影响不断显现，世界经济低速增长态势延续，而以计算机科学为代表的信息技术对国民经济的重要性进一步得到提升。十八大报告明确提出“建设下一代信息基础设施，发展现代信息技术产业体系，健全信息安全保障体系，推进信息网络技术广泛运用”。2013 年 7 月 12 日，国务院常务会议研究部署促进信息消费，推动经济转型升级。

今年的发展报告包含多媒体数据索引与检索技术研究进展、高性能计算机系统冷却技术、SDN 技术发展、计算机辅助设计与图形学研究进展与趋势、嵌入式系统抗恶劣环境技术的研究进展、普适计算的研究进展与趋势、微处理器芯片可靠性设计研究进展、软件测试技术研究进展、计算机存储技术的研究进展与趋势、计算系统虚拟化的研究新进展与趋势、社会化协同计算研究进展报告、多模态物理过程仿真研究进展与趋势 12 篇报告，详细介绍了相应研究方向的国际研究现状、国内研究进展、国内外研究进展的比较以及发展趋势与展望等。

本报告将对进一步推动我国计算机科学技术的发展和信息化进程起到重要作用，同时，也是广大计算机科学技术人员了解当前计算机科学技术发展动态的一个渠道。

由于策划组稿时间较短，报告在形式和内容安排上都有待进一步改进，希望广大读者对本报告的编写工作多提宝贵的意见和建议，特别希望广大计算机学会会员献计献策，以便在今后的年度报告中逐步改进。

本年度发展报告的组织和策划工作得到了中国计算机学会各专业委员会的大力支持，共有 22 个专委会提交了年度发展报告。经学术工委的审查和评选，共有 12 篇报告收入本文集。在此，要特别向本年度所有发展报告的执笔人表示感谢，也要感谢各专业委员会的主任和秘书长付出的大量辛苦劳动。中国计算机学会郑纬民理事长、杜子德秘书长等对本报告的出版给予许多指导和支持，学术工委副主任陈熙霖研究员、过敏意教授，特别是学术秘书黄华教授在发展报告的策划、组织、评审等方面做了大量的工作，学会秘书处马琳女士协助处理了烦杂的事务性工作，在此一并表示感谢。

胡事民
中国计算机学会学术工作委员会主任
2013 年 8 月

目 录

前言

多媒体数据索引与检索技术研究进展 CCF 多媒体专业委员会

1 引言	1
2 国际研究现状	2
2.1 视频事件识别与标注	2
2.2 高维索引结构	6
2.3 跨媒体检索理论与技术	11
3 国内研究进展	17
3.1 视频事件识别与标注	17
3.2 高维索引结构	18
3.3 跨媒体检索理论与技术	20
4 国内外研究进展比较	21
4.1 视频事件识别与标注	21
4.2 高维索引结构	22
4.3 跨媒体检索理论与技术	24
5 发展趋势与展望	25
5.1 视频事件识别与标注	25
5.2 高维索引结构	26
5.3 跨媒体检索理论与技术	27
6 结束语	27
参考文献	28
作者简介	36

高性能计算机系统冷却技术 CCF 工程与工艺专业委员会

1 引言	38
2 国际研究现状	39
2.1 国外典型高性能计算机系统的冷却技术	39
2.2 计算机等冷却技术的前沿研究	49
3 国内研究进展	51
3.1 国内典型高性能计算机系统的冷却技术	51
3.2 国内冷却技术的相关研究	56

4 国内外研究进展比较	58
4.1 研发模式的比较	58
4.2 研发内容的比较	59
4.3 研发水平的比较	59
5 发展趋势与展望	59
5.1 芯片液体冷却新技术的持续发展	60
5.2 液体冷却系统与空调系统的有机结合	60
5.3 冷却系统的节能和绿色化发展	60
5.4 冷却系统的商业化发展	60
6 结束语	61
参考文献	61
作者简介	62

SDN 技术发展报告 CCF 互联网专业委员会; CCF 网络与数据通信专业委员会

1 引言	63
2 软件定义网络的基本思想和架构	65
2.1 SDN 主要思想和架构	65
2.2 研究问题与主要挑战	66
2.3 SDN 实例	67
3 控制平面: 软件定义网络控制器	68
3.1 控制平面功能	68
3.2 主要技术挑战	69
3.3 主要研究进展	69
4 数据平面: 可编程网络转发设备	72
4.1 数据平面功能	72
4.2 技术挑战	72
4.3 关键技术研究进展	74
4.4 原型系统实现进展	79
5 软件定义网络的应用	80
5.1 基于 SDN 的流量工程	80
5.2 基于 SDN 的网络虚拟化	82
5.3 基于 SDN 的网络节能	85
6 国内研究进展	87
7 总结与展望	90
参考文献	91
作者简介	96

计算机辅助设计与图形学研究进展与趋势…… CCF 计算机辅助设计与图形学专业委员会

1	引言	99
2	国际研究现状	101
2.1	基于知识的建模/高层语义分析	101
2.2	数字城市建模	103
2.3	体网格生成与编辑	105
2.4	可视媒体	107
3	国内研究进展	107
3.1	基于知识的建模/高层语义分析	107
3.2	数字城市建模	110
3.3	体网格生成与编辑	111
3.4	可视媒体	112
4	国内外研究进展比较	113
4.1	基于知识的建模/高层语义分析	113
4.2	数字城市建模	115
4.3	体网格生成与编辑	115
4.4	可视媒体	116
5	发展趋势及展望	117
5.1	基于知识的建模/高层语义分析	117
5.2	数字城市建模	118
5.3	体网格生成与编辑	118
5.4	可视媒体	119
6	结束语	119
	致谢	120
	参考文献	120
	作者简介	127

嵌入式系统抗恶劣环境技术的研究进展 …… CCF 抗恶劣环境计算机专业委员会

1	引言	129
2	国外研究现状	131
2.1	体系架构	131
2.2	容错技术	132
2.3	可靠网络技术	132
2.4	片上系统技术	133
2.5	健康监控技术	134
2.6	安全软件技术	134

2.7 工程化技术	136
3 国内研究进展	137
3.1 体系结构	137
3.2 容错技术	137
3.3 可靠网络技术	138
3.4 片上系统技术	139
3.5 健康监控技术	139
3.6 安全软件技术	140
3.7 工程化技术	141
4 国内外研究进展比较	141
5 发展趋势与展望	142
6 结束语	144
参考文献	144
作者简介	146

普适计算的研究进展与趋势 CCF 普适计算专业委员会

1 引言	148
1.1 背景：新技术革命推动普适计算发展	148
1.2 未来普适计算愿景——iCPS	150
2 国际研究现状	152
2.1 “人—机—物”模型	152
2.2 智慧终端	153
2.3 人机交互	156
3 国内研究进展	158
3.1 用户模型	158
3.2 基础设施和服务平台	159
3.3 新型普适智能服务	162
3.4 智能终端感知和交互技术	165
4 国内外研究进展比较	167
5 发展趋势与展望	168
6 结束语	169
参考文献	169
作者简介	172

微处理器芯片可靠性设计研究进展 CCF 容错计算专业委员会

1 引言	174
------------	-----

2 国际可靠性设计研究现状	174
2.1 容芯片制造缺陷	174
2.2 容芯片中的瞬态故障	175
2.3 容芯片中的永久故障	178
2.4 国际高可靠微处理器实例：HP NonStop 系统	178
3 国内可靠性设计研究现状	179
4 可靠性评估体系	181
4.1 可靠性评估概述	181
4.2 电气级可靠性估计	182
4.3 门级可靠性估计	183
4.4 系统级可靠性估计	187
5 国内外研究现状比较	192
5.1 可靠性设计方法现状比较	192
5.2 可靠性理论研究现状比较	192
6 可靠性设计技术展望	193
6.1 近阈值电压计算模式下的可靠性设计	193
6.2 三维芯片的可靠性设计	194
6.3 面向应用的可靠性设计	195
7 结束语	196
参考文献	196
致谢	201
作者简介	201

软件测试技术研究进展与趋势 CCF 软件工程专业委员会

1 引言	203
2 国际研究现状	206
2.1 测试用例设计技术	207
2.2 缺陷定位与预测技术	212
2.3 面向特定领域的测试	213
3 国内研究进展	219
4 国内外研究进展比较	223
5 进展趋势与展望	224
6 结束语	228
参考文献	228
作者简介	236

计算机存储技术的研究进展与趋势 CCF 体系结构专业委员会

1 引言	238
------------	-----

2 国际研究现状	239
2.1 存储介质	239
2.2 存储结构	244
2.3 存储系统	247
2.4 存储应用	250
3 国内研究进展	253
3.1 新型存储器件	253
3.2 存储系统	254
4 国内外研究进展比较	255
5 发展趋势与展望	256
6 结束语	256
参考文献	257
作者简介	261

计算系统虚拟化的研究新进展与趋势 CCF 系统软件专委会

1 序言	262
2 嵌入式虚拟化技术的研究进展与趋势	264
2.1 引言	264
2.2 嵌入式虚拟化	265
2.3 国外研究现状	267
2.4 国内研究现状	270
2.5 国内外研究进展比较	271
2.6 发展趋势与展望	272
2.7 小结	272
3 GPU 虚拟化技术的研究进展与趋势	273
3.1 引言	273
3.2 国外研究现状	274
3.3 国内研究现状	277
3.4 国内外研究进展比较	278
3.5 发展趋势与展望	278
3.6 小结	279
4 云计算平台中虚拟资源性能计量方法的研究进展与趋势	279
4.1 引言	279
4.2 常用的云计算资源性能计量方法	279
4.3 存在的问题	281
4.4 国内外研究现状	282

4.5	发展趋势与展望	283
4.6	小结	284
5	云计算资源调度的研究进展与趋势	284
5.1	引言	284
5.2	国外研究现状	285
5.3	国内研究现状	290
5.4	国内外研究进展比较	292
5.5	发展趋势与展望	293
5.6	小结	294
6	云计算平台中虚拟化安全的研究进展与趋势	294
6.1	引言	294
6.2	虚拟云平台安全威胁	295
6.3	问题、挑战与支撑技术	296
6.4	多层次虚拟化平台安全解决方案	301
6.5	国内外研究进展比较	307
6.6	小结	307
7	总结	308
	参考文献	308
	作者简介	317

社会化协同计算研究进展报告 CCF 协同计算专委会

1	引言	319
2	国际研究现状	320
2.1	面向协同工作的社会网络分析	320
2.2	群体思维与社会智能	325
2.3	情感挖掘与人人协同	329
2.4	移动协同环境	333
3	国内研究进展	337
3.1	国内社会化协同系统开发与应用现状	337
3.2	国内主要研究团体和成果	341
	参考文献	343
	作者简介	349

多模态物理过程仿真研究进展与趋势 CCF 虚拟现实与可视化技术专业委员会

1	引言	351
2	国际研究现状	352

3 国内研究进展	356
4 国内外研究进展比较	362
5 国内外研究发展趋势展望	364
6 结束语	366
参考文献	366
作者简介	372
关键词索引	373
作者索引	375

多媒体数据索引与检索技术研究进展

CCF 多媒体专业委员会

摘要

随着计算机网络、社交媒体、数字电视和多媒体获取设备的快速发展，多媒体数据的生成、处理和获取变得越来越方便，多媒体应用日益广泛，数据量呈现出爆炸性的增长，已经成为大数据时代的主要数据对象。然而，由于多媒体数据本身的非结构化特性，使得多媒体数据的处理和检索相对困难。如何有效地存储、组织和管理这些多媒体数据，如何有效地按照多媒体的内容和特性去存取和检索这些数据，已经成为一种迫切的需求。本报告针对大数据时代多媒体数据的索引和检索问题，围绕高维索引结构、视频事件识别和标注、跨媒体检索三个研究方向，系统阐述 2012 年度多媒体数据索引与检索技术的发展状况，并对未来的发展趋势进行展望。

关键词：视频，事件识别，索引，跨媒体，检索

Abstract

With the rapid development of computer network, social media, digital television and communications technology, generating, processing and access to multimedia data has become more convenient. Multimedia applications are increasingly widespread, and the amounts of multimedia data are showing the explosive growth. In the era of big data, multimedia data has become a major data objects. However, due to its unstructured nature of multimedia data, the processing and retrieval of multimedia data is relatively difficult. How to effectively store, organize, and manage these data, how to access and retrieve data in accordance with the content and features of multimedia data, has become an urgent need. Around high-dimensional index structure, video event recognition and annotation, cross-media retrieval three research directions, this report elaborates their development status in 2012 and discusses the future development trend.

Keywords: video, event recognition, index, cross-media, retrieval

1 引言

当前，全球网络的发展和普及已经达到空前的规模，网络已成为人们日常生活环境中的重要部分。来自微博、手机、社交网站、新闻网站以及多媒体共享网站中的文本、图像、视频和空间属性等多媒体数据每天以惊人速度迅猛增长，网络多媒体数据与现实社会之间相互作用的范围之广和程度之深史无前例，在社会、经济和文化等诸多方面发挥着巨大效用，影响着无数人的思维、表达和行为方式。与文本数据相比，多媒体数据

提供了大量有用的信息，其内容更加丰富、直观和生动。一方面，丰富的多媒体数据包含海量信息，这是其他媒体类型所无法比拟的；但在另一方面，多媒体数据日益庞大的数据量、非结构化的数据形式和内容的多义性，又为其方便用户交互操作设置了障碍，影响了它发挥更好的作用。人们已经习惯于在互联网上查找各种信息，目前，很多搜索引擎（如谷歌、百度和雅虎等）已经能很好地解决文本的搜索问题，但是对图像和视频数据的检索仍无有效的方法，其主要原因是这些数据的数据量大而且内容复杂，导致我们现在还缺乏有效的手段对其建立适合检索的索引。要对这些数据进行高效的操作，如浏览、访问和缩略等，需要对多媒体内容进行基于内容的分析和合适的索引。

为了克服文本检索技术的局限性，基于内容的图像和视频检索在 20 世纪 90 年代应运而生。不同于文本检索使用图像的文本描述信息，基于内容的图像和视频检索利用图像和视频的视觉特征或其他语义特征来构建索引结构。基于内容的图像和视频检索一般分为三个部分：特征提取、高维索引构造和检索系统设计。本文将对 2012 年度的多媒体信息的索引和检索技术的发展情况做综述性介绍，内容涉及高维数据索引、视频事件识别与标注和跨媒体搜索。

2 国际研究现状

2.1 视频事件识别与标注

视频事件识别与标注就是根据视频所体现的内容，通过提取、存储、分析视频中的运动特征，并用来分析视频中的事件，从而按概念对其赋予标号。视频事件识别与标注是建立视频索引，进而实现高效视频数据处理（包括缩略、浏览和检索等）的必要基础。

视频事件识别是计算机视觉领域中备受关注的前沿方向，它是通过计算机对图像序列中的内容进行语义描述，自动地分析理解以知道场景中在哪里发生了什么，并以此来模拟人对视觉信息处理的全过程。视频标注的核心任务是给某个视频或镜头标注特定的语义概念，从而建立起视频片段集与概念集的映射。目前，视频标注的主要采用方法可以分为三类：纯手工标注方法、基于规则的标注方法、基于机器学习方法的标注方法。目前，视频标注的研究主要集中在如何利用视频的特点同时结合机器学习理论的相关知识来提高标注的准确性。从机器学习的角度看，语义标注通常可以看成一个监督学习的过程。首先将视频分割成不同粒度的片段，然后从中提取低层特征，针对各种低层特征的训练集建立模型，实现对未知片段的标注。因为概念之间并不互斥，所以一个片段可以有多个概念，每个概念的标注可以看成一个二分类问题，根据标注的片段是否具有该概念特征将其标注为“正”或者“负”。对视频进行语义标注，可以利用机器学习的理论结合视频本身的特点，尽量减少人工标注过程，实现半人工或者全自动的标注，最终达到可实际应用的目标。在统计机器学习理论中，为保证模型的泛化能力，往往需要大

量有标记的训练样本对模型进行训练。由于对视频语义标注时面临着训练样本不足的情况，因此利用大量未标注的样本进行训练分类的半监督学习方法成为了视频标注研究的热点。随着多媒体视频内容在广播和互联网爆炸性地增长，视频事件识别与标注具有高度的学术价值和广泛的应用前景。其研究成果已经应用在智能监控、高级用户接口、虚拟现实、运动分析、视频管理和检索等生产生活领域，并逐渐展现出更为广阔的应用前景。

目前，在世界范围的很多高校和科研机构都开展了视频事件识别与标注方面的研究。例如，法国国家信息与自动化研究院（INRIA）的 WILLOW 小组研究的一个课题是 3D 物体和场景建模、分析以及检索，同时对人体行为进行分类和场景识别；感知理解学习系统的行为识别（Perception Understanding Learning Systems for Activity Recognition, PULSAR）研究小组通过场景理解来识别人体行为，并开发了一个用于行为识别的软件架构；行为检测中感知、识别和集成（Perception, Recognition and Integration for Observation of Activity, PRIOA）研究小组正在研究一个交互系统，该系统可以检测单个目标，求出它们的位置和运动轨迹信息，并能识别单个个体或者人群的行为。英国雷丁大学（University of Reading）先后开展了在公共空间中对人们进行监控和理解的稳定方法（Robust Methods for Monitoring and Understanding People in Public Spaces, REASON）、视频场景理解评测（Evaluation of Video Scene Understanding, ETISE）、公共安全中拥挤区域中综合监控（Integrated Surveillance of Crowded Areas for Public Security, ISCAPS）等项目以研究基于视觉的单人行为以及人与人之间的交互行为识别和场景理解等。另外，如美国的麻省理工学院（MIT）、加州大学伯克利分校（UC Berkeley），英国的牛津大学（Oxford University）、剑桥大学（Cambridge University）等也开展了视频事件识别与标注的研究工作。另外，当前国际上一些权威期刊如 IJCV（International Journal of Computer Vision）、CVIU（Computer Vision and Image Understanding）、PAMI（IEEE Transaction on Pattern Analysis and Machine Intelligence）、IVC（Image and Vision Computing）和重要的学术会议如 ICCV（International Conference on Computer Vision）、CVPR（IEEE Computer Society Conference on Computer Vision and Pattern Recognition）、ECCV（European Conference on Computer Vision）、IWVS（IEEE International Work shop on Visual Surveillance）等将视频事件识别与标注研究作为主题内容之一，为该领域的研究人员提供了更多的交流机会。

视频事件识别与标注的关键问题是如何有效地抽取和表述事件的特征，以及如何根据得到的特征进行建模和分类。近些年来，大量研究人员对以上问题进行了深入的研究，并取得了一定的成果。Aggarwal 与 Cai 在文献[1] 中总结了 1997 年以前关于人的运动行为分析与识别的工作。Gavrila 在文献[2] 中介绍了 1999 年以前关于整个人体和手的运动行为方面的工作，从没有直接建立形状模型的 2D 方法、利用了形状模型的 2D 方法和 3D 方法三个方面进行了总结。Wang 与 Singh 在文献[3] 中总结了 2002 年以前的工作，回顾了包括脸、手、腿等人体部分和整体的跟踪方法与行为分析识别等方面内容；Wang 等人在文献[4, 5] 中从底层到高层的各个方面分别总结了 2003 年和 2004 年以前的运动分析方面的工作，包括目标检测、分类、跟踪以及人的行为分析和理解。Ronald Poppe 在文

献[6]中回顾了2007年以前的人体运动分析方面的工作。Pavan等在文献[7]中总结了2008年以前行为分析的工作，对行为分析的几个相关技术进行了细致的讨论。Ronald在文献[8]中总结了2010年以前行为分析相关的工作。视频中的行为事件是一个时变信号，事件识别与标注问题就是一个时变信号的分类问题。对于行为分类的方法，按照不同的标准分成不同的类型，如基于规则的方法与基于统计的方法；基于产生式模型的方法和基于判别式模型的方法；模板匹配方法和状态空间方法；基于监督学习、无监督学习和半监督学习的方法等。本节将对国外视频事件识别与标注的相关工作进行分析总结，对不同的方法进行对比分析。

2.1.1 基于模板匹配的分析方法

模板匹配方法^[9~17]可以分为帧对帧匹配方法和融合匹配方法。采用模板匹配技术的行为识别方法将图像序列转换为一组静态形状模式，在识别过程中和预先存储的行为样本进行比较，以确认人的行为。例如，Bobick和Davis^[9]将行为图像序列表示为运动能量图和运动历史图，然后提取矩特征作为行为模板。Wang等人^[14]采用变形模板匹配的方法计算不同姿态间的距离，因此，相似的姿态能够被聚类在一起。Babu和Ramakrishnan^[18]提出了一种利用运动历史图和运动光流历史特征进行识别的方法，这种特征不仅给出了运动的位置信息，而且给出了运动的方向、快慢、大小等信息。Carlsson和Sullivan^[19]利用每类代表性的例子与视频的每帧匹配来识别网球中的正手和反手等姿态。Dedeoglu等人^[15]利用基于关键帧的方法建立了一个实时系统来识别各种行为。Daniel和Edmond^[16]采取基于例子的方法将长度不同的视频转换成一个固定长度的描述，基于这个描述训练分类器。Laptev和Lindeberg^[20]引入了Gallilean变化得到了速度变化自适应的时空兴趣点，解决了不同相对运动速度下的基于时空兴趣点的事件行为匹配问题。基于模板匹配的方法通常计算量比较小，但对于时空域的变化比较敏感，鲁棒性差。

2.1.2 基于状态空间的分析方法

基于状态空间模型的分析方法^[21~25]定义每个静态姿势作为一个状态，这些状态之间通过某种概率联系起来。任何运动序列可以看做这些静态姿势的不同状态之间的一次遍历过程，在这些遍历期间计算联合概率，其最大值被选择作为分类事件行为的标准。这类方法主要以隐马尔可夫模型和动态贝叶斯网络为基本手段进行事件分析与理解。例如，Yamato等人^[23]利用二维运动区域作为学习和识别的低层特征，通过引入隐马尔可夫模型(Hidden Markov Model, HMM)识别网球运动中的基本动作。Starner和Pentland^[24]提出了一种基于HMM的系统来识别美式手语。HMM是一种有效的时变信号的处理方法，它隐含了对时间的校正，并提供了学习机制和识别能力。但是，如果对于每一类事件都建立一个模型，那么所需要的训练样本将是巨大的。另外，因为这类方法往往没有封闭解，需要非线性优化，所以计算量相对较大。Luo等人^[26]将动态贝叶斯网络(Dynamic Bayesian Networks, DBN)引入事件识别，DBN是沿时间轴展开的贝叶斯网络。该文还对HMM和DBN进行了比较：在一个时间切片上，HMM只能含有一个隐含结点和一个观测

结点；而 DBN 在一个时间切片上是一个贝叶斯网络，可以包含多个有因果关系的结点；HMM 在一个时刻需要将所有的特征压缩到一个结点中，那么所需要的训练样本将是巨大的（相当于联合概率密度函数）；而 DBN 用多个结点描述，即用条件概率来形成联合概率，训练相对要简单；但是 DBN 的设计要比 HMM 复杂得多。总之，基于状态空间的事件识别算法可以避免对时间间隔建模的问题，但是需要的训练样本量大，计算复杂。

2.1.3 基于参数模型的分析方法

基于参数模型的事件分析算法是通过统计不同事件的概率分布，以包容现实中事件上的某些模糊变化。该方法包括产生式模型和判别式模型。

产生式模型^[26~29]是一种基于概率的不确定性推理模型，是公认的学习时间序列数据效果很好的模型。其主要思想是通过一个联合概率函数模拟随机观测数据和隐藏状态之间的关系。通常是根据事先标注好的数据对不同的事件方式进行训练，得到各自模型的参数。基于概率论的模型对处理由随机性引起的事件不确定性十分有效。在事件识别过程，新的观测数据要根据已经获取的多个事件模型参数计算似然，比较取其大者为识别结果。用于事件识别的产生式模型多为有向图，最常见的就是混合高斯模型（Gaussian Mixture Model, GMM）、马尔可夫随机场（Markov Random Field, MRF）^[27]、动态贝叶斯网络^[26]、隐马尔可夫模型^[28]、概率随机文法^[29]等。

判别式模型与产生式模型不同，产生式模型不能直接对条件概率建模，而是需要大量的训练数据去学习输入的联合概率分布。另外，产生式模型存在着严格的独立性假设，每个观测值被作为独立的单元对待，与序列中所有其他观察值无关，且认为当前观察值是由当前状态生成的，同其他状态相互独立，这样就无法表现观察值序列中存在的长距离依赖关系。与之不同的是，判别式模型直接对条件概率进行建模，它认为观察值决定状态，克服了产生式模型的严格独立性假设，对于特征数据之间的相互关系没有任何的限制，可以融合各种特征到模型中。在事件识别领域，最广泛应用的判别式模型包括支持向量机（Support Vector Machine, SVM）、条件随机场（Conditional Random Fields, CRF）和神经网络（Neural Networks, NN）^[30~32]。

2.1.4 基于半监督学习的分析方法

许多现有的方法^[33~38]是基于监督学习的方式，为了得到好的识别效果，这些方法需要大量的标签样本训练事件模型^[36~38]。但是，有标签的样本通常需要大量的人工标注，因此得到这些样本是比较困难和昂贵的。随着信息量的日益增加，传统的依赖于人工标记大量训练样本进行事件建模与识别的方法，在很多应用场合显得越来越不适应。因此，近几年来，许多研究者将注意力转移到了基于无监督和半监督学习的事件分析方法的研究中。由于事件比较复杂，因此无监督的学习方法效果并不好，而半监督的学习方法是监督学习方法和无监督学习方法的折中，是目前事件分析的一个热点研究方向。基于半监督方法能够根据少量的标签样本和大量的无标签样本得到一个好的训练模型。无标签的数据很容易从公共的监控摄像头得到，因此数据的获得极其简单，不需要太多人工代

价。由于这类方法能够自动地标记样本，更加适用于实际场合的事件识别与检测以及未知事件的发现。Zelnik-Manor 与 Irani 在文献[39]中首次提出了一种自动建模的方法。Zhong 等在文献[40]中提出了一种无监督的异常检测方法。Xiang 与 Gong 在文献[41]中提出了一种新的无监督的建模与异常检测方法。类似于 Zelnik-Manor 与 Irani 的方法，他们也采用谱聚类的方式来获取事件的样本。Niebles 等人在文献[13]中提出了一种概率图方法，这种方法利用了时空词特征和无监督的学习方法，能够自动地学习事件的类别。文献[14, 42]等中也是直接根据无标签的样本采用无监督的学习方式训练事件模型识别各种事件。但是，这里很少的方法是基于半监督的，而这种方法能够同时利用有标签和无标签的样本。Zhang 等提出了一种半监督建模方法^[43]，即首先通过基于监督学习的方法建立一种正常事件模型，然后在此基础上，通过迭代自适应的方法自动地建立异常事件的模型。Guan 等人^[44]提出了 en-co-training 的方法利用无标签的样本，他们的方法显示通过利用无标签的样本，能够提高学习算法的性能，但是他们的方法没有在公共的数据集上进行测试，和现有的方法也没有比较。

2.2 高维索引结构

随着互联网和多媒体技术的快速发展，以图像为代表的多媒体信息呈爆炸性增长。面对海量的图像库，只有对图像进行有效的组织以便于浏览、搜索和检索，人们才能快速并准确地获取自己感兴趣和喜爱的图片。按照研究角度和侧重点的不同，图像检索主要有两个研究方向：数据管理和计算机视觉。前者是基于图像的文本描述，而后者是基于图像的视觉内容。

基于文本的图像检索始于 20 世纪 70 年代末期，数据库中所有图像首先用文本进行描述，然后构建一个基于文本的图像管理系统用于检索。沿着该研究方向，如数据建模、多维索引和查询评价等许多图像检索相关的理论被提出并逐渐成熟。然而，当图像库的规模很大（图像数量在百万级或者亿级以上）时，人工标注既费时又费力，其局限性越来越明显。此外，由于图像内容的丰富性以及图像文本描述本身带有强烈的主观色彩，因此不同的人可能对同一幅图像有不同的理解，导致图像的人工描述存在差异。人工标注的主观性和不准确性可能会导致图像检索中不可纠正的误匹配。

为了克服文本检索技术的局限性，基于内容的图像检索在 20 世纪 90 年代应运而生。不同于文本检索使用图像的文本描述信息，基于内容的图像检索利用图像的视觉特征来构建索引结构，常见的图像视觉特征有颜色直方图、Gist 和 sift 等。基于内容的图像检索一般分为三个部分：视觉特征提取、高维索引构造和检索系统设计。对于大规模图像库，对大量的图像视觉特征进行某种处理并直接存储在计算机内存以提高图像检索速度已成为当前的研究趋势。更为关键的是设计一个高效的高维索引结构来组织图像的视觉内容，实现图像的准确和快速检索。

2.2.1 基于树结构的高维索引

如图 1 所示，按照所采用的相似度度量方法的不同，树结构索引分为向量空间索引

和度量空间索引两类。向量空间索引使用欧式距离来度量对象之间的相似度，例如 R 树^[45]和 KD 树^[46]。度量空间索引使用非欧氏距离来度量对象之间的相似度，例如 M 树^[47]。

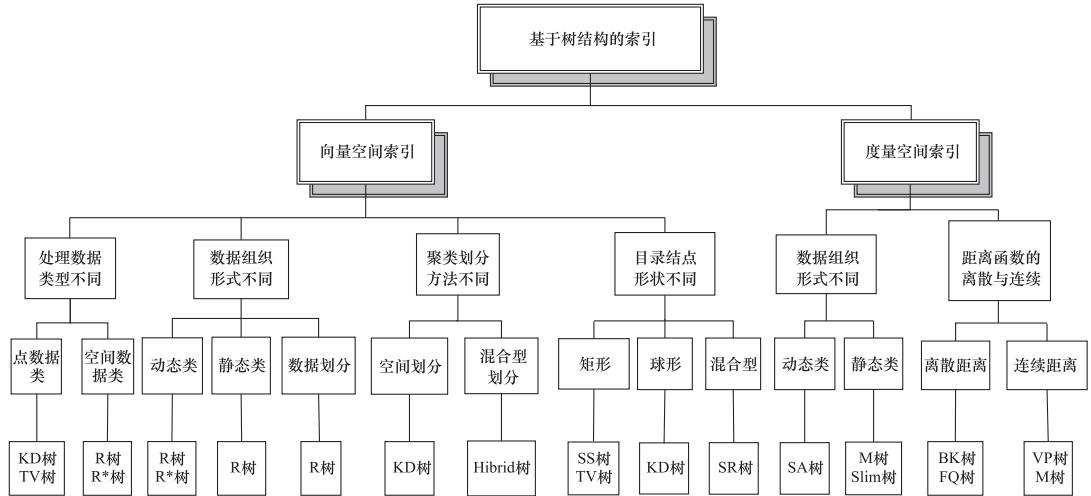


图 1 树结构索引的分类

KD 树虽然在低维空间中能有效降低搜索复杂度，但当特征向量维度超过 20 时，性能迅速下降，搜索复杂度接近于线性搜索，面临“维度灾难”问题。其原因在于面对高维数据，KD 树通常需要花费很多时间进行回溯搜索以寻找最近的特征点。为了提高 KD 树的性能，可以通过限制回溯搜索的次数，并用基于概率的近似最近邻特征点代替精确最近邻特征点来实现。通过构造多个 KD 树并且同时检索这些 KD 树，Silpa 等人^[48]对 KD 树进行了改进并对包含 500 000 个 sift 的特征集构建索引。

近年来，随着数据集的规模越来越大，由于树形索引结构通常需要比较大的存储空间来存储树的中间结点，并且其检索效率随着特征维度的增加而下降，因此国内外的研究者大都将关注点放在基于哈希的高维索引结构与基于视觉单词的倒排索引的研究，其原因在于这两类方法都可以在保证较好的检索速度和准确率的情况下，有效地降低构建索引所需的存储空间需求。

2.2.2 基于哈希的高维索引

基于哈希的高维索引分为两类：以 E2LSH 为代表的将特征点映射到低维欧式空间，使用欧式距离度量特征点之间的相似度；以谱哈希为代表的将特征点映射到低维汉明空间并保证欧式空间中相近的特征点具有相似的二进制编码，使用汉明距离度量特征点之间的相似度。

1. 位置敏感哈希

位置敏感哈希 (Locality Sensitive Hashing, LSH)^[49]最初是以严格的数学描述形式提出。随后，Indyk 等人^[50]在汉明空间上实现了面向低层高维特征向量的 LSH。LSH 利用

哈希函数将特征点映射到哈希表的相应桶来建立索引结构，其检索效率是亚线性的。

基于 p-stable 分布，Datar 等提出了精确欧式位置敏感哈希（Exact Euclidian Locality Sensitive Hashing, E2LSH）^[51]。相比 LSH，E2LSH 不需要对特征点进行任何空间嵌入，就可以直接应用于欧式空间的特征点。针对欧式空间的 c 近似最近邻问题，Andoni 等^[52]提出了近似优化哈希算法。

为了降低 E2LSH 的存储空间需求，Panigrahy 提出了基于熵的 LSH^[53]。基于熵的 LSH 随机产生与查询特征点相近的扩展对象并作为新的查询，并将所有查询对象的查询结果的合集作为最终查询结果，其核心思想是用时间换空间。Lv 等人在基于熵的 LSH 的基础上提出了 multi-probe LSH^[54]。为了提高查询准确率，multi-probe LSH 使用 step-wise 和 query-directed 两种探查序列来探查哈希表中多个临近桶，使得查询对象的近邻特征点尽可能被检索到。

Jegou 等人提出了查询自适应 LSH 方法（Query-Adaptive LSH, QA-LSH）^[55]，为查询特征点从一组哈希函数中在线选择最合适的哈希函数。其提出的相关性准则是计算每个哈希函数的准确率期望值和在线选择哈希函数的关键。QA-LSH 虽然仅仅在哈希函数的构造方式上不同于 E2LSH，但是在查询速度和准确率上可以获得更好的平衡。

利用 K-Means 聚类方法构造哈希函数，Pauleve 等人提出了 KLSH^[56]。其中，哈希函数将每个特征点映射到距离最近的聚类中心。鉴于原始 K-Means 聚类方法虽然具有非常好的聚类效果，但是当聚类数量 k 很大时，K-Means 的训练和聚类的时间开销很大，层次 K-Means^[57]和近似 K-Means^[58]可以用来代替原始 K-Means 以提高效率。

2. 哈希编码

哈希编码的思想是利用少量的二进制编码使得相似的特征点具有相同或者相似的二进制编码，通常使用汉明距离度量特征点之间的相似度。Slakhutdinov 等人^[59]提出的语义哈希是哈希编码研究中一个经典研究工作。语义哈希使用带有若干隐藏层的自动编码器，其架构是一个局限型玻尔兹曼机（Restricted Boltzmann Machine, RBM），其中，只有层内结点与层外结点之间存在关联。将 RBM 和 Boosting 相结合用于学习二进制编码，Torralba 等人^[60]实现了对百万级以上图像库进行编码。在此基础上，Weiss 等人提出了谱哈希（Spectral Hashing, SH）^[61]。SH 在编码效率上与 Torralba 的编码方法相同，但更易于实现。

在文献[62, 63]的基础上，文献[64]提出了球形哈希（Spherical Hashing, SH），将空间上位置邻近的特征点映射到相同或者相似的二进制编码。球形哈希使用超球体来划分特征空间，每个超球体与一个哈希函数相关联。此外，球形哈希提出了一种球形汉明距离来提高汉明距离的区分度。

利用监督信息，监督式哈希索引方法可以获得比非监督哈希索引方法更好的检索准确率，例如，LDAH^[65]、BRE^[66]和 MLH^[67]。然而，监督式哈希索引方法在训练学习阶段通常比非监督哈希索引方法花费更多的时间，不仅如此，当图像的标签数据太少或者有太多噪声时，这类方法可能会出现过度拟合的现象。为了提升监督式哈希索引方法的性能，Wang 等人^[68]提出了半监督哈希（Semi-Supervised Hashing, SSH）方法，在最大

化标签数据和无标签数据哈希码的差异性和独立性的同时最小化经验误差。在后续研究中，他提出了 3 种不同的 SSH 方法^[69]，分别是正交哈希、非正交哈希和顺序哈希。类似研究还包括基于核的监督哈希（Kernel-based Supervised Hashing，KSH）^[70] 编码方法。KSH 仅需要少量的监督信息和训练开销就可以生成高质量的紧凑哈希编码，使得相似特征点的汉明距离最小化和不相似特征点的汉明距离最大化。

哈希编码方法将特征点映射为二进制编码后，通常根据整段编码将其插入哈希表的相应桶中。然而，当编码长度超过 32 比特时，这类方法将不会显得那么高效。对此，Norouzi 等人提出了多重索引哈希机制^[71]，其将整段二进制编码划分为若干个子串编码并对每个子串编码构建哈希表。多重索引哈希支持汉明空间里的精确 K 最近邻检索和 r 邻近范围检索。当原始特征经过哈希编码得到的二进制编码服从均匀分布时，多重索引哈希的检索时间复杂度是亚线性，比线性检索快 100 倍以上。

2.2.3 基于视觉单词的倒排索引

词袋模型（Bag-Of-Feature，BOF）作为首个基于视觉单词的倒排索引，引入自文本检索系统。对于一幅图像，首先提取局部描述符 SIFT；然后根据事先训练好的视觉词汇，将局部描述符量化到相应的视觉单词；最终形成一个用以描述图像的高维 BOF 向量。BOF 图像描述符用一个倒排文件对其进行构建索引，索引文件的每个索引列表对应一个视觉单词，索引列表中每个结点存储的信息是图像 ID 和该视觉单词在图像中发生的次数。BOF 图像描述符之间的相似度通常用欧氏距离或者角度距离度量。图像描述符是影响图像索引与检索的根本因素，因而如何快速并准确地计算图像描述符非常重要。此外，构建一个高效的倒排索引以提高图像检索速度，以及对图像描述符进行压缩和编码，使得大规模图像描述符可直接存储在计算机内存是基于内容的大规模图像检索的关键。

1. 基于视觉单词的图像描述符

基于视觉单词的图像描述最初是由 Sivic 等人^[72]从文本检索引入，利用图像的 sift 局部特征计算 BOF 图像描述符。BOF 图像描述符向量的每个分量的权值根据 tf-idf 机制计算。根据 BOF 的各维度分量的权值计算方式不同，还有另外两种 BOF 描述符：一种是词频 BOF，其分量表示视觉单词在图像中发生的次数；另一种是二进制 BOF，其分量表示视觉单词是否在图像中发生。

(1) 构造视觉词汇

BOF 图像描述符的高维度意味着需要训练大量的视觉单词。标准 BOF^[72] 使用的 K-Means 聚类方法在大规模视觉单词的训练和聚类上需要耗费大量时间，变得非常低效。通过构建词汇树，Nister^[57] 提出分层 K-Means 聚类方法，提高了大规模视觉词汇训练和聚类的效率。Philbin 等人^[58] 提出了近似 K-Means（Approximate K-Means，AKM）聚类方法，其通过使用 8 个随机 KD 树来构建词汇树，提高了聚类速度。Avrithis 等人^[73] 将高斯混合的灵活性与所需的尺度属性相结合，提出了近似高斯混合（Approximate Gaussian Mixture，AGM）用于图像索引与检索中大规模视觉词汇的构造。

为了降低视觉词汇的存储空间和提高时间效率，文献[74]提出了随机位置敏感词汇

(Random Locality Sensitive Vocabulary, RLSV)。RLSV 将 LSH 和随机森林 (random forest) 相结合，通过一系列地随机线性二部划分来构造视觉词汇。同 K-Means 相关方法相比，RLSV 不需要训练和最近邻聚类的过程。

(2) 特征分配

传统的 BOF 描述符特征量化方法采用硬分配，即将局部特征分配到距离最近的视觉单词。然而，硬分配存在视觉单词不确定性和视觉单词疑惑性两个缺陷。为了克服硬分配的这两个缺陷，文献[75]提出了基于核的软分配，其首先计算特征点到每个视觉单词的权值，然后将特征点分配到权值大于预先设定阈值的视觉单词。文献[76]提出了描述符空间软分配，将特征点分配到最近的 k 个视觉单词来减少量化误差。类似的方法还有文献[77]提出的软分配方法，将特征点分配到最近的 k 个视觉单词。虽然特征点到每个视觉单词的权值计算方法与文献[76]类似，但其对权值进行归一化处理，使得特征点到所有视觉单词的权值总和为 1。

(3) 其他基于视觉单词的图像描述符

基于 BOF 和 Fisher 核，Jegou 等人^[78]提出了局部聚合描述向量 (Vectors of Locally Aggregated Descriptors, VLAD)。VLAD 将 SIFT 描述符量化到少量的视觉单词并形成紧凑的全局图像描述符，其由 SIFT 特征提取、视觉词汇训练、特征分配、向量残差聚合和残差向量串联 5 个步骤构成。VLAD 比 BOF 具有更高的准确性但维度更低。在其后续研究中^[79]，作者利用最新的 Fisher 核研究进一步提升了 VLAD 的准确性。将每个视觉单词对应聚类中最远特征点与聚类中心之间的距离作为阈值，Chen^[80]通过移除位于两个视觉单词之间边界上的异常特征点，进一步提高了 VLAD 的准确性。文献[81]将图像的全局属性描述和 Fisher 向量相结合，进一步提高图像描述符的准确性。

2. 描述符压缩与编码

哈希编码可以被认为是对致密描述符向量压缩与编码的有效方法。

基于汉明嵌入 (Hamming Embedding, HE) 和弱几何一致性 (Weak Geometric Consistency, WGC)，Jegou 等人^[82]提出了比 BOF 图像描述符更加紧凑和准确的图像描述符。HE 将特征描述符从欧式空间映射到汉明空间并对其进行编码，同时保证描述符在汉明空间上与其最近邻特征的距离在欧式空间上也是最近的。WGC 用于过滤与查询图像的角度和尺度比例不一致的图像。

基于 HE，文献[83]提出了一种近似 BOF 图像描述符 miniBOF。其利用 m 个预先定义的稀疏投影函数将 BOF 图像分解为 m 个低维描述符，并利用 HE 分别对这 m 个 miniBOF 描述符进行编码。一幅图像的标准 BOF 图像描述符通常需要 10K 字节左右存储空间，而当 BOF 描述符分解成 16 个 miniBOF 描述符时，一幅图像只需 320 字节。

建立在 Dong^[84] 和 Gorda^[85] 将非对称距离应用于 LSH 的思想上，Jain 等人^[86] 提出了非对称汉明嵌入 (Asymmetric Hamming Embedding, AHE) 机制，提高了 HE 的编码性能。AHE 在没有增加任何开销的基础上提高了 HE 的编码效率。

描述符量化同样可以用于描述符的编码，接下来将对近年来大规模图像索引与检索中典型的描述符量化和编码方法进行论述。

文献[87]提出了积量化 (Product Quantization, PQ) 将特征描述符向量分割为 m 个子向量并分别对其进行编码。在每段子向量对应的特征空间上，利用 K-Means 方法训练量化器并对特征子向量进行量化与编码。与全局量化器相比，PQ 不仅使用更少的聚类中心还明显减少了存储空间。PQ 只是将描述符向量按照固定长度进行均分得到子特征向量，没有考虑到特征向量各个维度分量的实际分布。Brandt^[88]提出了将转换编码和 PQ 相结合的描述符量化方法。

3. 基于视觉单词的倒排索引

基于视觉单词的倒排索引结构最初是在文献[72]中提出，每个视觉单词对应一个倒排列表，倒排列表中结点存储的信息是图像 ID 和视觉单词在该图像中发生的次数。对于一幅查询图像，在计算 BOF 图像描述符后，需要检索所有在该图像中发生的视觉单词对应的索引列表，这会造成某些图像将被重复检索多次。针对这种现象，一些研究者开始用包含较少数量聚类中心的粗量化器构建倒排索引结构。

基于 PQ，文献[87]提出了基于非对称距离计算的倒排索引 (InVerted File with Asymmetric Distance Computation, IVFADC) 用于特征描述符的最近邻搜索。IVFADC 利用 K-Means 聚类方法训练包含 k 个聚类中心的量化器，每个聚类中心对一个倒排列表，形成一个包含 k 个倒排列表的倒排索引结构。倒排列表中结点存储的信息是特征描述符 ID 及其 PQ 编码。

索引结构中各个倒排列表中结点数量是否均衡同样会影响查询时间。Jegou 等人^[89]提出了均衡聚类机制，使得各索引列表对应聚类的规模更加均匀。Jegou 引入惩罚选项，通过人为增加特征描述符和较大聚类的聚类中心之间的距离，提高特征描述符落入较小聚类的概率。这种方法在不影响准确率的前提下，使得倒排索引中倒排列表的规模更为均衡。

基于 PQ，Babenko 等人^[90]利用多维表来构建倒排多维索引。Babenko 首先将特征描述符平分为 2 个子向量，然后分别在子向量所在特征空间上用 K-Means 训练包含 k 个聚类中心的量化器。将分别来自两个量化器的聚类中心进行组合为聚类中心对，形成一个二维表，从而构建了包含 $k \times k$ 个倒排列表的多维倒排索引结构。每个特征描述符根据量化得到的聚类中心对，插入到对应的倒排列表。

2.3 跨媒体检索理论与技术

有别于传统的结构化和非结构化数据，网络上的数据是以文本、图像和视频等多种形式体现的，这些从不同渠道获取的文本、图像和视频等不同类型媒体及其与之相关的社会属性信息更加紧密混合在一起，以一种新的形式，更为形象地表示综合性知识，反映个体和群体的社会行为，这种新的媒体表现形式称为“跨媒体” (cross-media)。

跨媒体表现出如下 4 个基本属性：

1) 固有的自然属性。网页、图像、视频和用户评价等结构化或非结构化信息具有异构、多阶和高维等自然属性。

2) 与现实生活密切相关的社会属性。网络平台上的跨媒体数据通过不同方式被赋予了评价、热度和偏好等社会属性，这些社会属性汇集成社会影响力，对跨媒体内容的语言理解产生主观影响。

3) 数据来源多样以及不同类型媒体数据相互关联和动态演化。来源于不同渠道的结构化与非结构化数据之间交叉关联，且跨媒体数据所蕴含的主题和事件通过一定机制传播并动态演化。

4) 丰富的表达和呈现能力。与人脑通过不同感官渠道对外界认知一样，通过跨媒体形式，能更为自然地展示客观世界及其包含知识。

2010 年 1 月《Nature》杂志发表的对下一个 10 年科技发展进行展望的“2020 Vision”报告中指出：今后媒体内容将由来自不同渠道的文本、图像、语音、视频和亲朋好友交互属性所组成，即以“跨媒体”这一形式组合。

本节将对跨媒体搜索中所涉及的跨媒体度量、跨媒体索引和跨媒体排序技术的发展现状进行系统阐述。

2.3.1 跨媒体度量

在跨媒体检索中，如何挖掘不同类型数据之间的内在联系，进而对跨媒体数据之间的相似度进行计算，是跨媒体检索要解决的重要内容。

Hotelling 提出的典型相关性分析 (Canonical Correlation Analysis, CCA)^[91] 被用来实现不同类型数据之间的跨媒体检索^[92]。CCA 是一种用来建模两个高维随机变量间线性关联关系的一种方法，它的主要思想是使得两个变量经过线性投影后得到的向量之间的皮尔森相关系数达到最大。但是，我们知道，由于不同类型异构数据分布的非线性，仅仅通过线性投影无法挖掘数据之间的关联。因此，核函数被引入 CCA，从而提出了 Kernel CCA (KCCA)^[93]。在不同类型数据的关联性建模过程中，由于从数据中提取的高维特征存在噪音和冗余，因此我们希望对高维特征进行选择之后再进行投影和计算数据之间的关联性，从而使得模型之间的关联性建模更加具有可解释性，Sparse CCA^[94] 由此被提出。同时，高维特征之中可能存在组和层次等内在先验结构，充分利用这些先验结构，可以提升关联性建模的性能，如 Structured Sparse CCA^[95] 就是将结构先验引入 Sparse CCA 中所形成的不同类型数据之间关联性建模的方法。同时，为了对两种类型以上的数据之间的关联进行建模，文献[96]将传统 CCA 拓展为多视图 CCA 算法，使其可以被用来对多种类型数据之间的相关关系进行建模。但是，在上述 CCA 及其拓展模型中，数据的类别信息没有被利用。在实际应用中，很多情况下，我们还知道数据所具有的类别信息。如泛化多视图分析 (Generalized Multiview Analysis, GMA)^[97] 拓展了 CCA 算法，GMA 不仅在关联建模中可利用类别信息，还引入了核函数来反映数据的非线性分布。除此之外，还有一些基于 CCA 的非参数化方法如文献[98]。

在将文本或图像表达成单词词袋 (bag of textual words) 或视觉词袋 (bag of visual words) 后，一些通过潜在狄利克雷分配 (Latent Dirichlet Allocation, LDA)^[99] 来实现不同类型数据之间关联建模的方法也陆续被提出。原始的 LDA 针对的仅是单一类型数据

(通常状况下是文本), 用 LDA 处理多模态 (多类型) 数据需要做些改进^[100~102]等。Correspondence LDA (Corr-LDA)^[100]是最早对 LDA 进行拓展来处理两种不同类型数据的方法, Corr-LDA 对图像及其伴随标签文本之间的关联关系进行建模。但是在 Corr-LDA 这种方法中, 图像中所蕴含“主题”是由其对应文本所蕴含“主题”直接得来的, 这就限制了图像 - 文本关联关系的灵活性。文献[101]是对这一不足进行改进的一种方法, 其认为图像中所蕴含“主题”可由文本所蕴含“主题”线性组合而来, 这种方法比 Corr-LDA 灵活不少, 但是仍然存在图像中所蕴含主题受到文本中所蕴含主题限制的问题。为了解决这一问题, 文献[102]提出了多模态文档随机场 (Multi-modal Document Random Field, MDRF) 的方法。MDRF 采用马尔可夫随机场来刻画文档之间的关联关系。在 MDRF 中, 一个文档可以是任意一种模态数据所构成, 不同文档之间通过马尔可夫随机场相联系, 区别在于不同模态字典间的差异。在 MDRF 中, 不同模态数据之间完全是平等的, 这个方法可以更加自然刻画不同类型数据之间的关系, 但这一方法的缺陷在于先验知识仅通过马尔可夫随机场来影响模型, 使得其利用先验知识有限。在现实生活中, 不同类型数据之间的信息并不是完全相关的, 有一部分信息可能是仅属于本模态数据所私有的, 并不是不同模态数据中的共同信息。为了刻画这一客观现实, 文献[103]提出了一种将“主题”分为公共“主题”和私有“主题”的方法。该方法在对不同类型数据之间关联关系进行建模时具有更大灵活性。同时, 非参数化机制也被引入该方法, 用来自动确定“主题”数目。

随着字典学习这一方法的兴起, 通过字典学习来对不同类型数据之间关联性进行建模的手段也被陆续提出。如文献[104]应用字典学习来学习不同类型数据之间的关联关系。这一方法先假设存在一个全局共有系数空间, 每个模态数据所存在的空间是全局共有系数空间的子集。应该说, 这种方法并没有显式地求取不同模态系数之间的关联关系, 而是通过假设全局共有系数空间的存在, 从空间角度去考虑问题, 其可看成对 CCA 方法的一个拓展。

2.3.2 跨媒体索引

如何对海量高维数据进行高效的检索是一个有意义的研究方向。为了实现海量数据的高效检索, 通常需要设计合理的索引结构。传统的多维树形索引方法 (如 R 树、KD 树等) 在面对高维的多媒体数据时会遭遇“维度灾难”的问题, 从而大大影响索引的性能。为了避免传统树形索引遇到的“维度灾难”的问题, 需要设计更为合理的索引策略。

近年来, 高维索引策略中较为流行的方法是哈希索引。哈希索引的基本原理是通过设计哈希函数将高维数据映射到低维的汉明空间 (哈希编码), 使得高维空间内相近的数据具有相近的哈希编码。最终两个高维数据的相似度计算可以转化为高效的低维的哈希编码之间的距离计算, 从而避免了“维度灾难”^[49]。

现今的多媒体数据哈希索引的研究方向大致可以划分为如下三类:

1) 单一类型特征哈希索引: 单一类型特征哈希索引是指以单一类型的高维特征为输入的一类哈希算法的总称。

2) 多视图哈希索引: 对从数据中提取的不同类型特征和属性等进行索引的方法可归纳为多视图索引。

3) 跨媒体哈希索引: 对包含异构类型数据的信息资源进行哈希索引可称为跨媒体哈希索引。

这三个研究方面存在如下层层递进的联系: 从原始数据类型的角度看, 由单一类型数据(例如根据用户提交的样图查找与其相似图像)向多种不同类型的数据(根据用户提交的样图查找与其相关的音频或文档)演变; 从特征的角度看, 从单一类型特征向多类型异构特征演变, 如图 2 所示。

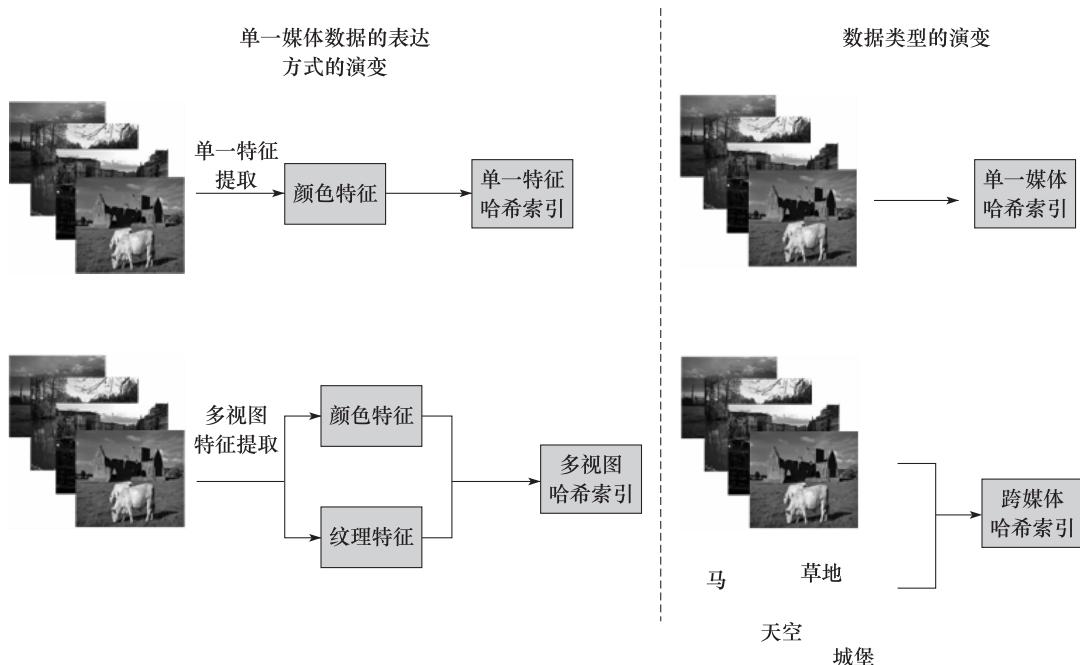


图 2 三类哈希索引演变关系图

跨媒体索引是传统高维索引的扩展, 它的基本思路是将不同模态的高维数据映射到一个统一表达的低维索引空间, 使得不同模态内具有相近“语义”的数据在索引空间内具有相近或相同的表达。

2010 年, Bronstein 等人提出了一种跨媒体哈希索引 (CMSSH) 并在不同类型的医疗图像(核磁共振, 断层扫描)以及 3D 模型数据上实现跨媒体检索^[105]。跨媒体哈希索引问题在 CMSSH 中首次被提出并给出如下解决思路: 对于两个维度不同的不同类型数据, CMSSH 目标是分别学习一组哈希函数, 使得对于存在已知语义关联关系的两个不同类型数据, 它们经过各自哈希函数映射后编码得到的哈希编码相似。CMSSH 通过标准的 Ada-Boost 方法依次对每一位哈希编码学习得到对应的哈希函数, 使得后一个哈希函数能够对前一个函数编码能力进行补充和加强。对每一个哈希函数学习的问题, 最终转变成一个 SVD 分解的问题进行求解。

作为最早的多视图哈希的算法，CMSSH 尽管还存在一些不足，例如：

- 1) 在学习哈希函数的时候只考虑了不同类型的数据的相似度，而没有考虑同一类型数据的相似度。
- 2) 只适用于两种数据类型之间的跨越，不能扩展到通用的多种数据类型。但是，它为后面的研究提供了一种跨媒体哈希索引的实现思路。

2011 年，Kumar 等人在谱哈希 (Spectral Hashing)^[61] 的基础上扩展为多模态的场景，提出了一种跨视图哈希索引方法 (Cross- View Hashing, CVH)^[106]，并解决了 CMSSH 算法存在的两点不足。CVH 在进行哈希函数的学习时，提出了将同种模态间 (intra-modality) 和不同模态间 (inter-modality) 的两种相似度都纳入考虑中来进行哈希函数学习。哈希函数旨在最小化相近数据对应的哈希编码的距离的同时最大化不相近数据的哈希编码的距离。值得一提的是，CMSSH 给出了两个特殊情况下 CVH 的退化版本，分别对应了流形学习中的经典方法局部保留映射 (Locality Preserving Projection, LPP) 和典型相关分析 (Canonical Correlation Analysis, CCA)。

2012 年，Zhen 等人^[107]提出了一种基于概率图模型学习哈希函数的算法——多模潜在二元嵌入 (Multimodal Latent Binary Embedding, MLBE)，从而实现跨媒体的检索。和以往利用谱图分割 (spectral graph partition) 来学习哈希函数的算法不同，MLBE 完全采用概率图模型来对学习得到哈希函数，并且同时考虑了相同模态和不同模态间的两种相似度。对于来自不同模态的数据，相应的哈希函数将其由各自的原始高维空间中映射到一个二值隐空间 (binary latent space) 中，并保证模态内的相似性得以保持。同时，模态间相似性的约束施加在相应的隐空间之间，使得有语义关联的不同模态的数据具有相近的哈希编码。

同年，Zhen 等人^[108]又提出了一种协同正则的哈希函数学习的方法 (Co-Regularized Hashing, CRH) 来解决跨媒体哈希索引的问题。CRH 旨在最小化多个模态的数据在哈希映射过程中的损失，从而学习得到每个模态的数据最优的哈希函数。由于目标函数的非凸性，求解时需要迭代求解，CRH 可通过凹凸过程 (Concave- Convex Procedure, CCP) 方法求解。

2.3.3 跨媒体排序

在跨媒体检索中，排序技术一直扮演着很重要的角色。当跨媒体索引将语义相关的跨媒体数据集合检索出来后，跨媒体排序的目的在于以更精细的粒度对集合里的跨媒体数据进行排序，使得与用户检索意图更为相关的文本（或图像）排序靠前。与传统排序学习不同的是，传统排序学习中已有的相似度模型是人为定义的，学习的目的是对不同相似度模型的加权组合；而在跨媒体检索中，由于特征表达的空间不一样，即跨媒体数据之间的“异构鸿沟” (heterogeneity-gap) 已成为跨媒体检索中一个障碍，不同模态数据之间的相似度并不能人为地去定义，跨媒体排序学习的目的在于用排序数据来学习两种不同模态之间的基于语义相似度的排序模型。

值得指出的是，与传统基于语义标注的跨媒体检索方法不同，跨媒体排序学习的训

练数据是比较容易获取得到的：排序训练样本只需要在已有的跨媒体搜索引擎中加入一点点额外开销以获取并存储用户的查询请求和点击流量数据即可^[109]。跨媒体排序方法的假设是，与查询词越相关的图片会被用户点击得越多。因此，它们都在寻求一个解决方案：如何更好地建立跨媒体数据之间的关联，使得对于同一查询词，被用户点击越多的图像（也就是越相关的图片）被排序得越靠前，且这种关联关系可以推广到未知的查询词和未知的检索文档上。

在排序学习中，Liu^[110]分析了已有的排序学习算法，根据对训练数据输入表达和损失函数的不同将这些排序学习算法分为三大类：基于数据点（Point-wise）的排序学习算法、基于有序对（Pair-wise）的排序学习算法和最新提出的基于列表（List-wise）的排序学习算法。研究结果表明，基于有序对和基于列表的排序学习算法在排序性能上往往较优于基于数据点的排序学习算法^[111]。

在文献[109]中，Joachims 基于在有序对空间中构建排序模型，训练了一个排序支持向量机（Ranking Support Vector Machine，RankSVM）用于学习各个排序特征函数的权重，其中训练数据来自于搜索引擎中用户点击数据的有序对。和传统支持向量机一样，RankSVM 的最小化目标函数是结构风险和训练数据中有序对的倒置（即经验风险）的折中。因此，RankSVM 也被认为是基于有序对的排序学习算法。

基于列表的排序学习算法以一个列表为基本的学习单元（基于列表的训练数据和损失函数）。由于现有排序性能准则均是基于列表的，因此基于列表的排序学习算法可以直接针对最终的排序性能进行优化，从而取得了比基于有序对的排序学习算法更好的排序结果。Cao 等人在文献[112]中第一次提出基于列表的排序学习算法 ListNet，并给出其损失函数和对应的评价标准；同时，根据不同的优化目标，他们分别训练了两个概率模型，一个是基于全排列的概率模型，另一个是基于 Top-K 的概率模型。

在文献[113]中，Yue 等人提出了另一种基于列表的排序学习算法 SVM-MAP，其同时基于结构性支持向量机（Structural SVM）框架^[114]。该方法的损失函数被定义为基于列表的平均查准率（Mean Average Precision，MAP），并针对优化目标提出了启发式的优化算法，其直接对排序中常用的评价准则进行优化从而更易于取得较好的排序性能（尤其是对平均查准率而言）。与此类似的还有文献[115]对另外两个常用排序评价准则 MRR（Mean Reciprocal Rank）和 NDCG（Normalized Discounted Cumulative Gain）直接进行优化。

需要强调的是，跨媒体排序学习和传统排序学习不一样的地方：传统排序往往是学习一些人为构造特征（如 TF-IDF 相似度，PageRank 值）的权重，其中的相似度是人为预先定义好的，如 TF-IDF 相似度被定义在余弦相似度上；而对于跨媒体排序，由于异构模态数据之间的语义鸿沟，学习的往往是一个跨媒体映射函数，使得不同模态数据被映射到同一个特征空间，进而再进行语义相关性排序。

图像检索中被动攻击模型（Passive Aggressive Model for Image Retrieval，PAMIR）算法^[116]首次提出用排序学习的方法对跨媒体图像进行排序。PAMIR 将跨媒体排序问题用一个类似于 RankSVM 的模型进行重新表达（因此它是个基于有序对的排序算法），并拓

展了用于求解传统支持向量机的 Passive Aggressive 算法^[117]以对海量的训练样本进行训练。同时，PAMIR 还可以结合核技巧（kernel trick）对图像的相似度进行不同的建模，再而选择在排序性能上表现最优的核函数。

在文献[118]中，Zhai 等人提出了最近邻异构相似度测量（Heterogeneous Similarity Measure with Nearest Neighbors，HSNN）方法用于计算不同模态媒体间的相似度以进行跨媒体排序。此异构的相似度是根据计算两个多媒体数据属于同一个语义类别的概率得来，而概率是根据分析异构每个多媒体数据的最近邻得来。在 HSNN 中，组合多个弱排序器并通过 AdaRank 来最终学习得到排序模型。

跨媒体排序中，文本和图像通常都用单词词袋（bag of words）模型或者视觉词袋（bag of visual words）模型来进行特征表达建模，因而往往被表达为高维特征向量。然而，高维向量空间表示引入了两个传统的问题，即“一词多义”和“一义多词”的问题。在传统文本检索中，为了捕获数据所包含的隐语义，将文本中的单词嵌入到一个低维隐空间是一种经典的解决方法，例如隐语义分析（Latent Semantic Analysis，LSA）^[119]以及概率隐语义分析（Probabilistic LSA）^[120]。类似的思想被引入到跨媒体排序中，如 SSI 方法^[99]和 PSI 方法^[100]。有监督的语义索引（Supervised Semantic Indexing，SSI）方法定义了一类的低秩线性模型来捕获模态内以及模态间的“一词多义”和“一义多词”。与基于无监督学习的 LSI 不同的是，SSI 基于有监督的排序学习，使用了有序对作为训练样本。多项式语义索引（Polynomial Semantic Indexing，PSI）方法则扩展了 SSI，使之不仅能捕获线性关联，还能捕获更高阶的关联。

结构学习（Structured Learning）是机器学习中的一个热点问题，其研究的内容是学习输出一个复杂的结构，如列表和树状结构，而不是传统机器学习研究的分类问题（输出为类别）或者回归问题（输出是实数）。排序学习的输出是一个列表，其自然符合结构学习的输出，因此近年来基于结构学习，特别是结构支持向量机的排序学习成为热点，如文献[113]和文献[115]。文献[121]将度量学习以排序学习的视角来进行研究：他们基于结构性支持向量机提出了一个通用的模型来学习得到单模态数据内的度量；通过排序学习的方法，该学习得到的度量可以使得给定查询文档，相关的文档会基于某种排序评价准则被排序在前面。与此类似的方法还有文献[122]，它与文献[121]不同的是其不仅考虑了输出端的稀疏性质，还考虑了输入端的稀疏性质。

3 国内研究进展

3.1 视频事件识别与标注

在我国，视频事件识别与标注方面的研究可以说是个新领域，近几年国内才有很多高校和研究机构开展起来。比如中科院自动化所模式识别国家重点实验室、北京大学视

觉与听觉信息处理国家重点实验室、微软亚洲研究院视觉小组和清华大学等对视频事件识别与标注进行了相关的研究。另外，当前国内一些权威期刊（如计算机学报、软件学报、中国图像图形学报、自动化学报等）将视频事件识别与标注研究作为主题内容之一，为该领域的研究人员提供了更多的交流机会。

第一届全国智能视觉学术会议于 2002 年 5 月 25~26 日在北京市西郊宾馆成功举行，会议由国家自然科学基金委员会、中国自动化学会、中国科学院自动化研究所等单位举办，并由中国科学院自动化研究所模式识别国家重点实验室承办。来自全国各地高校、科研院所、公安部、军队、企事业单位的科研人员、学者就视觉内容语义理解进行了广泛而又深入的讨论。2004 年 6 月，中国图像图形学会主办了第一届机器视觉国际展览会，在产业界和学术界产生了广泛的影响。2006 年 3 月 18~19 日，北京图像图形学会和公安部第一研究所在北京举行了首届中国城市公共安全视觉听觉科技研讨会，就视频监控系统中的运动目标识别、人群聚集与突发行模式分析、生物特征识别与认证等技术在国内的研发和应用现状进行了研讨，以提升 2008 年北京奥运公共安全对策。

中国科学院自动化研究所模式识别国家重点实验室对交通场景的视觉（基于三维线性模型定位、基于扩展卡尔曼滤波器的车辆跟踪算法）、人的运动视觉监控（基于步态的远距离身份识别）、行为模式识别（提出了对目标运动轨迹和行为特征的学习的模糊自组织神经网络学习算法）和视频标注（提出了多实例学习算法）进行了较为深入的研究，取得了一定的成果^[123~127]。开发的原型系统，在不需要人的干预或者只需要很少干预的情况下，通过对摄像机拍录的视频序列进行分析，实现车辆的定位、识别和跟踪，并在此基础上分析和判断车辆的行为，对车辆的行为给出语义描述，并对视频进行语义压缩。国内还有一些高校也进行了这方面的研究，如上海交通大学、北京航空航天大学、北京理工大学等。此外，国内在面向体育视频^[128~129]、新闻视频^[130]等视频事件的处理与识别中也开展了一些有成效的研究工作。

而国内智能视频事件识别与标注方面的产品只是简单地检测出有无运动目标，而没有对运动目标做任何分析。通过对国内银行、交通等视频分析系统应用单位的了解，目前对视频内容做进一步智能分析，从而辅助管理决策的系统几乎没有。在进行视频事件检索时也只靠时间线索人工查找，智能程度不高。但是，对智能视频产品的研发和需求关注的单位已有不少，并有越来越多的实体加入这个行列。特别是在智能交通监控方面已有一些低层次的智能功能，如流量统计、违章检测等产品进入实际应用。

以上内容都说明了视频事件识别与标注在国内的起步较晚，应用需求近几年才逐步显现，相关的视频和图像处理与分析技术的研究正在逐步受到学者、政府和商家的关注和重视。

3.2 高维索引结构

国内在高维索引结构方向的研究也取得了相应的研究成果，分别表现在基于树结构的高维索引、基于哈希的高维索引以及基于视觉单词的倒排索引，与国际上的研究相互

呼应。

3.2.1 基于树结构的高维索引

Zhuang 等人^[131]提出了基于对称编码的混合距离树 (Symmetrical Encoding-based Hybrid Distance Tree, EHD-Tree) 高维索引结构, 将 k-NN 近邻搜索从高维空间转换到一维空间。EHD-Tree 首先利用 K-Means 方法对数据集进行聚类; 然后利用基于双距离驱动 (dual-distance-driven) 的编码机制来获得每个特征点的 UID; 最后结合特征点的 UID 和中心距离计算得到它的索引关键字, 并利用 B+ 树对索引关键字构建索引。

相对于用单一的特征描述视频或图像, 用多种特征能更准确地描述它们, 进而呈现给用户更为满意的检索结果。为了对多特征构建有效的高维索引, He 等人^[132]提出了一种面向权值的多特征索引树 (Multi-Feature Index Tree, MFI-Tree)。MFI-Tree 是度量空间的一种动态索引结构, 通过与之匹配的 ADD-kNN 检索算法可以实现权值可变的多特征检索。

3.2.2 基于哈希的高维索引

在 E2LSH 相关索引方法的研究中, 文献 [133] 提出了两种查询扩展策略来改进 E2LSH: 内部扩展和交互扩展。内部扩展是获得更多与查询特征点相似的特征点并用作新的查询; 交互扩展是挖掘查询结果图像中与查询图像不同的特征点作为新的查询。

针对 LSH 相关索引方法对高位稀疏的 BOF 图像描述符构建索引的问题, 文献 [134] 引入图像分解模型 (Image Decomposition Model, IDM), 将 BOF 图像描述符分解为 3 个向量: 低维图像主题信息分布向量、图像细节信息分布向量和图像背景信息分布向量。LSH^[51]被用于对低维致密的图像主题信息分布向量构建索引结构。

对于哈希编码方法的研究, Shao 等人提出了稀疏谱哈希 (Sparse Spectral Hashing, SSH)^[135], 将稀疏主成分分析 (Sparse PCA) 和增强相似敏感哈希 (Boosting Similarity Sensitive Hashing, Boosting SSH) 引入 SH。通过限制非零特征的数量, SSH 将二进制编码问题公式转化为拉普拉斯图的特征向量子集阈值。另一个以 SH 为基础的研究工作是超图谱哈希 (Hyper-graph Spectral Hashing, HSH)^[136]。HSH 针对图像的社交内容设计哈希函数, 实现大规模社交图像的快速检索。Zhou 等人^[137]提出了一种标量量化方法, 其根据每个特征向量的中位值进行量化和编码, 无需任何训练过程。

3.2.3 基于视觉单词的倒排索引

由于 BOF 描述符对不同视角、光照条件和部分遮挡图像的区分性不高, 因此一些研究者在构造视觉词汇时引入一些额外信息来提高视觉词汇的准确性。然而, 这些额外的人工信息会显著地增加监督式训练的开销。为了解决这种问题, Kuo 等人^[138]提出了无监督的辅助视觉单词 (Auxiliary Visual Words, AVW) 发现机制。AVW 较传统 BOF 显著地改进了在不同光照和拍摄角度等条件下图像的检索性能, 能检索到更加多样化的结果。

为了减少计算 BOF 过程中的总体量化误差, Cai 等人^[139]提出了约束性特征点量化。

如果一个特征点到所有视觉单词的欧式距离都大于一个预先设定的阈值，则约束性特征点量化就认为该特征点为异常点，不分配到任何视觉单词。针对硬分配方法存在的问题，文献[140]提出了球形软分配将特征点自适应地分配到近邻视觉单词。每个视觉单词对应一个聚类的中心，每个聚类定义为一个超球体。超球体半径用聚类中最远特征点与聚类中心之间的距离表示。根据特征点的空间位置是否位于超球体内，特征点被分配到相应的视觉单词。

在对描述符量化和压缩的研究方面，Chen 等人^[141]提出了残差向量量化（Residual Vector Quantization, RVQ）方法，利用多个低复杂度的量化器来减少量化误差，使得对结构化描述符和非结构化描述符都能进行高效的量化。RVQ 由多层量化器构成，每层量化器利用 K-Means 方法训练聚类中心，并且每层量化器的误差用于训练下一层的量化器。RVQ 的量化层数越多，其编码的准确性就越高。由于 RVQ 使用全局量化器，因此 RVQ 在时间效率上要略逊于 PQ。构造索引结构时，RVQ 用前 L 层量化器构建倒排索引的外层粗量化器。如果每层量化器包含 k 个聚类中心，那么就可以构造具有 k^L 个倒排列表的倒排索引结构。每个特征描述符根据其前 L 层的编码插入对应的倒排列表，倒排列表中结点存储的信息是特征描述符 ID 及其整体编码。

3.3 跨媒体检索理论与技术

国内在跨媒体检索这一方向的研究与国际上相互呼应，在跨媒体度量、跨媒体索引和跨媒体排序三个领域也取得了一些研究成果。

3.3.1 跨媒体度量

在典型相关性分析基础上，Zhuang 等人^[142]提出了一种基于改进的稀疏典型相关性分析（Sparse CCA）方法，并将其应用于跨媒体度量学习领域。和文献[104]中提出的针对“原始 - 对偶”数据问题的 Sparse CCA 略有不同，该算法更为通用、简洁且易于实现。

在基于字典学习的跨媒体度量学习方法中，全局共有系数空间的假设会导致不同模态数据在字典学习中重建系数被迫相同，从而限制了算法本身。因此，如何将字典学习更加自然应用于多模态数据，已成为一个热点问题。如文献[143]提出了一种基于结构信息的监督耦合字典学习方法，将其应用于跨媒体检索。这一方法通过耦合字典学习来求取不同类型数据的字典及每一数据所对应的稀疏系数，同时求取一个映射函数来实现不同类型数据的映射。

互联网跨媒体数据检索的一个核心问题是数据间的关联进行建模，常用的模型有图和矩阵等。然而，由于互联网中异构数据间的关联通常是复杂高阶的，传统的建模方法并不能够充分表达这些关联。例如，在图像社区中，当某个用户对某幅图像增加一个标签时，图像、用户、标签这三个对象形成了三阶的复杂关联，这样的关系很难基于图或矩阵进行表达。与传统的图表达不同，在超图（hyper graph）中，每一条超边可以连

接任意多个图中的顶点，从而将存在高阶关联的数据进行关联。因而，超图非常适合对跨媒体数据之间的关联进行建模，并在近年来成功地应用到了跨媒体检索中。在文献[144]中，Yu等人通过引入超图表达将图像的属性信息与图像的底层视觉特征进行建模，并且应用在了图像相似性的检索中。此外，作者还提出了基于超图的图像排序算法，达到了更好的图像检索效果。在文献[145]中，Hong等人提出了语义关联超图，对图像及其关联进行建模。在文献[146]中，Hong等人将超图应用在了多视图数据的图像块对齐工作中，通过将图像块看做超图中的超边并将相邻块的相似度看做超边权重，作者定义了多视图拉普拉斯矩阵并基于此实现了图像块对齐。

3.3.2 跨媒体索引

在文献[147]中，Liu等人利用统一超图（unified hypergraph）对Flickr网站中的图像及其上下文进行建模。在该模型中，图像和图像的上下文属性（如用户、标签、地理位置等）都被统一地看做超图中的结点，而异构数据间的高阶复杂关联则可以通过连接这些顶点的超边来自然表达。在此基础上，Liu等人将谱哈希算法扩展到超图表达中，并借鉴自学习哈希（Self-taught Hashing）的思想将哈希算法应用到了新数据的计算。对来自不同模态的数据，分别利用各自模态学习到的哈希函数，可以映射为统一的哈希编码，从而进行高效的跨媒体检索。

跨媒体检索不局限于跨越不同模态的数据检索，还将来自不同模态数据作为查询进行统一检索也输入跨媒体查询的范畴。Liu等人^[148]提出了一种针对图像+关键词的混合查询的哈希索引。在哈希函数学习的同时学习每个哈希函数和关键词的关联关系，在给出图像+关键词的查询条件时就可以先利用关键词找出最相关的哈希函数，并用这些哈希函数对图像进行编码，查找最相关的结果。

3.3.3 跨媒体排序

Lu等人在文献[149]中提出以排序学习视角的潜在语义跨媒体排序（Latent Semantic Cross-Modal Ranking，LSCMR）方法：LSCMR通过两个不同的映射函数分别将两种不同模态的数据映射到同一个特征空间，使之可以在同一个特征空间里进行排序。和其他方法不同的是，LSCMR基于最大化排序间隔的思想来学习得到这两个线性映射函数，因此能直接对最终的排序结果进行优化，从而提高了排序性能。

4 国内外研究进展比较

4.1 视频事件识别与标注

视频事件识别与标注是当前国内外非常重要的研究领域，拥有强大的应用推动力。

在国内，从事视频事件识别与标注的单位主要有浙江大学、清华大学、中国科学院计算技术研究所、中国科学院自动化研究所、微软亚洲研究院等。相关课题受到国家的高度重视，得到了国家高技术研究发展计划（863计划）、国家重点基础研究发展计划（973计划）和国家自然科学基金等的资助。与前几年相比，国内在这方面的理论研究和产业应用水平逐年提升，在国际学术刊物（International Journal of Computer Vision, IEEE Trans on Pattern Analysis and Machine Intelligence）和重要国际会议（International Conference on Computer Vision, IEEE Conference on Computer Vision and Pattern Recognition）上发表了越来越多的论文，申请了更多的国际专利，并且与国际同行的合作也更加紧密。

在国外，如在美国，随着网络技术的发展和大量视频的产生，政府全力推进相关技术的发展，开展了大量关于视频事件识别与标注项目的研究。

2001年，国家标准与技术协会（National Institution of Standards and Technology, NIST）开始组织TREC视频检索评测（TREC Video Retrieval Evaluation, TRECVID）评测，并且从2002年开始将视频标注列为一项单独的任务：高级特征提取任务（High-Level Feature Extraction Task），这标志着大家对这个领域开始了系统性的研究。在这之后的几年，参加TRECVID的Group数目呈递增趋势，同时大家对视频标注的研究也越来越深入，TRECVID选取的评价概念的复杂度越来越高，同时覆盖范围也越来越广，稀有概念（即出现频率相对较低的概念，通常在训练集中正例很少）有增加的趋势。然而，大家在这个任务上的结果仍然是越来越好，标志着这个领域中研究的快速发展。值得一提的是，国内很多高校和研究机构，如清华大学、复旦大学、中国科学技术大学以及微软亚洲研究院等，很早就开始关注这个领域，积极参加TRECVID，并且在各项任务评比中取得了很好的结果。与国际研究水平相比较，国内的这些高校和研究机构毫不逊色，甚至在有些技术的研究上处于领先地位。尽管我国在视频事件识别与标注方面取得了一些成果，但在总体上与美国、法国、德国等国际先进国家相比，无论是在研究还是在产业应用上，还是有一定的差距。

总的来说，视频事件识别与标注技术仍处于理论研究阶段。由于一般的识别方法需要处理庞大的数据量，计算代价非常高，而且这些方法多数会极大地受到视频数据复杂性等因素的影响，因此，要做到灵活、实时的应用仍存在许多难以解决的问题。目前，视频事件识别与标注的研究还是集中于视频中目标的跟踪、特征提取、简单事件识别等问题，虽然近年来利用机器学习、人工神经网络等工具构建视频理解的统计模型的研究有了一定的进展，但依据事件识别与标注进行视频的有效管理和理解还是处于初级阶段。而且，对于视频事件识别与标注来说，目前国际上不仅没有通用的、权威的大规模的数据库，算法的评估与实验都是在各自小样本数据库上模拟进行，而且数据库在构建时限定了许多约束条件，比如：静止的摄像机，相对比较简单的背景，简单的事件类别等。

4.2 高维索引结构

作为基于内容的大规模图像检索研究的关键技术之一，基于内容的高维索引结构是

当前国内外在这一研究领域的研究重点和热点，拥有广阔的应用前景。在国内，开展高维索引结构的研究单位主要有华中科技大学、浙江大学、清华大学、中国科学院计算技术研究所、微软亚洲研究院以及 IBM 中国研究院等，企业单位主要有百度、腾讯、360 和搜狗等，并且都已推出了基于内容的图像搜索服务。近年来，随着国际合作越来越紧密，国内的研究者在国际重要学术期刊以及国际顶级学术会议上发表了越来越多的高水平论文，已与国外的研究处于同步阶段。目前，华中科技大学数字媒体处理与检索实验室正在开展这一领域的一项国家自然科学基金研究：面向海量图像搜索的高维索引结构与快速检索算法研究。这一项目拟对图像搜索引擎面临的如何在有限的计算机内存中构建高维索引问题进行深入研究。具体研究内容有：构造具有较高区分度的图像描述符、对图像描述符进行量化与编码以满足大规模图像描述符在计算机内存中的存储需求；对大规模描述符构造高效的索引结构以及快速自适应检索方法。

在国外，率先关注高维索引结构这一领域的研究单位主要有：美国麻省理工学院、哥伦比亚大学、伊利诺伊大学以及英国牛津大学等，他们也各自研发出了一些基于内容的图像检索系统。企业单位包括微软以及谷歌等，并且都已在搜索引擎中推出了基于内容的图像搜索服务。虽然国外研究者在高维索引结构的研究上比国内关注得更早，但随着高维索引结构成为国内的研究热点以及更紧密的国际合作，目前国内的研究水平和研究成果已接近国外。近年来，法国著名的国家信息与自动化研究所 INRIA 在这一领域的研究非常活跃，开展了一些关于大规模图像索引与检索的研究项目：“ICOSH-D”对视频内容的高维描述符构建索引和进行可扩展性压缩；“Quaero”对多媒体（图像、视频等）和多语言文档的自动分析与分类，其中就包括对图像和视频检索的研究。

总体而言，国内外对构建高维索引结构的方法：基于树结构的高维索引、基于哈希的高维索引和基于视觉单词的倒排索引都投入很大精力进行研究。相对于基于树结构的高维索引，基于哈希的高维索引和基于视觉单词的倒排索引受到国内外研究者更多的关注。在后两种高维索引方法的研究中，相对来讲，国内研究者更侧重于基于哈希的高维索引，尤其是哈希编码方法成为最近的热门研究方向。虽然目前国内外的研究取得了很多研究成果以及各大搜索引擎都推出了基于样例的图片搜索，但距离成熟阶段，尤其是针对海量图像的索引和检索，尚有一段距离，仍然存在许多问题亟待解决。在基于哈希的高维索引方面，其通过一系列的哈希函数，将欧式空间中的相近的描述符映射到汉明空间中相同或者相近的二进制编码。然而，二进制编码的汉明距离的相似度区分能力受限于编码长度，因此，在汉明距离有限区分度的前提下，如何设计哈希映射函数，进一步提高检索准确率值得进一步研究。在基于视觉单词的倒排索引方面，目前的倒排索引的构建过程是通过对一个样本集进行训练完成的。然而，互联网上的数据每时每刻都在快速增长，而已构建好的索引结构只能反映当时的数据分布，而不能实时地反映当前得到的数据。因此，如何设计对动态插入或者删除具有健壮性的倒排索引值得进一步探索。

4.3 跨媒体检索理论与技术

2012年，我国科技部正式启动“面向公共安全的跨媒体计算理论与方法（项目编号：2012CB316400）”973项目，这一项目拟对跨媒体数据统一表示与建模、跨媒体数据关联推理与深度挖掘和跨媒体数据综合搜索和内容合成等三个方面进行重点研究，面向社会公共安全这一重大需求，对跨媒体数据所隐含的社会属性和特定事件的动态演变进行分析，挖掘和预测特定社会事件，进而以跨媒体形式对其起源、现状及发展进行全过程呈现。这一项目由浙江大学牵头，北京交通大学、西安交通大学、中国科学院计算技术研究所、中国科学院西安光学精密机械研究所等单位参加。

对不同类型媒体数据的综合利用及搜索也成为近年来国际上的一个热点趋势。2010年1月，《自然》(Nature)杂志邀请20位搜索、医药、能源、天文学、化学、人类健康、激光、石油、生态学的专家和政策制定者对下一个十年科技发展进行展望，形成的报告《2020Vision》发表于《自然》杂志2010年第1期。谷歌的研究主管Peter Norvig博士应邀写了对未来多媒体计算与搜索技术的预测：文本、图像、视频将更加紧密混合在一起（跨媒体），亲朋好友之间交往的社会化属性等信息将被记录和分析（社会网络），传感器网络在定位、医疗等方面应用更加普遍（物联网）。搜索请求不再是文本，而是语音和对大脑信号的理解（语音识别与人脑感知）；搜索结果不再是罗列网页，而是以图或表来表示的更为形象的综合性知识（数据挖掘）。国际期刊《Journal of Multimedia》于2012年2月推出了名为“Connected Multimedia”的特刊，讨论了多模态数据主题建模、基于多模态数据分类的路线推荐等问题，《Journal of Zhejiang University-Science C》也于2012年12月推出了名为“Societally connected multimedia across cultures”的特刊，对类似问题进行了深入讨论。同时，国外著名研究机构INRIA成立了关于多模态数据的内容检索项目“I-Search”，对来自各种媒体多模态的数据（文本、图像、视频、音频、3D模型等）进行统一检索。

总体而言，国内外对跨媒体检索理论和方法均抱以很大热情并投入很大精力进行研究，国内外在这一方向的研究处于同步阶段。但是，在跨媒体度量学习、索引和排序等方面尚存在巨大挑战需要克服。在跨媒体度量方面，目前的手段主要有典型相关性分析、跨模态字典学习和概率主题建模，它们分别从相关度学习、共享结构挖掘和稀疏重建等角度来解决跨媒体度量学习。因此，如何将三者结合起来，构建基于概率隐模型的跨媒体度量学习方法，是一个值得深入研究的问题。在跨媒体索引方面，哈希的方法是最近的热门研究方向，如何将来自不同模态的数据映射为统一的简短的哈希编码，从而实现高效的跨媒体检索是跨媒体哈希索引的目标。因此，如何对数据进行哈希编码的同时，保持模态内的相似性和模态间的相关性是至关重要的因素，值得深入探索。在跨媒体排序方面，目前的工作主要是基于相关性的排序，即输入一个模态的查询数据，返回和它相关的另一个模态的有序的反馈结果。此外，在此基础上还有一些值得进一步深化的研究问题，例如，支持更为复杂的查询模式（如复合模态的查询输入），或者返回多个模态的相关结果，并进行统一排序等。

5 发展趋势与展望

5.1 视频事件识别与标注

尽管视频事件识别与标注技术正展示出广阔的应用前景，众多的科研人员也已经投入了大量的精力进行相关方面的研究并产生了许多有效的算法，但是很多工作都是一些探索性的研究，它们仅限于对一些简单场景或严格限制的场景中简单的事件进行分析与理解，对于相对复杂的场景和行为，现有的研究成果还难以处理。目前，各自算法的测试和评估都是在少量的几个数据库上完成的，缺乏普遍性和说服力，很多技术迫切需要完善。影响其进一步发展的主要原因在于以下几个方面：

- 1) 事件表述。事件表述是事件分析的基础，因此很多研究工作着力于有效的特征提取和融合。但是，目前的特征提取和融合方法对于复杂场景下的事件分析仍然存在不足。所以，如何构造一种有效、描述准确度高并具有良好鲁棒性的事件表述也是事件分析研究中的难点之一。
- 2) 事件模型。就事件模型而言，不论产生式模型还是判别式模型，目前的模型结构都是研究者根据先验知识事先定义好的，很少自动获取事件的结构，这难免对事件本质产生一定程度的束缚，不能够完全地、真实地表达事件的本质。而且我们也知道模型越简单越有效，但是真实世界中的事件总是错综复杂的，到底哪一种模型才是有效的、可以泛化的，值得我们进行深入的研究探讨。
- 3) 高层语义理解问题。目前对事件分析的研究还是停留在约束条件下简单事件的分析，其在实际场景中的应用还有很大距离。事件分析的最终目的是要电脑分析和理解人的行为，这种理解不仅要识别的基本行为，还要能根据识别的结果做出反应。
- 4) 鲁棒性、实时性和精度。事件描述与事件的语义存在鸿沟，因此需要找到合适的学习方法来鲁棒地解决在事件描述不好的情况下的事件识别与标注问题。另外，对于不同的应用场合，事件行为识别算法需要兼顾鲁棒性、实时性和精度等多方面因素。通常情况下，离线的行为识别方法精度会相对高一些，然而计算量会比较大，不适合实时应用的场合；在线的行为识别方法，实时性比较好，然而往往精度偏低；有监督的行为识别方法，识别精度会相对高些，然而比较依赖训练样本；无监督的行为识别方法，精度则普遍较低；相比较，半监督的学习方法既能减少训练样本数量又能取得相对较好的识别效果。因此，针对不同的应用场合，如何实现鲁棒性、实时性和高精度等因素的折中，也是事件分析技术中亟待解决的问题之一。

虽然目前基于机器学习的视频识别与标注已经取得了很大进展，但是其距离实用化仍有一定距离，要真正实现大规模数据集高效而精确的事件识别与标注，仍需进行进一步的研究。这里简单阐述五种可能的改进和扩展。

1) 挖掘事件相关性的研究: 视频片段间的事件是具有一定关系的, 如果能够对这种关系加以挖掘, 必然能够提高标注的准确性。关于这个领域已有一系列的研究, 如文献[150~152]中的工作。然而, 目前的研究表明, 这些工作与传统单事件学习方法相比, 通常是仅有部分事件的标注结果会提高, 但是也会降低另一部分事件的标注性能^[152]。如何设计更稳定的多事件相关学习算法, 仍然值得进一步研究。

2) 自适应特征提取: 不同的事件概念需要通过不同的特征进行区分, 因此为了能对整个事件概念集合进行分析, 需要尽可能多地提取各种特征参数, 如颜色、边缘、纹理以及运动特征等。问题在于: 往往仅有部分特征与待标注概念相关, 冗余的部分不仅增加了计算负荷, 而且可能会对最终标注产生负面的影响。因此, 需要研究针对特定的事件选择合适的特征参数或映射准则的方法。这部分可以借助深度学习方法进行改进。

3) 鲁棒的相似性度量: 目前的学习算法中采用的相似性度量, 大部分都是直接根据样本间的距离定义的。然而, 相似度不仅仅与距离有关, 也应考虑到样本周围的分布的相似性。在文献[153]中, 我们已经提出了一种邻域相似度的概念, 其可以同时考虑到样本间的距离和周围样本分布的差异, 性能比传统的基于距离的相似度要好。然而, 该工作还可以继续扩展, 进一步考虑邻域内标号的分布, 这样设计出的新的相似性度量可以提升学习算法的区分性。

4) 有效的模型训练方法: 当前大多数方法为有监督的方法, 即采用人工标记的方式来训练行为模型。这种方法需要人工标注获取各类行为的样本, 因此将会浪费大量的人力。近年来, 人们提出了自动获取样本, 自动建立行为模型的方法。因此, 无监督或半监督的方法是重要的研究方向。

5) 复杂场景下复杂事件识别与标注: 目前许多工作主要是对简单场景下的简单事件进行识别和标注, 并且取得了较好的结果; 而对于复杂场景下的事件识别和标注, 则效率很低。现实生活中的很多场景都是比较复杂的, 甚至是时常变化的, 而且场景中的行为也不单单是一些基本的动作, 因此复杂场景下复杂事件的识别和标注问题是一个值得进一步研究的问题, 估计在很长一段时间内, 它将是一个很难解决的问题。

5.2 高维索引结构

尽管目前的高维索引技术采用如特征降维或者描述符特征分解等方法来应对高维描述符特征的“维度灾难”问题, 但暂时还没有彻底解决这个问题。因此, 我们仍然需要进一步研究一个高效和健壮的高维索引技术。

基于视觉单词的倒排索引利用描述符压缩和编码, 使得海量的特征描述符可以存在计算机内存。然而, 目前这类索引方法通常使用的是图像低层特征, 而检索得到的低层特征距离相近的图像对用户来说, 不一定是视觉感官上相似的。因而, 如何将图像低层特征与一些语义信息相结合, 构造更能准确描述图像内容的描述符可能需要进一步关注和研究。

由于一个索引列表对应一个聚类, 因此在构建索引之前需要样本集训练聚类中心。

然而，这些中心只能反映当前数据集的分布情况。当插入一批新数据或者从数据集中删除一批数据，由于数据分布已发生改变，因此已训练的聚类中心就不能反映当前的数据分布，进而影响检索准确率。针对这种问题，目前的索引方法采用重构索引来解决。但是，当数据集的规模很大时，重构索引需要花费大量的时间。因而，设计一个支持动态插入和删除的倒排索引仍然需要进一步的研究。

随着移动互联网和智能手机的出现与发展，人们之间的交流越来越依赖于智能手机。手机存储空间越来越大，此外，智能手机的处理器性能越来越强大，如今已有 4 核处理器甚至 8 核处理器。这意味着个人智能手机中的图像等多媒体信息也越来越多。因此，面对智能手机和移动互联网中的大规模图像，如何使人们可以从其中快速地检索到相似并感兴趣的图像或许会成为一个新的研究热点。

5.3 跨媒体检索理论与技术

在谈及跨媒体这一概念时，我们需要分清其与多媒体概念，它们有以下两个方面不同：

1) 跨媒体数据强调不同类型数据来源于不同渠道 (sources 或 domains)，在分析这些数据时要关注三个方面的“跨越” (cross)，即跨越模态 (cross-modality)、跨越来源 (cross-source) 和跨越网络空间和实体世界 (cross cyberspace to reality)。跨越模态指在检索过程中可以利用不同类型的异构底层特征 (如文本特征和视觉特征等) 来提升检索性能；跨越来源指不同来源数据 (如从 Flickr 和 Twitter 获取的数据) 从不同侧面来表达语义；跨越网络空间和实体世界指跨媒体数据通过人类交互行为来反映网络空间和实体世界的相互耦合、相互补充。

2) 在跨媒体分析过程中我们需要充分利用从不同来源所获取的不同类型数据来挖掘其蕴含的知识，如文本和图像数据中所内嵌的隐性关联、不同类型数据合成在一起作为摘要来表达多源头数据的内容侧面 (aspect) 等。

因此，在跨媒体研究中，如何去发现和挖掘从不同源头所获取的不同类型数据之间关联性是研究人员面临的巨大挑战，正如加州大学伯克利分校统计系前任系主任 Terry Speed 教授于 2011 年 12 月在《Science》杂志发表题为 “A Correlation for the 21st Century”的论文所提出的“21 世纪是关联性学习的时代”这一观点，在跨媒体检索中，从庞大数据集中发现数据之间所潜在的重要关系变得十分重要。

6 结束语

对媒体信息的索引和检索的终极目标是实现类似文本搜索的搜索引擎，通过文本可以找到任何想要的多媒体内容。在视频智能管理和智能视觉监控需求的推动下，视频事件识别与标注分析技术蓬勃发展，虽然有很多的系统实现了对简单事件的分析与理解，

但是视频中复杂场景下复杂的事件分析一直没有得到很好解决。随着数据量的增加，现有的高维索引结构仍然无法彻底克服“维度灾难”，无法达到文本倒排索引的效果，更无法满足搜索引擎的实际需求。跨媒体搜索为跨越“语义鸿沟”提供了很好的解决思路并已经取得较好效果，但远远无法满足基于内容搜索的需要。随着大数据时代的到来，多媒体信息的索引与检索的需求将日益迫切，面对众多挑战的同时，该研究领域将迎来前所未有的重大机遇，将会有越来越多的研究者关注该领域，也必将产生越来越多可以实用的研究成果。

参考文献

- [1] Aggarwal J, Cai Q. Human Motion Analysis: A Review[A]. Proceeding of IEEE Nonrigid and Articulated Motion Workshop[C]. 1997: 90-102.
- [2] Gavrila D. The Visual Analysis of Human Movement: A Survey[J]. Computer Vision and Image Understanding, 1999, 73(1): 82-98.
- [3] Wang J J, Singh S. Video Analysis of Human Dynamics-A Survey[J]. Real-Time Imaging, 2003, 9(5): 321-346.
- [4] Wang L, Hu W, Tan T. Recent Developments in Human Motion Analysis [J]. Pattern Recognition, 2003, 36: 585-601.
- [5] Hu W, Tan T, Wang L, et al. A Survey on Visual Surveillance of Object Motion and Behaviors[J]. IEEE Systems, Man, and Cybernetics Society, 2004, 349(3): 334-352.
- [6] Ronald P. Vision-Based Human Motion Analysis: An Overview[J]. Computer Vision and Image Understanding, 2007, 108(1): 4-18.
- [7] Turaga P, Chellappa R, Subrahmanian V S, et al. Machine Recognition of Human Activities: A Survey [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2008, 18(11): 1473-1488.
- [8] Poppe R. A Survey on Vision-based Human Action Recognition[J]. Image and Vision Computing, 2010, 28(6): 976-990.
- [9] Bobick A, Davis J. The Recognition of Human Movement Using Temporal Templates[J]. IEEE Transaction Pattern Analysis and Machine Intelligence, 2001, 23(3): 257-267.
- [10] Blank M, Gorelick L, Shechtman E, et al. Actions as Space-Time Shapes[A]. Proceeding of ICCV[C]. 2005, 1395-1402.
- [11] Ben A, Wang Z, Pandit P, Rajaram S. Human Activity Recognition Using Multidimensional Indexing[J]. IEEE Transaction Pattern Analysis and Machine Intelligence, 2002, 24(8): 1091-1104.
- [12] Wada T and Matsuyama T. Multiobject Behavior Recognition by Eventdriven Selective Attention Method[J]. IEEE Transaction Pattern Analysis and Machine Intelligence, 2000, 22(8): 873-887.
- [13] Niebles C, Wang H, L Fei-Fei. Unsupervised Learning of Human Action Categories Using Spatial-Temporal Words[J]. International Journal of Computer Vision, 2008, 79(3): 299-318.
- [14] Wang Y, Jiang H, Drew M S, et al. Unsupervised Discovery of Action Classes[A]. Proceeding of CVPR [C]. 2006, 1654-1661.
- [15] Dedeoglu Y, Toreyin B, Gudukbay U, et al. Silhouette-Based Method for Object Classification and Human

- Action Recognition in Video[A]. Proceeding of HCI[C]. 2006, 64-77.
- [16] Daniel W and Edmond B. Action Recognition Using Exemplarbased Embedding[A]. Proceeding of CVPR [C]. 2008 , 1-7.
- [17] Du T, Alexander S. Human Activity Recognition with Metric Learning[A]. Proceeding of ECCV [C]. 2008 , 548-561.
- [18] VenkateshBabu R, Ramakrishnan K R. Recognition of Human Actions Using Motion History Information Extracted from the Compressed Video[J]. Image and Vision Computing , 2004, 22(8) : 597-607.
- [19] Carlsson S, Sullivan J. Action Recognition by Shape Matching to Key Frames[A]. Proceeding of Workshop on Models versus Exemplars in Computer Vision[C]. 2001 , 1-8.
- [20] Laptev, Lindeberg T. Velocity Adaptation of Space- Time Interestpoints[A]. Proceeding of ICPR [C]. 2004 , 1-4
- [21] Brand, Kettner V. Discovery and Segmentation of Activities in Video[J]. IEEE Transaction Pattern Analysis and Machine Intelligence , 2000, 22(8) : 844-851.
- [22] Hongeng S, Nevatia R. Large-Scale Event Detection Using Semi-Hidden Markov Models[A]. Proceeding of ICCV[C]. 2003 , 1455-1462.
- [23] Yamato, Ohya J, Ishii K. Recognizing Human Action in Time- Sequential Images Using Hidden Markov Model[A]. Proceeding of CVPR[C]. 1992 , 379-385.
- [24] Starner T and Pentland A. Visual Recognition of American Sign Language Using Hidden Markov Models [A]. Proceeding of International Workshop on Automatic Faceand Gesture Recognition [C]. 1995 , 189-194.
- [25] Weinland D, Boyer E, Ronfard R. Action Recognition from Arbitrary Views Using 3d Exemplars[A]. Proceeding of ICCV[C]. 2007 , 1-7.
- [26] Luo Y, Wu T, Hwang J. Object Based Analysis and Interpretation of Human Motion in Sports Video Sequences by Dynamic Bayesian Networks[J]. Computer Vision and Image Understanding , 2003 , 92 : 196-216.
- [27] Ghahramani Z, Jordan M. Factorial Hidden Markov Model [J]. Machine Learning , 1997 , 29 (2) : 245-273.
- [28] Bui H, Venkatesh S, West G. Policy Recognition in the Abstract Hidden Markov Model[J]. Journal of Artificial Intelligence Research , 2002, 17: 451-499.
- [29] Michael S R, Aggarwal J K. Recognition of Composite Human Activities Through Context-Free Grammar Based Representation[A]. Proceeding of CVPR[C]. 2006 , 1709-1718.
- [30] Hu W. and Xie and T D. A Hierarchical Self-Organizing Approach for Learning the Patterns of Motion Trajectories[J]. IEEE Transaction on Neural Networks , 2004 , 15(1) : 135-144.
- [31] Vail D L, Manuela M V, Lafferty J D. Conditional Random Fields for Activity Recognition[A]. Proceeding of AAMAS[C]. 2007 , 1-8.
- [32] Sminchisescu C, Kanaujia A, Li Z, et al. Conditional Models for Contextual Human Motion Recognition [A]. Proceeding of ICCV[C]. 2005 , 1808-1815.
- [33] Laptev I, Marszalek M, Schmid C, et al. Learning Realistic Human Actions from Movies[A]. Proceeding of CVPR[C]. 2008 , 1-8.
- [34] Liu J, Luo J, Shah M. Recognizing Realistic Actions from Videos “in the wild”[A]. Proceeding of CVPR [C]. 2009 , 1996-2003.

- [35] Zhu G, Yang M, Yu K, et al. Detecting Video Events Based on Action Recognition in Complex Scenes Using Spatiotemporal Descriptor[A]. Proceeding of ACM Multimedia[C]. 2009, 165-174.
- [36] Xiang T, Gong S. Beyond Tracking: Modelling Activity and Understanding Behavior [A]. IJCV [C]. 2006, 67(1) : 21-51.
- [37] Niebles C, Li F. A Hierarchical Model of Shape and Appearance for Human Action Classification [A]. Proceeding of CVPR[C]. 2007, 1-8.
- [38] Natarajan P, Nevatia R. View and Scale Invariant Action Recognition Using Multiview Shape-Flow Models [A]. Proceeding of CVPR[C]. 2008, 1-8.
- [39] Manor Z, Irani M. Event-Based Analysis of Video[A]. Proceeding of CVPR[C]. 2001, 123-130.
- [40] Zhong H, Shi J, Visontai M. Detecting Unusual Activity in Video[A]. Proceeding of CVPR[C]. 2004, 819-826.
- [41] Xiang T, Gong S G. Video Behaviour Profiling and Abnormality Detection without Manual Labeling[A]. Proceeding of ICCV[C]. 2005, 1238-1245.
- [42] Yang S, Goncalves L, Perona P. Unsupervised Learning of Human Motion[J]. IEEE Transaction on PAMI, 2003, 25(7) : 814-827.
- [43] Zhang D, Perez D G, Bengio S, et al. Semi-Supervised Adapted Hmms for Unusual Event Detection[A]. Proceeding of CVPR[C]. 2005, 611-618.
- [44] Guan D, Yuan W, Lee Y, et al. Activity Recognition Based on Semi-Supervised Learning[A]. Proceeding of International Conference on Embedded and Real-Time Computing Systems and Applications[C]. 2007, 469-475.
- [45] Antonin G. R-Trees: A Dynamic Index Structure for Spatial Searching[A]. Proceeding of the ACM SIGMOD International Conference on Management of Data[C]. 1984, 47-57.
- [46] John T R. The K-D-B-Tree: A Search Structure for Large Multidimensional Dynamic Indexes[A]. Proceeding of the ACM SIGMOD International Conference on Management of Data[C]. 1981, 10-18.
- [47] Paolo C, Marco P, Pavel Z. M-tree: An Efficient Access Method for Similarity Search in Metric Space [A]. Proceeding of the 13th VLDB Conference[C]. 1997, 426-435.
- [48] Chanop S A, Richard H. Optimised KD-trees for Fast Image Descriptor Matching[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR)[C]. 2008, 1-8.
- [49] Piotr I, Rajeev M. Approximate Nearest Neighbor: Toward Removing the Curse of Dimensionality[A]. Proceeding of the 30th Annual ACM Symposium on Theory of Computing(STOC)[C]. 1998, 604-613.
- [50] Aristides G, Piotr I, Rajeev M. Similarity Search in High Dimensions via Hashing[A]. Proceeding of the 25th International Conference on Very Large Data Bases(VLDB)[C]. 1999, 518-529.
- [51] Mayur D, Nicole I, Piotr I, et al. Locality- Sensitive Hashing Scheme Based on p-Stable Distributions [A]. Proceeding of the 20th Annual Symposium on Computational Geometry[C]. 2004, 253-262.
- [52] Alexandr A, Piotr I. Near-Optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions[J]. Communications of the ACM, 2008, 51(1) : 117-122.
- [53] Rina P. Entropy Based Nearest Neighbor Search in High Dimensions[A]. Proceeding of the 17th Annual ACM-SIAM Symposium on Discrete Algorithm(SODA)[C]. 2006, 1186-1195.
- [54] Qin L, William J, Zhe W, et al. Multi-Probe LSH: Efficient Indexing for High-Dimensional Similarity Search[A]. Proceeding of the 33rd International Conference on Very Large Data Bases (VLDB) [C]. 2007, 950-961.

- [55] Herve J, Laurent A, Cordelia S, et al. Query-Adaptive Locality Sensitive Hashing[A]. Proceeding of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) [C]. 2008, 825-828.
- [56] Loic P, Herve J, Laurent A. Locality Sensitive Hashing: A Comparison of Hash Function Types and Querying Mechanisms[J]. Pattern Recognition Letters, 2010, 31(11): 1348-1358.
- [57] David N, Henrik S. Scalable Recognition with A Vocabulary Tree[A]. Proceeding of IEEE Computer Society Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2006, 2161-2168.
- [58] James P, Ondrej C, Michael I, et al. Object Retrieval with Large Vocabularies and Fast Spatial Matching [A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2007, 1-8.
- [59] Ruslan S, Geoffrey H. Semantic Hashing[A]. Proceeding of SIGIR Workshop on Information Retrieval and Application of Graphical Models[C]. 2007, 1-8.
- [60] Antonio T, Rob F, Yair W. Small Codes and Large Image Databases for Recognition[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition[C]. 2008, 1-8.
- [61] Yair W, Antonio T, Rob F. Spectral Hashing[A]. Proceeding of NIPS[C]. 2008, 1753-1760.
- [62] He J, Regunathan R, Chang S, et al. Compact Hashing with Joint Optimization of Search Accuracy and Time[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2011, 753-760.
- [63] Alexis J, Olivier B. Random Maximum Margin Hashing[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2011, 873-880.
- [64] Heo J, Lee Y, He J, et al. Spherical Hashing[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2012, 2957-2964.
- [65] Christoph S, Alexander M B. LDA Hash: Improved Matching with Smaller Descriptors[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(1): 66-78.
- [66] Brian K, Trevor D. Learning to Hash with Binary Reconstructive Embeddings[A]. Proceeding of NIPS [C]. 2009, 1-9.
- [67] Mohammad N, David J F. Minimal Loss Hashing for Compact Binary Codes[A]. Proceeding of the 28th International Conference on Machine Learning[C]. 2011, 1-8.
- [68] Wang J, Kumar S, Chang S. Semi-Supervised Hashing for Scalable Image Retrieval[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2010, 3424-3431.
- [69] Wang J, Kumar S, Chang S. Semi-Supervised Hashing for Large Scale Search[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2012, 34(12): 2393-2406.
- [70] Liu W, Wang J, Ji R, et al. Supervised Hashing with Kernels[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2012, 2074-2081.
- [71] Norouzi M, Punjani A, Fleet D J. Fast Search in Hamming Space with Multi-Index Hashing[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR) [C]. 2012, 3108-3115.
- [72] Sivic J, Zesserman A. Video Google: A Text Retrieval Approach to Object Matching in Video[A]. Proceeding of IEEE International Conference on Computer Vision(ICCV) [C]. 2003, 1470-1477.
- [73] Yannis A, Yannis K. Approximate Gaussian Mixtures for Large Scale Vocabularies[A]. Proceeding of EC-CV[C]. 2012, 15-28.
- [74] Mu Y, Sun J, Han T X, et al. Randomized Locality Sensitive Vocabularies for Bag-of-Features Model

- [A]. Proceeding of ECCV[C]. 2010, 748-761.
- [75] Jan C G, Cor J V, Arnold W M S. Visual Word Ambiguity[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2010, 32(7) : 1271-1283 .
- [76] James P, Ondrej C, Michael I, et al. Lost in Quantization: Improving Particular Object Retrieval in Large Scale Image Databases[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C]. 2008, 1-8.
- [77] Liu L, Wang L, Liu X. In Defense of Soft-Assignment Coding[A]. Proceeding of IEEE International Conference on Computer Vision(ICCV)[C]. 2011, 2486-2493.
- [78] Jegou H, Douze M, Schmid C, et al. Aggregating Local Descriptors into A Compact Image Representation [A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C]. 2010, 3304-3311.
- [79] Jegou H, Douze M, Schmid C, et al. Aggregating Local Image Descriptor into Compact Codes[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 34(9) : 1704-1716.
- [80] Chen D, Tsai S, Chandrasekhar V, et al. Residual Enhanced Visual Vectors for On-Device Image Matching[A]. Proceeding of 45th Asilomar Conference on Signals, Systems and Computers(ASILOMAR) [C]. 2011, 850-854.
- [81] Douze M, Ramisa A, Schmid C. Combining Attributes and Fisher Vectors for Efficient Image Retrieval [A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)[C]. 2011, 745-752.
- [82] Jegou H, Douze M, Schmid C. Improving Bag-of-Feature for Large Scale Image Search[J]. International Journal of Computer Vision, 2010, 87(3) : 316-336.
- [83] Jegou H, Douze M, Schmid C. Packing Bag-of- Features[A]. Proceeding of IEEE 12th Conference on Computer Vision(ICCV)[C]. 2009, 2357-2364.
- [84] Dong W, Charikar M, Li K. Asymmetric Distance Estimation with Sketches for Similarity Search in High-Dimensional Spaces[A]. Proceeding of the 31st annual International ACM SIGIR Conference on Research and Development in Information Retrieval(SIGIR)[C]. 2008, 123-130.
- [85] Gordo A, Perronnin F. Asymmetric Distances for Binary Embeddings[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition[C]. 2011, 729-736.
- [86] Jain M, Jegou H, Gros P. Asymmetric Hamming Embedding: Taking the Best of Our Bits for Large Scale Image Search[A]. Proceedings of the 19th ACM International Conference on Multimedia [C]. 2011, 1441-1444.
- [87] Jegou H, Douze M, Schmid C. Product Quantization for Nearest Neighbor Search[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(1) : 117-128.
- [88] Brandt J. Transform Coding for Fast Approximate Nearest Neighbor Search in High Dimensions[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition (CVPR) [C]. 2010, 1815-1822.
- [89] Tavenard R, Jegou H, Amsaleg L. Balancing Clusters to Reduce Response Time Variability in Large Scale Image Search[A]. Proceeding of 9th International Workshop on Content-Based Multimedia Indexing(CB-MI)[C]. 2011, 19-24.
- [90] Babenko A, Lempitsky V. The Inverted Multi-Index[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR)[C]. 2012, 3069-3076.

- [91] Hotelling H. Relations Between Two Sets of Variates[J]. *Biometrika*, 1936, 28(3/4) : 321-377.
- [92] Wu F, Zhang H, Zhuang Y. Learning Semantic Correlations for Cross-Media Retrieval[A]. Proceeding of 2006 IEEE International Conference on Image Processing[C]. 2006, 1465-1468.
- [93] Akaho S. A Kernel Method for Canonical Correlation Analysis. arXiv Preprint cs/609071, 2006.
- [94] Hardoon D R, Taylor J S. Sparse Canonical Correlation Analysis[J]. *Machine Learning*, 2011, 83(3) : 331-353.
- [95] Chen X, Liu H, Carbonell J G. Structured Sparse Canonical Correlation Analysis[A]. Proceeding of the 15th International Conference on Artificial Intelligence and Statistics[C]. 2012, 199-207.
- [96] Rupnik J, Taylor J S. Multi-view Canonical Correlation Analysis[A]. Proceeding of the Conference on Data Mining and Data Warehouses[C]. 2010, 1-4.
- [97] Sharma A, Kumar A, Daume H, et al. Generalized Multiview Analysis: A Discriminative Latent Space [A]. Proceedings of 2012 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)[C]. 2012, 2160-2167.
- [98] Balakrishnan S, Puniyani K, Lafferty J. Sparse Additive Functional and Kernel CCA. arXiv Preprint arXiv: 1206.4669, 2012.
- [99] Bai B, Weston J, Grangier D. Learning to Rank with(a lot of) Word Features[J]. *Information Retrieval*, 2009, 13(3) : 291-314.
- [100] Bai B, Weston J, Grangier D, et al. Polynomial Semantic Indexing[J]. *Advances in Neural Information Processing Systems*, 2009, 4(1) : 1-9.
- [101] Putthividhy D, Attias H T, Nagarajan S S. Topic Regression Multi-Modal Latent Dirichlet Allocation for Image Annotation[A]. Proceedings of 2010 IEEE Conference on Computer Vision and Pattern Recognition(CVPR)[C]. 2010, 3408-3415.
- [102] Jia Y, Salzmann M, Darrell T. Learning Cross-Modality Similarity for Multinomial Data[A]. Proceedings of 2011 IEEE International Conference on Computer Vision(ICCV)[C]. 2011, 2407-2414.
- [103] Virtanen S, Jia Y, Klami A, et al. Factorized Multi-Modal Topic Model. UAI, arXiv: 1210.4920, 2012.
- [104] Jia Y, Salzmann M, Darrell T. Factorized Latent Spaces with Structured Sparsity[J]. *Advances in Neural Information Processing Systems* 23, 2010, 982-990.
- [105] Bronstein M M, Bronstein A M, Michel F, et al. Data Fusion Through Cross- Modality Metric Learning Using Similarity-Sensitive Hashing[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR)[C]. 2010, 3594-3601.
- [106] Kumar S, Udupa R. Learning Hash Functions for Cross- View Similarity Search[A]. Proceeding of the 22th International Joint Conference on Artificial Intelligence(IJCAI)[C]. 2011, 1360-1365.
- [107] Zhen Y, Yeung D Y. A Probabilistic Model for Multimodal Hash Function Learning[A]. Proceeding of the ACM International Conference on Knowledge Discovery and Data Mining(SIGKDD)[C]. 2012, 940-948.
- [108] Zhen Y, Yeung D Y. Co-Regularized Hashing for Multimodal Data[J]. *Advances in Neural Information Processing Systems(NIPS)*[C]. 2012, 1385-1393.
- [109] Joachims T. Optimizing Search Engines Using Clickthrough Data[A]. Proceeding of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining[C]. 2002, 133-142.
- [110] Liu T Y. Learning to Rank for Information Retrieval[J]. *Foundations and Trends in Information Retrieval*, 2009, 3(3) : 225-331.
- [111] Li H. Learning to Rank for Information Retrieval and Natural Language Processing[J]. *Synthesis Lectures*

- on Human Language Technologies, 2011, 4(1): 1-113.
- [112] Cao Z. Learning to Rank: From Pairwise Approach to Listwise Approach[A]. Proceeding of the 24th International Conference on Machine learning[C]. 2007, 129-136.
- [113] Yue Y, Finley T, Radlinski F, et al. A Support Vector Method for Optimizing Average Precision[A]. Proceeding of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval[C]. 2007, 271-278.
- [114] Tschantaridis I, Hofmann T. Large Margin Methods for Structured and Interdependent Output Variables [J]. Journal of Machine Learning Research, 2005, 6: 1453-1484.
- [115] Chakrabarti S, Khanna R, Sawant U, et al. Structured Learning for Non-Smooth Ranking Losses[A]. Proceeding of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining [C]. 2008, 88-96.
- [116] Grangier D, Bengio S. A Discriminative Kernel-Based Model to Rank Images from Text Queries [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008, 30(8): 1371-1384.
- [117] Crammer K, Dekel O, Keshet J, et al. Online Passive-Aggressive Algorithms[J]. The Journal of Machine Learning Research, 2006, 7: 551-585.
- [118] Zhai X, Peng Y, Xiao J. Effective Heterogeneous Similarity Measure with Nearest Neighbors for Cross-Media Retrieval[A]. Proceeding of Advances in Multimedia Modeling[C]. 2012, 312-322.
- [119] Deerwester S, Dumais S T, Furnas G W, et al. Indexing by Latent Semantic Analysis[J]. Journal of the American Society for Information Science, 1990, 41(6): 391-407.
- [120] Hofmann T. Unsupervised Learning by Probabilistic Latent Semantic Analysis[J]. Machine Learning, 2001, 42: 177-196.
- [121] Mcfee B, Lanckriet G. Metric Learning to Rank[A]. Proceeding of the 27th Annual International Conference on Machine Learning[C]. 2010, 775-782.
- [122] Lim D K H, Mcfee B, Lanckriet G, et al. Robust Structural Metric Learning[A]. Proceeding of the 30th International Conference on Machine Learning[C]. 2013, 615-623.
- [123] 方显春, 王蕴红, 谭铁牛. 融合人脸轮廓和区域信息改进人脸检测[J]. 计算机学报, 2004, 27(4): 482-491.
- [124] 刘红毅, 王蕴红, 谭铁牛. 基于改进ENN算法的多生物特征融合的身份验证[J]. 自动化学报, 2004, 30(1): 78-85.
- [125] 王亮, 胡卫明, 谭铁牛. 基于步态的身份识别[J]. 计算机学报, 2003, 26(3): 353-360.
- [126] 王亮, 胡卫明, 谭铁牛. 人运动的视觉分析综述[J]. 计算机学报, 2002, 25(3): 225-237.
- [127] 徐成华, 王蕴红, 谭铁牛. 三维人脸建模与应用[J]. 中国图像图形学报, 2004, 9(8): 893-903.
- [128] 童晓峰, 刘青山, 卢汉清, 等. 基于语义单元的足球视频事件检测[J]. 自动化学报, 2005, 31(4): 523-529.
- [129] 吴川, 马字飞, 贺玉文, 等. 体育视频中基于语义推理的事件检测方法[J]. 清华大学学报(自然科学版), 2003, 43(4): 507-509.
- [130] 肖鹏, 吴玲达, 老松杨, 等. 一种基于可视化的新闻视频挖掘方法[J]. 情报学报, 2004, 23(3): 307-312.
- [131] Zhuang Y, Zhuang Y, Li Q, et al. Indexing High-Dimensional Data in Dual Distance Spaces: A Symmetrical Encoding Approach[A]. Proceeding of the 11th International Conference on Extending Database Technology[C]. 2008, 241-251.

- [132] He Y, Yu J. MFI-Tree: An Effective Multi-feature Index Structure for Weighted Query Application[J]. Computer Science and Information System, 2010, 7(1): 139-151.
- [133] Kuo Y, Chen K, Chiang C, et al. Query Expansion for Hash-Based Image Object Retrieval[A]. Proceeding of the 17th International Conference on Multimedia[C]. 2009, 65-74.
- [134] Zhang X, Li Z, Zhang L, et al. Efficient Indexing for Large Scale Visual Search[A]. Proceeding of the IEEE 12th Conference on Computer Vision(ICCV)[C]. 2009, 1103-1110.
- [135] Shao J, Wu F, Ouyang C, et al. Sparse Spectral Hashing[J]. Pattern Recognition Letters, 2012, 33(1): 271-277.
- [136] Zhuang Y, Liu Y, Wu F, et al. Hypergraph Spectral Hashing for Similarity Search of Social Image[A]. Proceeding of the 19th ACM International Conference on Multimedia(MM)[C]. 2011, 1457-1460.
- [137] Zhou W, Lu Y, Li H, et al. Scalar Quantization for Large Scale Image Search[A]. Proceeding of the ACM International Conference on Multimedia(MM)[C]. 2012, 169-178.
- [138] Kuo Y, Lin H, Cheng W, et al. Unsupervised Auxiliary Visual Words Discovery for Large-Scale Image Object Retrieval[A]. Proceeding of IEEE Conference on Computer Vision and Pattern Recognition(CVPR)[C]. 2011, 905-912.
- [139] Cai Y, Tong W, Yang L, et al. Constrained Keypoint Quantization: Towards Better Bag-of-Words Model for Large-scale Multimedia Retrieval[A]. Proceeding of the International Conference on Multimedia Retrieval(ICMR)[C]. 2012, 1-8.
- [140] Ai L, Yu J, Guan T, Spherical Soft Assignment: Improving Image Representation in Content-based Image Retrieval[A]. Proceeding of Pacific-Rim Conference on Multimedia[C]. 2012, 801-810.
- [141] Chen Y, Guan T, Wang C. Approximate Nearest Neighbor Search by Residual Vector Quantization[J]. Sensors, 2010, 10, 11259-11273.
- [142] Zhuang L, Zhuang Y, Wu J, et al. Image Retrieval Approach Based on Sparse Canonical Correlation Analysis[J]. Journal of Software, 2012, 23(5): 1295-1304.
- [143] Zhuang Y, Wang Y, Wu F, et al. Supervised Coupled Dictionary Learning with Group Structures for Multi-modal Retrieval[A]. Proceeding of the Twenty-Seventh Conference on Artificial Intelligence(AAAI)[C]. 2013, 1070-1076.
- [144] Yu Z, Tang S, Zhang Y, et al. Image Ranking via Attribute Boosted Hypergraph[A]. Proceeding of PCM[C]. 2012, 779-789.
- [145] Hong C, Zhu J. Hypergraph-Based Multi-Example Ranking with Sparse Representation for Transductive Learning Image Retrieval[J]. Neurocomputing, 2013, 101(4): 94-103.
- [146] Hong C, Jun Y, Jonathan L, et al. Multi-View Hypergraph Learning by Patch Alignment Framework[J]. Neurocomputing, 2013, 118(22): 79-86.
- [147] Liu Y, Shao J, Wu F, et al. Hypergraph Spectral Hashing for Image Retrieval with Heterogeneous Social Contexts[J]. Neurocomputing, 2013, In Press(<http://dx.doi.org/10.1016/j.neucom.2012.02.051>).
- [148] Liu X, Mu Y, Lang B, et al. Compact Hashing for Mixed Image-Keyword Query Over Multi-Label Images[A]. Proceeding of the 2nd ACM International Conference on Multimedia Retrieval(ICMR)[C]. 2012, 18-27.
- [149] Lu X, Wu F, Tang S, et al. A Low Rank Structural Large Margin Method for Cross-Modal Ranking[A]. Proceeding of the 36th Annual International ACM SIGIR Conference on Research & Development on Information Retrieval[C]. 2013, 433-442.

- [150] Smith J R, NaPhade M. Multimedia Semantic Indexing Using Model Vectors[A]. Proceeding of International Conference on Multimedia Expo[C]. 2003, 445-448.
- [151] Jiang W, Chang S F, Loui A. Active Concept-Based Concept Fusion with Partial User Labels[A]. Proceeding of TREC Video Retrieval Evaluation[C]. 2004, 2917-2920.
- [152] Hau Ptman A, Chen M Y, Christel M. Confounded Expectations: Informedata at Trecvid 2004 [A]. Proceeding of TREC video Retrieval Evaluation[C]. 2004.
- [153] Wang M, Mei T, Yuan X, et al. Video Annotation by Graph-Based Learning with Neighbourhood Similarity[A]. Proceeding of ACM Multimedia[C]. 2007, 325-328.

作者简介

本文的视频事件识别与标注部分由徐常胜和张天柱撰写，高维索引结构部分由于俊清和艾列富撰写，跨媒体检索理论与技术部分由吴飞和庄越挺撰写，于俊清负责统稿。以下以姓氏笔画排序。

于俊清 博士，华中科技大学计算机科学与技术学院，教授，博士生导师，主要研究方向为数字媒体处理与检索、多核计算与流编译，中国计算机学会高级会员，多媒体专业委员会副主任。



艾列富 硕士，华中科技大学计算机科学与技术学院，计算机应用技术专业博士研究生，主要研究方向为基于内容的大规模图像高维索引与检索。



庄越挺 博士，浙江大学计算机学院院长，教授，主要研究方向为跨媒体计算理论、计算机动画、数字图书馆和人工智能等。E-mail: yzhuang@cs.zju.edu.cn。



吴 飞 博士，浙江大学计算机学院人工智能研究所副所长，教授，主要研究方向跨媒体计算、稀疏表达与统计学习，中国计算机学会高级会员。E-mail: wufei@cs.zju.edu.cn。



徐常胜 博士，中国科学院自动化所模式识别国家重点实验室研究员，中国 - 新加坡数字媒体研究院执行院长。主要研究方向包括多媒体分析、计算机视觉和模式识别。中国计算机学会高级会员。



张天柱 博士，中国科学院自动化研究所助理研究员。主要研究方向包括多媒体分析、计算机视觉、模式识别与图像处理等。中国计算机学会会员。



高性能计算机系统冷却技术

CCF 工程与工艺专业委员会

摘要

随着计算机系统性能与计算密度的不断提高，冷却技术成为制约系统可靠性的重要因素。本文对高性能计算机系统的冷却途径和关键技术进行了介绍，分析了国际发展现状和趋势，对我国多家单位在此方面的进展进行了介绍，最后对于我国在此方面的后续发展提出了建议。

关键词：高效能计算，高性能计算机系统，冷却技术，数据中心

Abstract

With the computer system performance and compute density continues to increase, the cooling become an important factor restricting system reliability. This report introduced the cooling technical approach and key technologies, analyzed the international current situation and trends development, and introduced our country a number of units on progress in this regard. The last are the subsequent development recommendations for our country in this regard.

Keywords: high productivity computing, high performance computing system, cooling technology, data center

1 引言

高性能计算（High Performance Computing）是计算机科学的一个分支，研究开发高性能计算机及相关软件，把一个大的程序并行化以便在多处理器上同时运行，实现问题的快速求解。由于高性能计算便于高效率、低成本、大尺度、全过程且无损伤地研究众多科学与工程问题，高性能计算经过几十年的发展，已被公认为继传统科学方法——理论和实验之后的第三大科学方法，是科技创新的重要手段。如今几乎所有科学和工程领域，乃至商业和政府决策都离不开高性能计算。高性能计算技术在一个国家的发展中，特别是一些尖端科技的发展中，发挥着不可替代的作用。高性能计算发展及应用水平正日益成为一个国家科研实力的重要标志之一。

高性能计算机系统经历了电子管、晶体管、集成电路、大规模集成电路和超大规模集成电路的演变，而不断攀登高速度永远是计算机研究者执着追求的目标。从 1976 年美国的 Cray-1 向量计算机突破每秒 1 亿次运算速度到 2008 年美国的 Roadrunner 系统达到千万亿次量级，短短三十多年高性能计算的发展速度令人叹为观止。如今的高性能计算机系统具有性能高、规模大的突出特点，也面临着严峻的可靠性、可用性和经济性等问题。为了应对高速发展暴露的一系列重大问题，出现了从高性能计算到高效能计算的重大

转变。高效能计算概括了高性能、低功耗、可编程性、程序可移植性与鲁棒性等多个要素，对计算机的体系结构与工程实现提出了更多挑战^[1]。

冷却技术是指为了实现对计算机系统内元器件的温度控制，保证其正常、可靠地工作而采取的传热技术。理论和实践都表明，电子元器件的性能与可靠性与其工作温度密切相关。当超出器件的额定温度条件时，器件就可能丧失其规定的功能导致失效。一般地，降低工作温度、温度梯度和温度变化幅度，都有助于提高计算机系统的可靠性，但也会带来体积、成本等诸多问题。冷却技术的基本目的就是要在防止热失效的前提下，确定合理传热方式、传热材料和传热结构，将系统的温度条件控制在合理的阈值范围内。

高性能计算往往采用当时最高性能的元器件。为了获得更高的性价比，正如摩尔定律总结的那样，硅片的集成度不断提高。尽管单个门电路的功耗在减小，但由于工作频率和集成度的迅猛提高，芯片的热功耗仍不断提高。伴随着性能的突飞猛进，器件的热功耗也日益上升。同样是芯片的功耗问题，制约着单芯片的性能。基于 Gustafson 加速比公式，大规模并行蓬勃发展，成为高性能计算的主流技术。目前千万亿次规模的高性能计算机系统往往就包括数万颗处理器、数十万个功率元器件，设计什么样的冷却结构，实现温度控制的同时，减小冷却部件体积，提高组装密度，降低构建成本，对大规模计算机系统的冷却无疑是个重大挑战。随着并行计算机系统规模的不断扩大，其固有的弊病逐渐显现出来。一方面，依靠高性能商用微处理器的发展，集群系统的峰值性能不断增长；另一方面，所构建的高端系统在实际应用中表现出很低的实用效率，系统稳定性差，功耗不断攀升^[2]。目前千万亿次规模系统的功耗都达数兆瓦，日本的地球模拟器系统的功耗更高达 11.9 兆瓦。为了将巨大的热量迅速转移到大气环境，高性能计算机系统不得不配套建设庞大的制冷系统，制冷系统的建设和运行更进一步加剧了高性能计算机系统的成本压力。高性能计算机的冷却技术已成为高性能计算机发展的重要组成部分和制约计算机系统可靠性运行的重要因素，国内外相关的科研机构和公司都在积极研究和探索新的更高效、更节能的冷却技术和设备。

2 国际研究现状

2.1 国外典型高性能计算机系统的冷却技术

高性能计算机技术集信息技术的大成，既是衡量一个国家或地区综合实力的重要技术指标，也是迸发科技创新能力的高端基础设施。在高性能计算领域处于领先地位就等于占领了未来发展的制高点。世界经济强国对高性能计算无不十分重视，尤其是美国、日本和欧洲各国，为了增强其在国际上的综合竞争力，纷纷制定长期发展计划，政府投资建设与民间外包运营并举，竞相投入巨大人力、物力和财力争夺这一战略高地。长期

以来，美日等国均将研制和应用最高性能的超级计算机系统作为保持其科技领先优势的国家战略。2011 年美国总统财年预算报告强调：“持续研制顶级系统（leadership-class system），即最高性能的先进系统，用于能源、环境和国家安全等最前沿研究，以高端计算能力支撑学术机构、联邦实验室和工业领域的研究人员，推进诸如生物学、现代医学、化学、天气气候、能源、环境、材料科学、纳米科技、航空宇航、物理及其他领域的研究计划任务”。日本、欧洲和俄罗斯也纷纷加大了对高性能计算机研发的投入。下面介绍这些科研机构和公司最新高性能计算机的冷却技术和应用方案。

（1）IBM 高性能计算机冷却技术

IBM 在全球高性能计算机排名 TOP 500 中占据数量最多的公司，IBM 在高性能计算机的发展历史中具有非常重要的地位，从高性能计算机诞生时直到现在持续对高性能计算机进行研究和开发。IBM 高性能计算机经过一代代的技术发展演变，目前最具代表性、规模和数量也最多的就属蓝色基因系列（BlueGene）超级计算机。BlueGene 从 2004 年开始研制经历了 BlueGene/L、BlueGene/C、BlueGene/P、BlueGene/Q 不同阶段，基于最新一代的 BlueGene/Q 系列，IBM 构建了红杉（Sequoia）、米拉（Mira）、费密（Fermi）等超级计算机系统。

BlueGene/Q 以前的 BlueGene 系统采用强迫风冷技术冷却，组装与散热结构如图 1 所示^[3]。由于采用了低功耗的处理器（单处理器功耗约 25 瓦），BlueGene 采用板卡式组装结构。2 个处理器一个计算子卡，单卡功耗 54 瓦。16 个子卡按 8 行 2 列组装成一个结点插箱，单插箱功耗约 860 瓦。32 个插箱双面对插构成一个机柜，单机柜功耗约 28 千瓦。

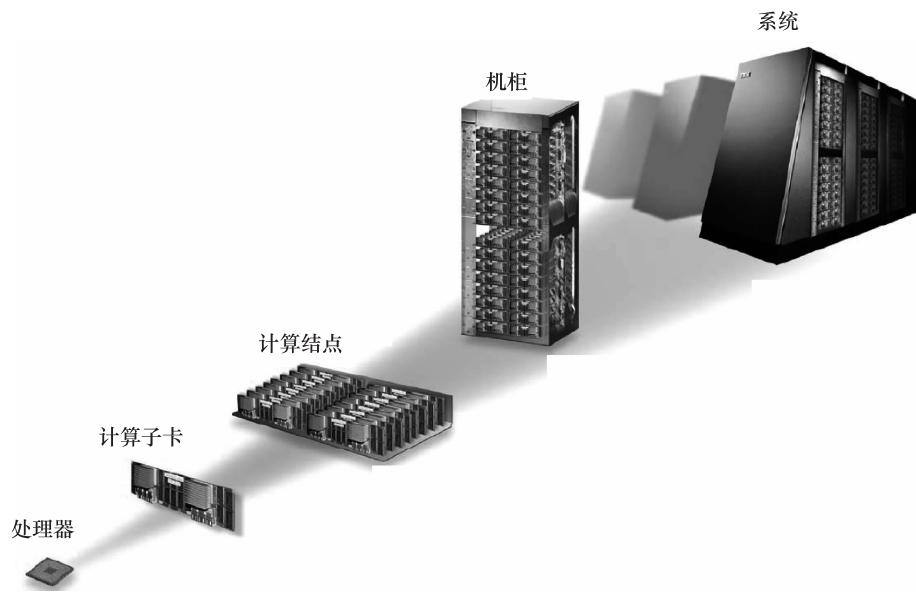


图 1 采用风冷技术的 Blue Gene 系统^[3]

采用风冷技术的 Blue Gene 系统在机柜左侧设计安装了风机模块，如图 2 所示。相邻机柜间设计了专门的风柜。风柜底板进风，顶部出风，进出风通道间设有斜隔板。

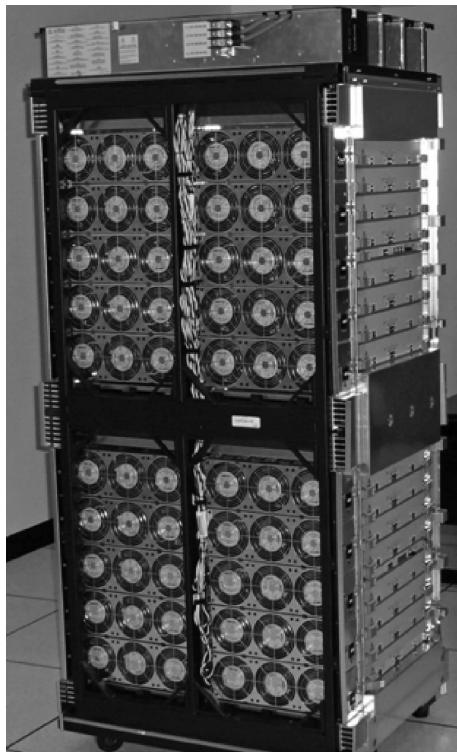


图 2 强迫风冷的 Blue Gene 机柜^[3]

BlueGene/Q 系统采用类似 BlueGene/L 的组装结构，从芯片到系统的层次式组装结构如图 3 所示^[4]。

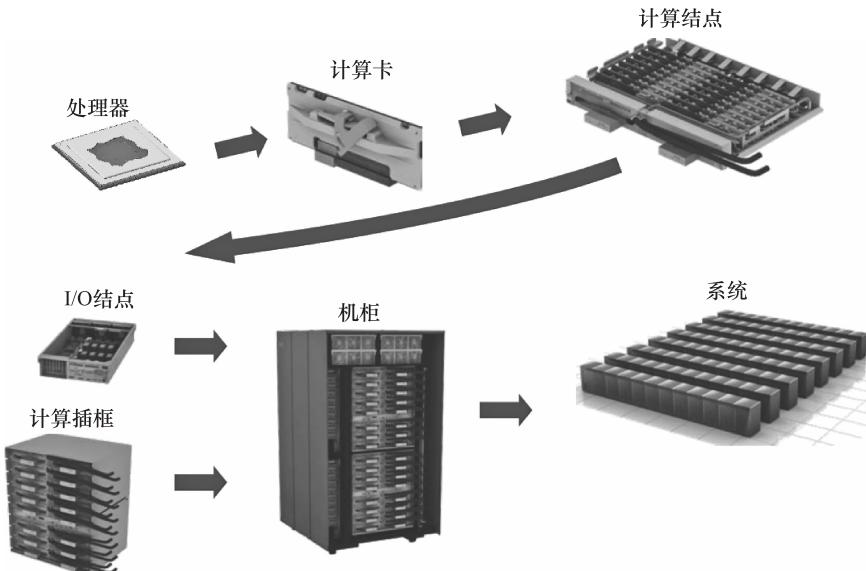


图 3 Blue Gene/Q 层次式组装结构

为了降低冷却系统的功耗, BlueGene/Q 系统采用间接水冷技术。1 颗 Cell 芯片、16GB DDR3 内存颗粒及相应的分立元件组装成一个计算子卡, 每个计算子卡设计一块导热片紧贴在功率器件上, 如图 4 所示。

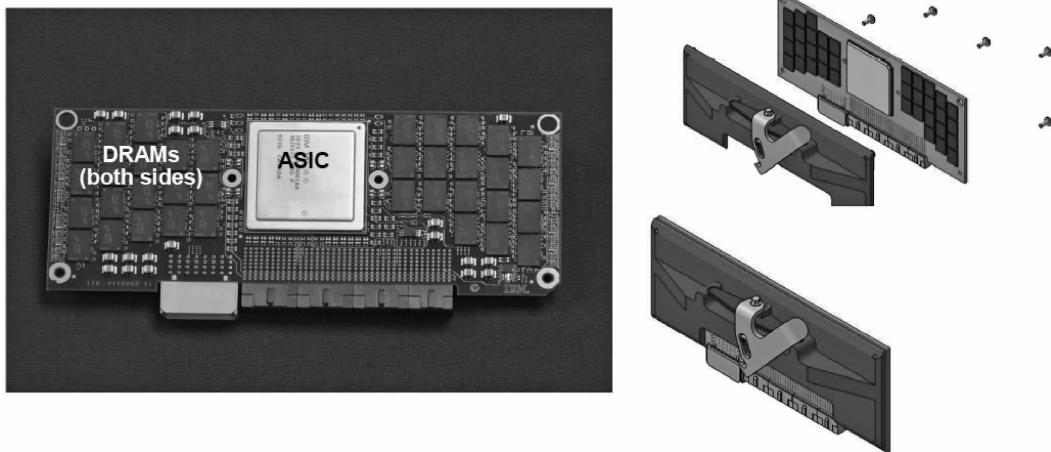


图 4 Blue Gene/Q 计算子卡

32 个计算子卡、互连部件、电源部件组装成一个结点。每个结点底板上设计有条状冷板, 上下插接组装的计算子卡导热片与冷板接触, 通过导热的方式将热量从功率芯片转移到条状冷板中。每个结点前面板有一对快换接头, 输入/输出冷冻水。结点组装和冷却结构如图 5 所示。



图 5 水冷散热的 Blue Gene/Q 的计算结点

16 个计算结点双面对插在 1 块互连中板上构成一个超结点, 2 个超结点以及最多 4 个 I/O 结点组装成一个机柜。单机柜峰值性能约 209TFlops, 功耗约 100 千瓦。96 个机柜

构成峰值 20PFlops 的系统。

相对于传统的风冷技术，水冷技术具有传热效率高，容量大，耗能少等优点，但也存在构建成本高、系统可维修性差以及一定的水泄露风险。

面对日益严重的功耗墙问题，近几年 IBM 的高性能计算机在关注系统高速度的同时也高度注重高性能计算机的低能耗。在欧洲苏黎世联邦理工学院，IBM 研制的 SuperMUC 超级计算机如图 6 所示。利用当地有利的气候条件，通过设计低热阻的水冷散热器，减小冷却工质与芯片工作温度间的传热温差，如图 7 所示，进水温度到达 60℃ 芯片也不至于过热。由于出水温度可达 65℃，冷却过芯片的温水可供居民和商业建筑供暖使用，温水冷却（Warm Water Cooling）实现了能源的循环利用，大大减少计算机系统的碳排放量^[5]。

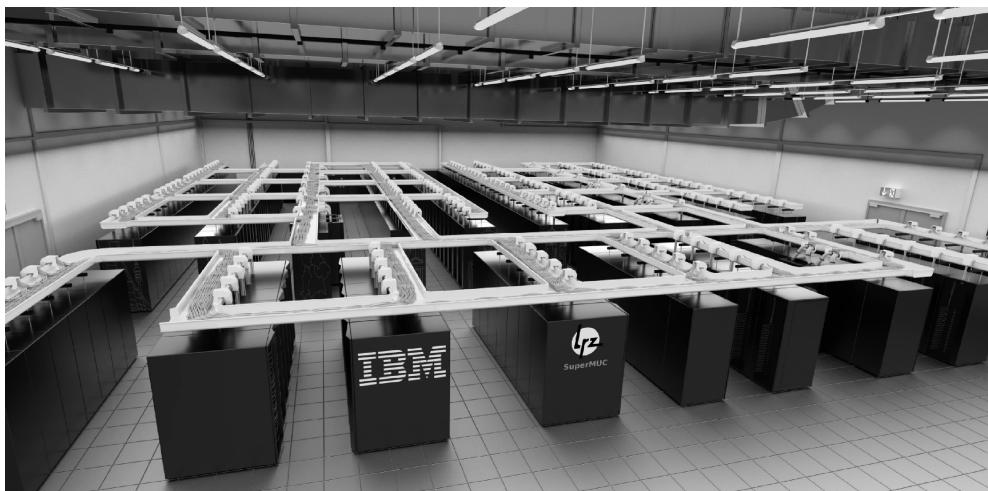


图 6 IBM SuperMUC 系统

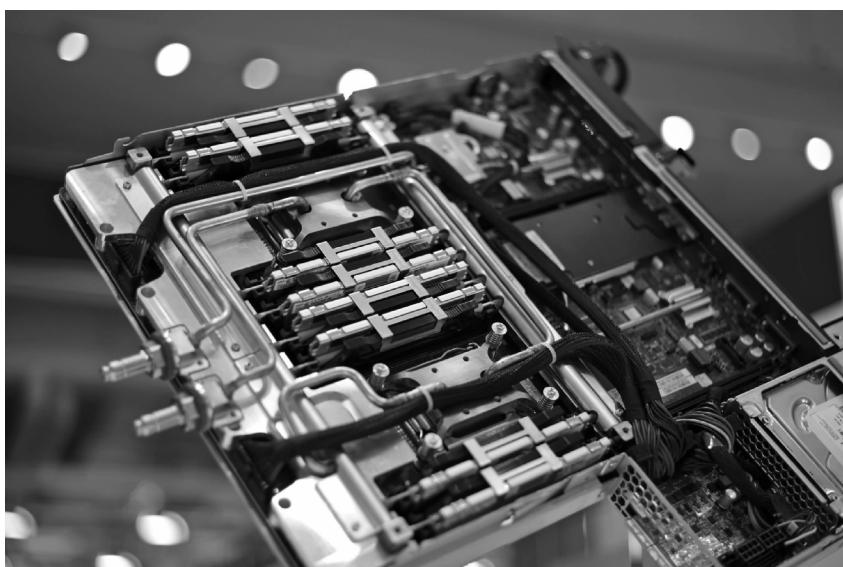


图 7 IBM SuperMUC 系统计算结点

(2) 克雷高性能计算机冷却技术

克雷 (Cray) 是美国另外一家非常重要高性能计算机的研究和设计公司，其创始人西摩·克雷 (Seymour Cray) 被称为超级计算机之父，其在 20 世纪 60 年代独霸超级计算机市场。

自成功推出 Cray XT5 计算机以来，近年来的 XE6、XK6、XK7 系统都采用了类似的冷却技术-ECOphlex^[6]。构建的典型高性能计算机系统包括 Jaguar、Cielo、Hopper、Titan 等，如图 8 所示。



图 8 采用 ECOphlex 冷却技术的 Cray XT 系列机柜

ECOphlex 可以看做带有蒸发器的计算机柜，典型的 ECOphlex 冷却结构如图 9 所示。冷却气流从机架底部进入机架，经过机架底部的离心风机加压后流过计算机设备，携带热量温度升高后的热气流在机架顶部被蒸发器冷却，降温后的气流进入机房内再次循环。不同于常规的单相管道内强制对流的表冷器，ECOphlex 蒸发器管道内发生的是沸腾换热，换热效率高，换热器体积小、重量轻，风侧压力损失小，由于可以利用相变潜热，带走同样热量需要的工质的流量小，循环泵消耗的功率也小。同时制冷剂 R134a 在管路出现异常泄漏时会闪发为气体，能够确保计算机的安全运行。

Cray XT5 系列巨型机插件级风冷，机柜级液冷，组装密度并不高，单机柜功耗约 45 千瓦左右。但对冷却作了很多细致的优化，其目标是降低成本和提高可靠性。在插件级，4 个串联用风的散热器各不相同，沿流动方向肋片依次增加，这样不仅可以降低总体流动阻力，而且可以提高各处理器工作温度的均匀性，代价是制造成本稍高些，维护稍麻烦些。在机柜级，由于功率密度并不大，Cray XT5 只使用了一个大的离心风机，冷却系统可靠性较高。使用相变换热器可以大大提高换热效率，提高了送水温度，减小换热器的体积，在省掉机房精密空调的同时还可以减小风机运行功耗，提高了能源利用效率。

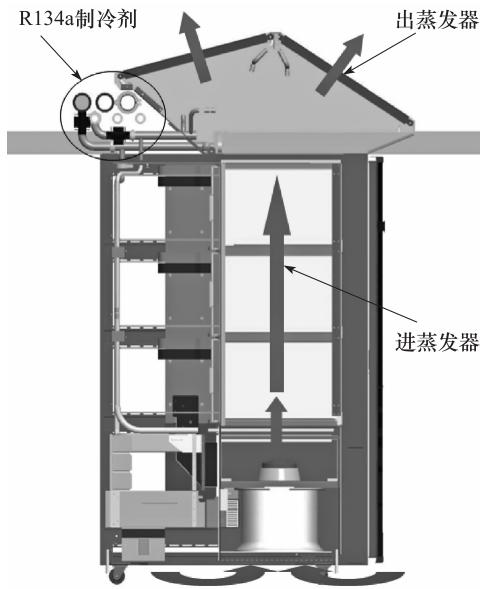


图9 Cray ECOphlex 冷却方案

2012年，Cray收购了美国另外一家重要的高性能计算机供应商“Appro”。Appro的产品在高端服务器市场占有非常地位，其部分产品采用了Asetek公司提供的Rack-CDU冷却技术方案——一种以机架为单元的芯片水冷技术集成解决方案。

RackCDU冷却技术在主板上使用了水冷芯片技术，芯片上集成了微型泵和冷却板的冷却组件对CPU和GPU进行液体冷却，如图10所示。液体冷却可以带走服务器设备产生的80%以上的热量，相对风冷技术，冷却效率高，能够有效降低冷却系统的功耗。

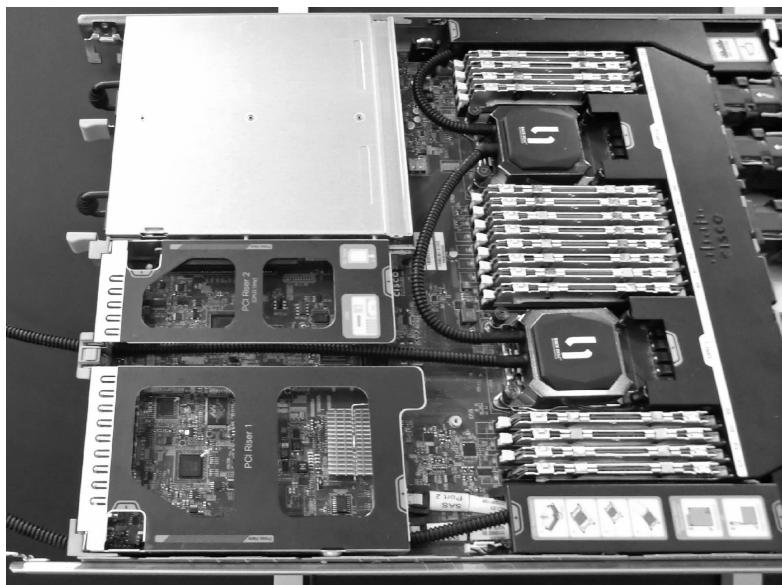


图10 RackCDU 冷却技术方案

Cray 公司不仅是高性能计算技术的佼佼者，也是高性能计算机冷却技术的重要开拓者。1976 年的 Cray-1 及其后的 Cray XMP 系统，采用了利用氟利昂作冷却介质的铜冷板间接液冷。1986 年，Cray-2 及其后的 Cray T90 采用了浸没式液冷。在 Cray Y-MP、Cray C90、Cray T3D、Cray T3E 等系统中，也尝试采用了氟利昂冷板冷却技术。Cray X1 系统采用的喷射相变冷却技术，更是将冷却与组装技术推上了一个新的高度^[7]。

图 11 展示了 CRAY-X1 在板级组装的一些主要组装工艺技术，包括多芯片陶瓷 MCM 封装向量处理器、三维高密度组装。

CRAY-X1 采用的向量处理器由几个 SSP 芯片陶瓷 MCM 封装。一块 34 层的印制板上有 4 个 MCM 芯片，如图 12 所示。单个处理器的峰值功耗在 600 瓦以上，最高表面功率密度达到 70W/cm^2 。要实现如此高功率密度的冷却，其他方法很难奏效。根据该系统的特点和 CRAY 在液体冷却上的技术积累，高效的喷射相变冷却。

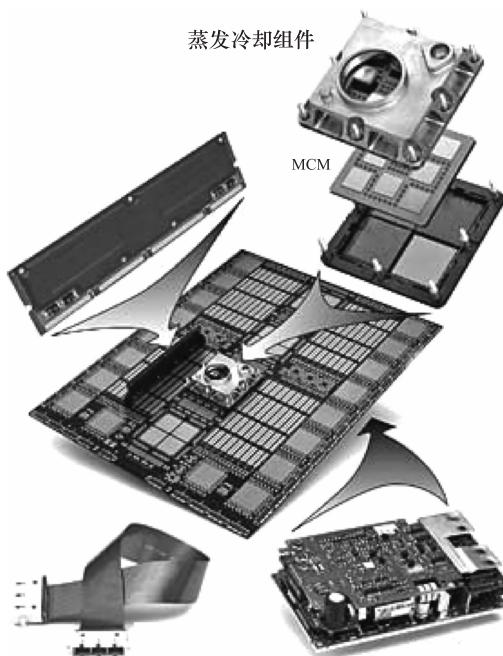


图 11 Cray X1 板级组装技术

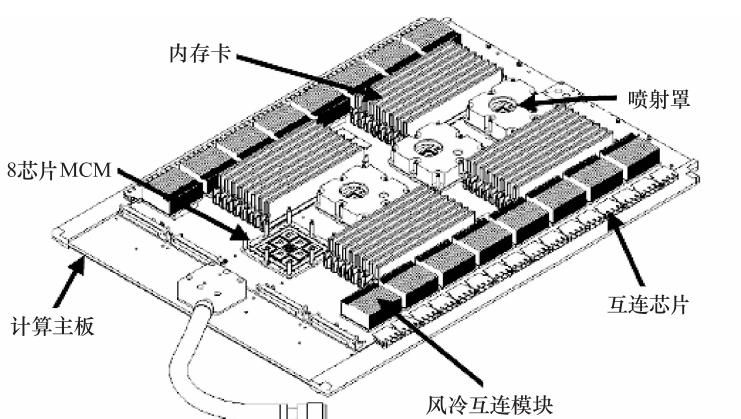


图 12 Cray X1 计算主板

CRAY X1 采用液冷和风冷相结合的冷却方式，对 CPU 采用液冷，其他器件用风冷。每个 CPU 有一个铝罩，4 个 CPU 共用一个喷、回不锈钢喷头，如图 13 所示。CRAY-X1 的液冷用 3M 公司提供的 FC-72 作为工质。FC-72 沸点温度约 56°C ，芯片设计工作壳温在 70°C 左右。

CRAY X1 总装机规模并不大，尽管其喷射冷却技术的运行复杂性没有经过大规模系统运行的实践检验。但其局部极高的散热能力值得称道。

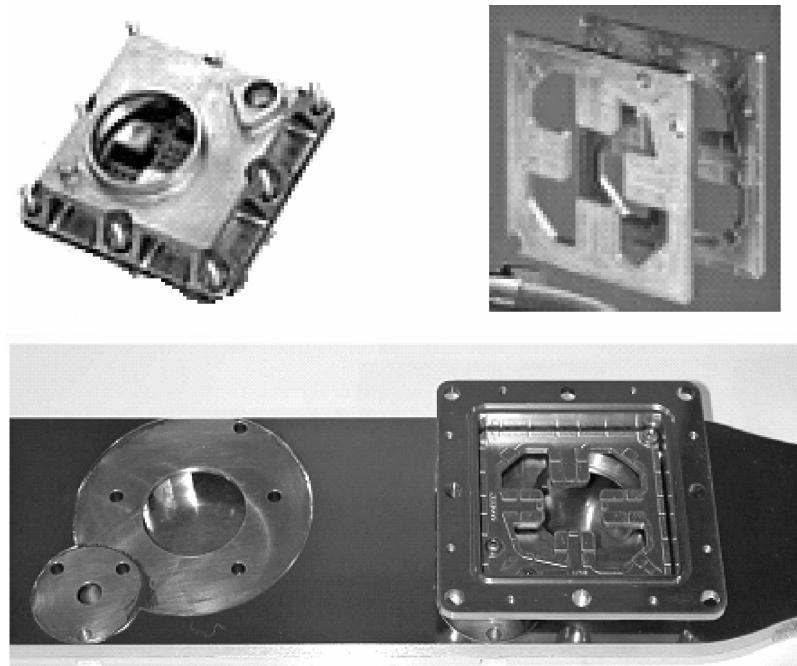


图 13 Cray X1 的喷射结构

(3) 日本高性能计算机冷却技术

日本研制高性能计算机主要以日本理化研究所、富士通、NEC 等公司，2002 年的 NEC 的地球模拟器（Earth Simulator）夺得当年的最快超级计算机排名，2011 年理化所与富士通联合研制“K（京）”超级计算机的再次夺冠。

NEC 的地球模拟器、燕子等系统都采用了强迫风冷技术。地球模拟器由日本 NEC 公司研制，峰值性能为 35.86TFLOPS，从 2002 年 6 月到 2004 年 6 月，在 TOP500 上一直排名第一^[8]。该系统采用了 5120 个定制向量处理器，功耗高达 12MW，地球模拟器的机房共有四层，机器存放在四楼，三楼布置了上百公里的铜质电缆用于全局互连，二楼是空调房，一楼则是电力房，这样布局的一个重要原因是由于它功耗太大。虽然地球模拟器在可编程性和系统实用效率方面有所变革，但是其极高的功耗和硬件成本，使得该机器成为迈向高效能计算的一个反面典型。

“京”计算机的冷却散热系统也是以芯片级水冷技术为主的冷却架构，如图 14 所示。每个计算结点的主板有 4 个 CPU 和 4 个 ICC 运算器分别用 8 块水冷冷板进行冷却，在主板前部有一个水冷模块（Water Cooling Module），将外部进入的冷却水分成两路，分别对左右两侧的 2 个 CPU 和 2 个 ICC 进行串联冷却。无论是单器件功耗、单主板功耗和单机柜功耗，“京”系统的热密度都不高，其水冷技术的实现难度也不大。K 系统使用水冷不是强调整节能或提高组装密度，而是侧重于将处理器工作温度控制得更低，从而降低漏电电流和提高可靠性^[9]。

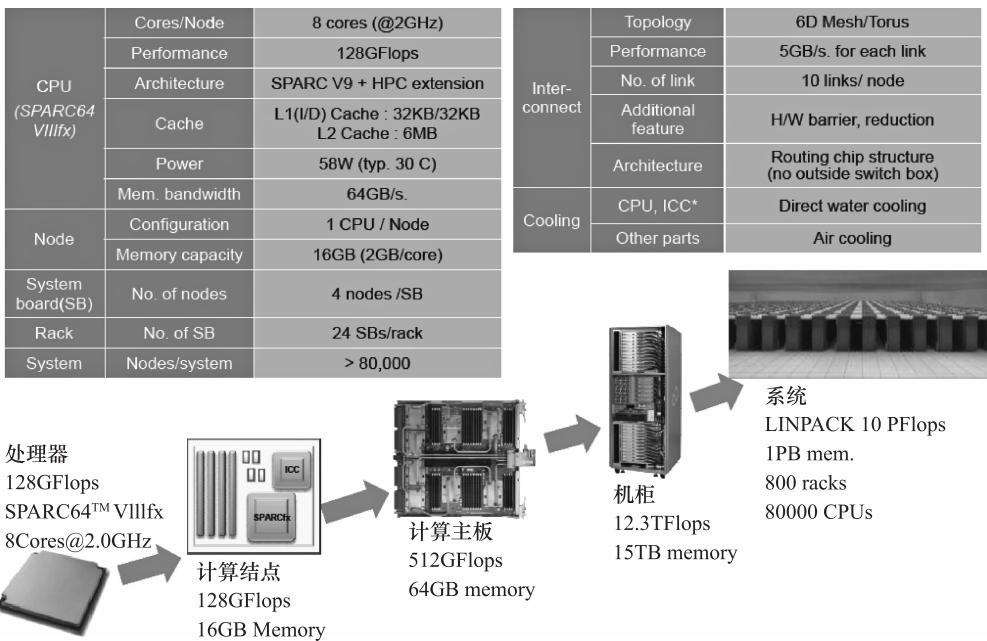


图 14 “京”计算机系统组装架构

“京”系统使用的处理器功耗并不大，仅 58 瓦。一块计算板 4 个处理器，每个处理器带 8 个存储条。一个机柜 24 块计算板，96 个处理器。由于组装密度不高，做出 10P 的系统需要 800 个机柜，规模很大。K 系统的计算主板尺寸大约 460×560 ，包括 4 个处理器，4 个互连芯片， 4×8 个内存条。处理器和互连芯片使用水冷散热器冷却，两个处理器和两个互连芯片为一组，全部水冷散热器用铜管连接成一体。存储条采用风冷，水平送风。采用一对快换接头实现供回水，正面手动操作。计算主板如图 15 所示。

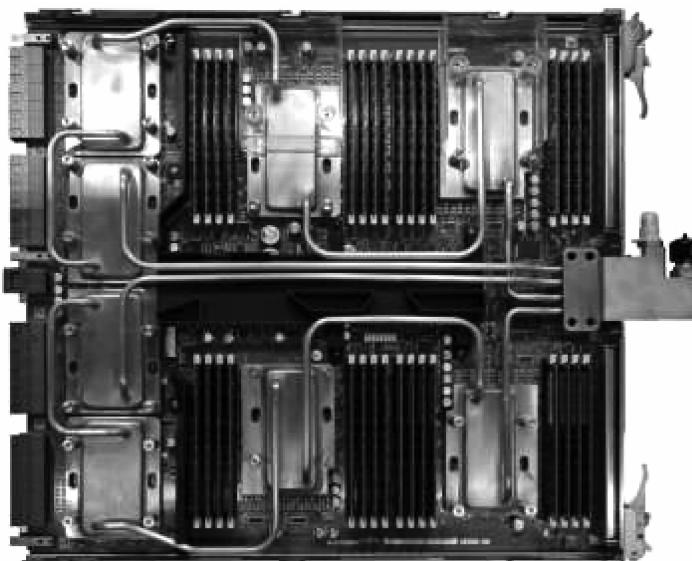


图 15 “京”计算机系统计算主板

2.2 计算机等冷却技术的前沿研究

针对电子设备冷却的严峻形势，DARPA 组织一些院校、企业和国家实验室开展的系列研究工作。项目名称为 HERETIC (Heat Removal for Thermo- Integrated Circuits)。项目简要情况如下：

- Embedded Droplet Impingement for Integrated Cooling of Electronics (Lead Institution: Carnegie Mellon University)。
- Integrated Thermal Management Using Laminate and Ceramic- MEMS Technologies (Lead Institution: Florida International University)。
- Microfluidic Technologies for Integrated Thermal Management: Micromachined Synthetic Jets (μ JETS) and Vida Heat Transfer Cells (Lead Institution: Georgia Institute of Technology)。
- Chip Level Thermal Management Using MEM Thermoacoustic Refrigerators (Lead Institution: Rockwell Science Center)。
- Electrokinetic Micro Coolers (Lead Institution: Stanford University)。
- Microfabrication Alliance for Innovative Cooling of Microelectronics (MAICOM) (Lead Institutions: Georgia Institute of Technology/University of Maryland)。
- Integrated Micro- Cooler Module for High Thermal Flux Removal (Lead Institution: University of California, Berkeley)。
- Modular Micromachined Si Heat Removal (MOMS- Heat Removal) (Lead Institution: University of California, Los Angeles)。
- Monolithically Integrated Thermoelectric Coolers for Mid- IR Laser (Lead Institution: Jet Propulsion Laboratory)。
- Integrated Microelectronics and Photonics Active Cooling Technology (IMPACT) (Lead Institution: University of California, Santa Cruz)。
- Heat Removal By Inverse Nottingham Effect (Lead Institution: University of North Carolina)。

EDIFICE 项目研究集成喷雾冷却系统，主要承担单位有卡内基 - 梅隆大学、Raytheon Systems 公司、Applied Thermal Technologies 公司及 IBM 公司。该项目旨在开发一个集成液滴喷射装置即 EDIFIC (Embedded Droplet Impingement for Integrated Cooling of Electronics)，研究目标是冷却热功率密度 $50 \sim 100\text{W/cm}^2$ 的军用或商用移动电子设备。主要特色在于采用高效相变冷却方式和器件控制在各局部按需要获得适量液滴。其基本思路是通过调节控制液滴的大小、频率及着点来实现局部动态温度控制。

Microfluidic Technologies for Integrated Thermal Management 项目与 EDIFICE 项目类似，也是要实现一个集成的对器件不同部位实现动态冷却。不同的是它立足于微观对流换热而不是相变冷却。该项目由佐治亚理工大学、Honeywell 公司和 Intel 公司合作承担，侧重

研究冷却部件的微加工和封装，特别是要研究一个旨在冷却局部热密度很高的激光二极管阵列的微喷头。

Monolithically Integrated Thermoelectric Coolers for Mid-IR Laser 项目旨在开发一个用于红外激光的集成热电制冷器以取代传统的冷冻液。项目的重点是单板集成的高性能热电制冷器的设计实现，它必须满足在恶劣条件下对激光器的动态冷却。

空气是绝大多数电子设备冷却的最终热沉，采用强制对流冷却是就需要风机等机械驱动部件。Chip Level Thermal Management Using Thermoacoustic Refrigerators 项目试图利用声波驱动空气，从而为提高组装密度创造条件。项目的关键技术包括狭小空间内热声效应的理论基础研究、高效声波发生器用压电材料及变送器研究以及热声核心部件小型化微加工技术研究。

Electrokinetic Micro Coolers 项目重点也是要开发流体驱动装置。鉴于微沟道对流换热具有体积小、换热效率高的特点，在电子设备冷却中具有很大潜力，但在系统实现时还需要有一个小体积高压头的泵。Electrokinetic Micro Coolers 项目的设计目标就是为一个冷却能力约 200 瓦的微沟道散热器配套一个耗电 1 瓦左右的电动泵。

为解决高功率密度的电子设备的散热问题，需要采用一些新的冷却技术。目前研究的热点技术主要有冷板技术、热管技术、热电制冷技术、微通道冷却技术、喷射冷却技术、热声制冷技术等。尽管在原理上它们都不是最近出现的技术，只是由于还没有在电子设备冷却中广泛应用但又具有很大潜力，我们姑且称为新型热设计技术。

冷板技术应用较早，包括风冷冷板和液冷冷板。由于管道内强制对流的研究比较充分，冷板的设计相对简单成熟。但冷板的应用主要受限于板级组装。从冷板的角度希望元器件平面组装，而从硬件性能的角度器件需要采用三维立体组装。同时冷板在解决局部高热流密度方面潜力有限。

热管技术可以用极低的温差传输大量的热量，在电子设备冷却中应用越来越广泛。目前普遍的使用方式是与风冷相结合，如“地球模拟器”的强迫风冷就应用了热管技术。热管技术正向大型化和微型化两个方向发展。通过应用更大的管芯传输更大的热量，微型热管则可以提高散热的功率密度。将微热管技术与芯片封装技术相结合，有可能突破器件散热瓶颈，考虑到可靠性和抗恶劣环境的能力，其市场化应用仍需假以时日。

热电制冷技术可以增大发热器件的传热温差，却存在制冷效率低的致命缺陷。再考虑到可靠性因素，该技术不大适合功耗几千瓦的大型电子设备。

微通道冷却技术可以在很小的空间实现高功率密度器件的散热，但由于流动阻力太大，需要很高的配送压力。相关外围配套设施的研发却大大滞后。当前微沟道热沉的应用主要集中在激光阵列的冷却上。

喷射冷却技术可以有各种形式。依据工质，可以分为气体喷射和液体喷射，依据传热机理，可以分为单相对流和蒸发相变。带相变的喷射冷却技术具有效率高、温差小、实现相对简单等诸多优点，在大功率电子设备的冷却中具有很大优势。

热声制冷技术是利用声波在传播过程中因压缩膨胀导致温度变化的原理工作的。目前该项技术还处于实验室研究阶段。由于空气比热小，该技术在大功率器件的冷却方面

不具优势。

总的来说，这些研究主要围绕大功率密度器件的高效冷却，即如何在较小的空间内带走大量的热量。项目的立足点大多基于微加工技术（如 MEMS）。主要冷却技术包括：喷射冷却、热声制冷、微沟道等。所有这些项目仅仅取得阶段性成果，走出实验室应用于工程实际仍需假以时日。

3 国内研究进展

3.1 国内典型高性能计算机系统的冷却技术

我国高性能计算和高效能计算已引起政府足够的重视。国家从“九五”到“十二五”通过 863 计划对该领域持续投入机器研制和应用环境建设。“十五”期间，“高性能计算机及其核心软件”重大专项研制两台万亿次以上的超级计算机，并以此为主要计算资源初步构建中国国家网格环境。2004 年研制成功的曙光 4000A 计算机达到了 10 万亿次浮点运算/秒的峰值速度，落户于上海超级计算中心。另一台 4 万亿次深腾计算机落户于中科院网络中心，加上国内其他 6 个院所的资源，初步形成了以上海超级计算中心和中科院网络中心为主结点的中国国家网格环境。“十一五”期间，设立了“高效能计算机和网格服务环境”重大专项。该专项明确提出研制千万亿次计算机系统，以及网格软件、网格服务环境与高性能计算应用系统，推动我国高效能计算机的研究、制造和应用。2010 年 11 月，我国国防科学技术大学研制成功“天河一号”超级计算机，采用异构可扩展阵列并行计算机体系结构，包含 CPU 和 GPU 共 23552 个，峰值性能达 4.7PFlops，Linpack 实测性能达到 2.566PFlops，位居 2010 年 11 月 TOP 500 排行榜第一，装备于国家超算天津中心。其后又相继建立了国家超算深圳中心、长沙中心和济南中心。经过“863 计划”连续支持，我国高效能计算机研制的整体技术水平已进入了国际先进行列。2012 年的典型系统见表 1^[10]。

表 1 2012 年中国 HPC 排行榜 TOP 5^[10]

排名	厂商	型号	安装地点	年份	处理器核	Linpack (Tflops)
1	国防科大	天河一号 A	国家超级计算天津中心	2010	202752	2566.00
2	国家并行计算机工程技术研究中心	神威蓝光	国家超级计算济南中心	2011	137200	795.90
3	国防科大	天河一号 A	国家超级计算长沙中心	2011	53248	771.70
4	曙光	曙光星云	国家超级计算深圳中心	2011	52416	749.20
5	中科院过程所	Mole-8.5 Cluster	中国科学院过程工程研究所	2010	33120	496.50

“天河一号”计算机使用的是高效的水风冷混合冷却技术。为了提高连接的可靠性，减小互连印制线长度，提高组装密度，减小系统占地面积，计算插框采用双面水平对插

组装。计算插框包括 16 块计算刀片、1 块 NRM 交换刀片、1 块 SMC 监控刀片、8 个电源模块、6 个风机模块、1 块计算互连中板、一组支持双面对插的机框。计算插框组装结构如图 16 所示。

计算单元互连通道与冷却通道分离以满足互连需要，前后为互连通道，左右为冷却通道，消除常规设计中互连插座对通风散热的影响，保证计算存储刀片内功率器件都能得到均匀流畅的冷却气流。

为了满足安全和电磁兼容的需要，电源模块必须设计成金属包络的整体，势必是通风散热的阻碍。通过电源模块与计算存储刀片交错布局，在保证可靠供电的同时预留足够的通风通道，在极小的空间实现 AD/DC 转换、配电、冷却和电磁兼容。每个电源模块提供 3 个计算存储刀片所需的电能，布置在相邻的两个计算存储刀片中间，在水平方向上通过中板实现与计算存储刀片的电连通，上下方向相邻电源模块间保留足够空间作为冷却空气的流动通道。

每个风机模块 3 个直流风机， $2+1$ 冗余设计，以吹风方式实现对应计算存储刀片的散热。冷却风扇根据对应计算存储刀片的功耗特征与运行温度自动调整转速，在满足可靠性要求的同时减小风机功耗。4 个电源模块 $3+1$ 并联冗余，每个电源模块高度 44mm，电源模块与计算存储刀片交错 22mm，相邻电源模块间距 44mm，为每个计算刀片预留 22mm 高度的通风空间。旋转对称封闭循环的冷却架构如图 17 所示。

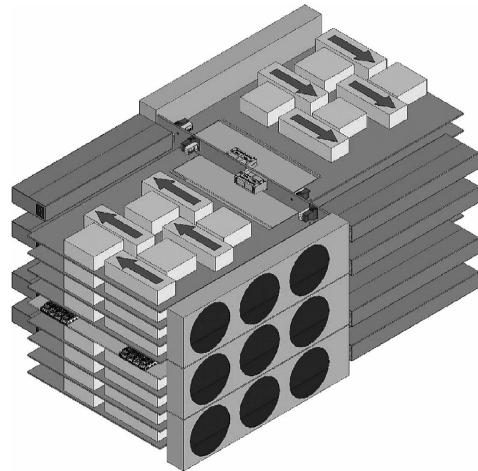


图 16 “天河一号” A 计算插框组装结构

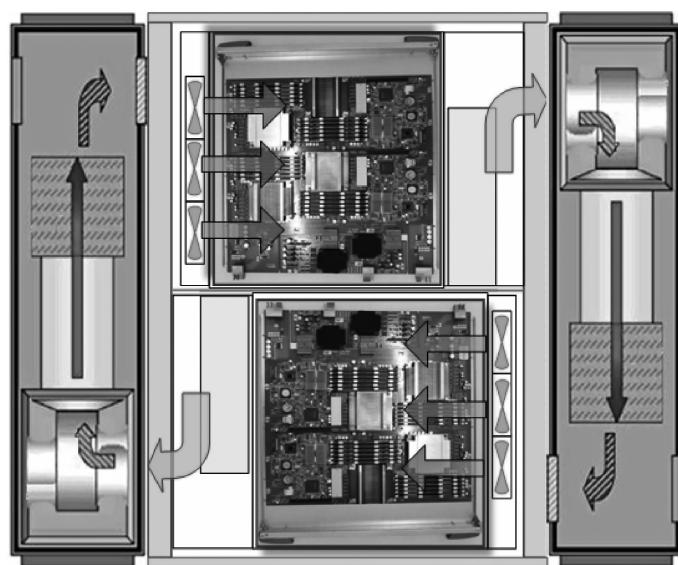


图 17 “天河一号” A 计算单元旋转对称封闭循环散热方案

计算机柜组装热密度超过 50 千瓦，采用常规通风散热模式，开孔地板或通风格栅都很难提供足够的循环风量，需要庞大的面积安装机房空调系统和送风通道。结合机房空调技术的发展，瘦结点机柜采用列间空调实现散热，如图 18 所示。

每个计算机柜利用 2 台冷冻水列式空调，与计算机柜构成一个封闭的组合体。从右侧的冷冻水空调流出的冷空气，在冷却风机驱动下由右向左流过计算存储刀片，吸收功率器件的热功耗后变为热空气，通过电源模块间的间隙后流入左侧的冷冻水空调，冷冻水空调的风机强制热空气流过空调的水 - 空气换热器后又恢复为冷空气，冷空气在背部风机模块驱动下实现对背部计算存储刀片的散热并升温后，通过背部电源模块间间隙流回左侧的冷冻水空调，在该空调内换热器上实现与冷冻水的热交换变为冷空气后流出。如此循环往复，形成一个顺时针循环气流。机柜内冷却气流组织如图 19 所示。

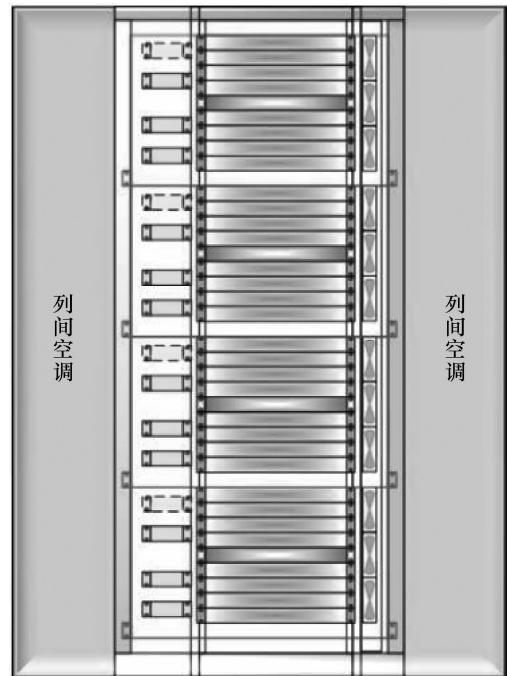


图 18 “天河一号” A 计算机柜散热方案

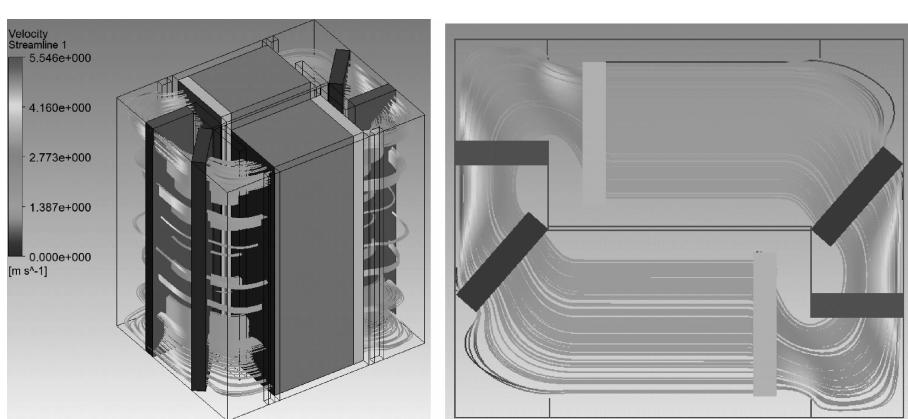


图 19 “天河一号” A 计算机内部冷却气流模拟图

计算机柜与列式空调构成封闭体，冷却空气在封闭体内循环，避免了冷却气流旁通，相对于常规开放式冷却结构，降低了冷却系统功耗。空调与计算机柜紧密相邻，输出的冷风很快就被计算机柜利用，减小了冷却空气输送距离，降低了空调风机运行功率。冷却空气在封闭组合体内循环，不仅降低了噪声，而且放宽了机房温湿度要求，减小了机房空调系统的热负荷。

中科院曙光在中国 TOP 100 排名中数量占据第一名，“魔方”“星云”是曙光代表性

的高性能计算机，其中“星云”是我国第一台实测性能超千万亿次的超级计算机。

曙光 5000A 高性能计算机是国家 863 计划高性能计算机及其核心软件重大专项支持的研究项目，是面向网格的高性能计算机，可以为网格提供计算服务，同时也是面向信息服务的超级服务器，可以提供多目标的系统服务。系统于 2008 年落户“上海超级计算中心”，该系统的计算结点机采用四路四核 AMD Barcelona（主频 2.0GHz）处理器，每个结点 64GB 内存，胖结点机采用 8 路 4 核 AMD Barcelona（主频 2.0GHz）处理器，每个结点 128GB 内存，整个系统由 1650 个结点组成庞大机群，共有 6000 颗 CPU 与 100TB 海量内存，其理论浮点峰值为每秒 230 万亿次。为提高曙光 5000A 的系统密度，研制了具有自主知识产权的 4CPU 4 核高密度刀片主板（参见图 20）和基于该主板的世界上第一款 7U 高度的 10 片 4 路 4 核 64 位 CPU 的工业标准刀片服务器。该刀片服务器采用全内置的 ConnectX DDR 二层交换模块，全内置的管理 Ethernet 网交换模块，可以实现单机箱 40CPU，160 核的超高计算密度，并实现了单结点内部的并行计算。



图 20 曙光 5000A 计算刀片和刀片服务器

曙光 5000A 采用了机柜级的水冷系统，通过使用水冷系统，降低了高性能计算系统对客户机房的要求，同时提高了冷却效率。水冷系统的设计实现了水电分离，保证了系统的安全性。系统采用了一体化的水冷机柜 CoolTherm，机柜底部有一个 V35 高性能空气 - 水热交换器，背部有 3 个离心风机。刀片服务器排出的热风由机柜风机导入机柜底部的热交换器，降温后输送到机柜前门与刀片服务器之间，机柜内的气流组织如图 21 所示。

“星云”是由曙光公司、中科院技术研究所、国家超级计算深圳中心共同研制的千万亿次超级计算机系统，理论峰值 3000 万亿次，实测峰值每秒达 1271 万亿次，采用了自主设计的 HPP 体系结构，使用了 Intel Xeon 5650 通用处理器 + Nvidia Tesla C2050 加速处理器。星云由星云由大约 464 个 TC3600 服务器通过 IB 交换机连接而成。每个 TC3600 刀片服务器支持 10 个高性能刀片、2 个高速交换模块、2 个管理模块、4 个电源模块、4 个风机模块，如图 22 所示。



图 21 CoolTherm 水冷机柜



图 22 曙光 TC3600 刀片服务器

星云的一个计算机柜可以安装 4 个 TC3600 服务器，单机柜最多组装 80 个通用处理器和 40 个加速处理器。星云亦采用风冷 + 水冷的混合冷却方式散热，每个机柜侧面配一个 CoolLoop 冷冻水空调。TC3600 采用的是前后通风方式，热风从背面排除，通过 CoolLoop 水冷空调变成冷风送到服务器前面循环封闭运行，单机柜冷却能力可达 30 千瓦，如图 23 所示。



图 23 “星云”高性能计算机系统

神威蓝光（Sunway BlueLight MPP）由国家并行计算机工程技术研究中心制造，系统全部采用自主设计生产的 CPU（ShenWei processor SW1600）共 8704 个 CPU，峰值 1.7016PFlops，持续性能 795.9TFlops，总功耗 1074 千瓦。系统于 2011 年 9 月安装于国家超算济南中心，由山东省科学院计算中心负责建设、管理和运营。

神威蓝光采用水冷技术散热。一块铝制水冷冷板、冷板两面各贴装 2 块双 CPU 计算子板，构成 8CPU 的计算结点，如图 24 所示。冷板侧面安装有一对快换接头供接入冷水。

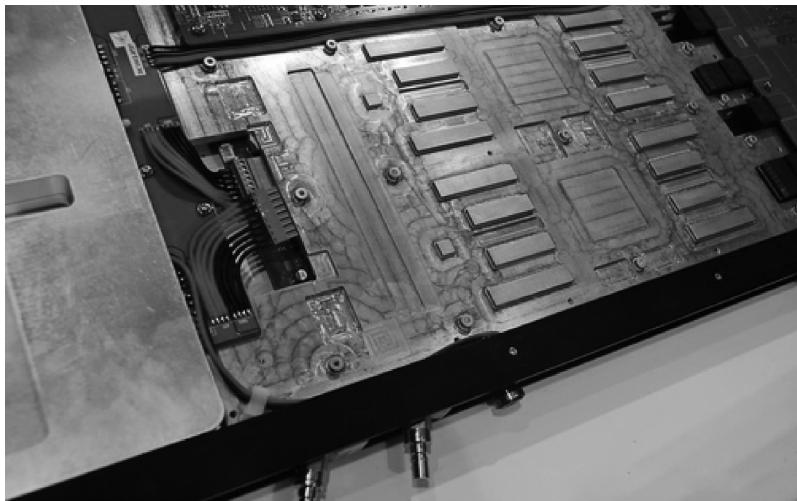


图 24 “神威蓝光”水冷计算结点

3.2 国内冷却技术的相关研究

近几年，中国科学院电工研究所展开了将具有我国自主知识产权的蒸发冷却技术应用于超级计算机的试验与应用研究。蒸发冷却作为一种高效冷却技术，它是依据冷却介质蒸发时的汽化潜热来带走发热体的热量，使其降温。运行时，发热体将热量传递给蒸发冷却介质，蒸发冷却介质吸收热量，温度升高，当达到一定压力下的饱和温度时，开始沸腾，介质汽化吸热，汽化后的介质蒸汽在冷凝器中将热量传递给二次冷却系统，同时介质蒸汽被冷却为液体，回到冷却系统中，继续循环，形成密闭式循环系统。

将蒸发冷却技术应用于超级计算机不仅可以实现对高集成度芯片的良好冷却，同时因蒸发冷却介质具有良好的物理、化学稳定性和绝缘性能，系统的安全性、可靠性也可得到有效保证。

2009 年中国科学院电工所承担中国科学院重大装备研究项目“超级计算机蒸发冷却系统研究”，首次将蒸发冷却技术应用于超级计算机。在该项目中，成功研制出了喷淋式蒸发冷却超级计算机的真机样机和热模拟样机，从结构可行性，冷却效果，运行可靠性以及巨大的带热潜力等多方面验证了蒸发冷却技术应用于超级计算机的优势。图 25 所示为喷淋式蒸发冷却超级计算机



图 25 喷淋式蒸发冷却超级计算机的 D 真机样机

的真机样机。此外，在该项目中，中科院电工所研制的低温强迫循环蒸发冷却系统在中科院过程工程研究所 GHPC1000 超级计算机系统上成功进行了负荷试验和长期运行测试。结果显示，机柜内计算结点满负荷工作状态下，完全满足冷风温度为 18~22℃ 的设计要求，系统的安全性、可靠性也可得到有效保证。该项目已于 2012 年 6 月 6 号通过项目验收。

除以上的喷淋式和低温强迫蒸发冷却方式，中科院电工所目前正在研制的还有浸泡式和贴壁式等自循环蒸发冷却技术，其节能优势更明显，且具有更强的运行自调节能力，有望成为未来最新的高效、安全、节能的高性能计算机系统冷却技术。

在对于中科院过程工程研究所 GHPC1000 超级计算机的冷却系统改造项目中，中科院电工所研制的低温强迫循环蒸发冷却系统初次显示出了蒸发冷却技术应用于高性能计算机的可行性、安全性、可靠性。之后，中科院电工所研制出了喷淋式蒸发冷却超级计算机的真机样机和热模拟样机（冷却负荷 50 千瓦）各一台，其运行效果充分体现了冷却效率高，热分布均匀，低噪声，微功耗，可靠性高，基本免维护，冷却介质用量少等技术特色。当前，中科院电工所正着力于多台冷却负荷更高的喷淋式蒸发冷却超级计算机的研制工作，与此同时，中科院电工所还针对计算机冷却进行了新型节能高效冷却技术的拓展研究，如浸泡式蒸发冷却技术，贴壁式蒸发冷却技术等。

浸泡式蒸发冷却技术是将计算机整体浸泡在冷却介质中，使计算机部件产生的热量直接传递给冷却介质，冷却介质通过汽化相变换热的方式将热量传递给位于顶部的冷凝管，然后再由空气带走，图 26 所示为单刀片的浸泡式蒸发冷却展示模型。该冷却方式可去掉原计算机中的所有风扇，完全避免了由风扇引起的噪声问题；同时降低了系统的冷却能耗，节能高达 30%。而且由于采用全密封结构，即使在恶劣环境（高温、高湿、灰尘）中也可以保证正常运行。

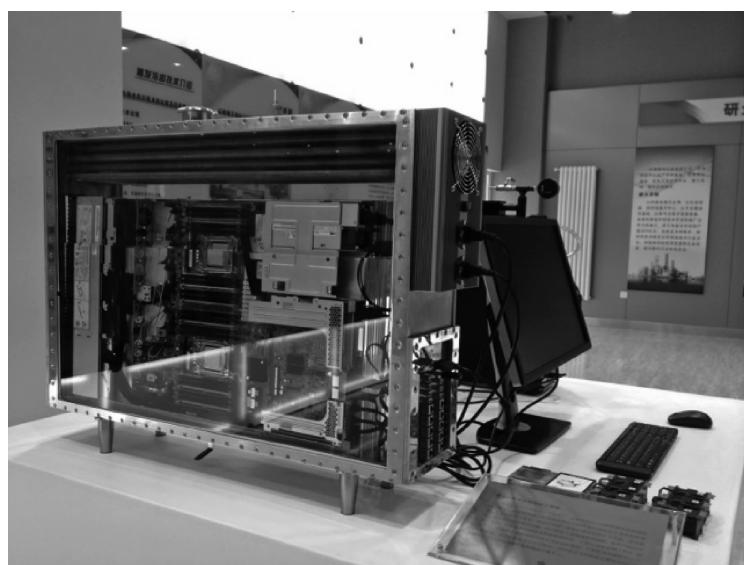


图 26 单刀片的浸泡式蒸发冷却展示模型

贴壁式蒸发冷却技术是指冷却介质不与计算机直接接触，而是在计算机主板的主要发热部件上固定一个结构尺寸与其相匹配的液盒，发热部件产生的热量主要通过热传导传递至液盒贴附面上，再由液盒传递至冷却介质，冷却介质受热后通过相变换热的方式将热量传递给机箱外部的冷凝器，最后由空气带走。贴壁式蒸发冷却系统中的液盒可以作为标准构件，对计算机主板高发热部件进行有针对性的冷却，适用范围较广，可应用于 PC、服务器工作站，也可应用于大型超级计算机，系统建立较为简单，且介质用量很少。图 27 所示为贴壁式蒸发冷却 PC 样机。



图 27 贴壁式蒸发冷却 PC 样机

作为新型节能高效的计算机冷却技术，浸泡式和贴壁式蒸发冷却技术均利用冷却介质吸收的热量做功，循环的动力来自系统内的热工转换，实现了无泵自循环，因而节能效果更加突出。相对于传统风冷技术，这两种冷却方式均可通过取消原计算机中主要发热部件上的风扇，解决由风扇引起的噪声和能耗问题，而且由于蒸发冷却技术在换热效率上高于传统冷却方式，计算机发热部件可以长期工作在 50~60℃，远远小于各种器件的最高允许运行温度，计算机的可靠性大幅提高，使用寿命也相应延长。

4 国内外研究进展比较

4.1 研发模式的比较

从研发模式上看，国外对高性能计算机冷却技术的研究更系统更有组织性。一方面，从事高性能计算机研制的大型商业公司如 IBM、Cray、HP 等联合冷却散热相关公司如

Emerson、Paker、3M 等针对具体项目开展联合研究；另一方面，相关组织、协会根据行业发展现状和趋势征集和发布研究课题，组织不同类型的研究机构开展不同方向的冷却技术研发。比较而言，国内的相关研究尤其是基础性研究更偏重于个体，缺乏有组织的管理和持续的发展。

4.2 研发内容的比较

从研发内容上看，国外对高性能计算机冷却技术的研究更全面。从芯片内微观尺度的散热到全系统绿色数据中心的建设，从传统冷却技术的改良到新型冷却技术的创新，从实用性强的新技术到概念型的新探索，国外对计算机相关冷却技术的研究更深入更全面。受限于产业发展水平，国内在处理器内部微观散热的研究仅限于理论研究，宏观冷却技术（如液冷、水风冷混合冷却技术）的研究与实现更成熟。

4.3 研发水平的比较

在宏观冷却技术上，无论是水冷、风冷、水风冷混合冷却还是蒸发冷却，无论是器件级、插件级、机柜级的散热能力，我国的研究水平都不落后于美国等国外高性能计算机系统。但在芯片内散热，特别是面对未来 3D 封装、SoP 的散热技术研究方面，我们缺乏相应的背景支撑和需求驱动，研究水平大大落后。

总之，近几年随着我国高性能计算的飞速发展，对高性能计算机系统冷却方面的投入不断增多，在空冷、水冷和混合冷却等取得了较大成果。在新型冷却技术研发方面，我国也展开了一定规模的研究，一些高效的新型冷却技术初露头角，并显示出了相当的应用潜力。目前国内高性能计算机散热冷却技术的发展情况，已经接近或处于国际一流的技术应用水平。但对未来微观尺度冷却技术的研究有待加强。

5 发展趋势与展望

预计到 2020 年以前，高性能计算机系统的主要元器件仍将基于硅工艺。随着芯片集成度的不断提高、工艺尺寸的不断缩小、工作电压的不断降低、信号速度的不断升高，高性能计算机热可靠性将面临诸多的不确定因素。研究分析未来芯片的热失效机理、明确未来芯片的工作温度应力阈值无疑是高性能计算机冷却技术研究的基础与方向。

预计未来单芯片的热设计功耗将提高到 200 瓦左右，单机柜的热功耗在 100 千瓦左右，全系统的功耗在 20 兆瓦左右。高性能计算机散热冷却技术未来的发展将关注冷却系统冷却效率提升、冷却系统与电子设备的融合、冷却系统的安全可靠运行以及冷却系统对环境的影响等方面。

5.1 芯片液体冷却新技术的持续发展

在芯片液冷技术中水冷技术目前在高性能计算机中应用较多，水冷系统器件的安全可靠性和与外部热量交换的方式将会有更多的研究和探索。水冷芯片技术中冷板的材料研究、液体流道的强化换热设计、冷板与芯片之间的表面传热材料（Thermal Interface Materials）的研究及选择是提高水冷系统传热效率关键的技术研究方向。液体冷却技术的应用将显著地降低高性能计算机运行的能耗。

热管换热技术也是芯片液体冷却的重要发展方向，未来热管冷却技术的热点研究问题将主要集中在：1) 提高处理器芯片散热能力；2) 改善热管技术；3) 取消整体式的散热片（IHS），降低热阻；4) 改善蒸发流道；5) 以泵驱动的微通道两相换热循环。

5.2 液体冷却系统与空调系统的有机结合

未来高性能计算机运算速度和规模仍旧不断扩大，而为了更有效地冷却计算机内的电子器件，冷却系统将更加深入到计算机内部，甚至深入到芯片内部，液冷冷却系统将热量从芯片的热量传递到服务器外部，而从服务器外部再将热量转移到室外有很多形式上的选择，通过水系统或风系统都可以转移这部分热量。对于计算机内部发热量较小的电子元器件，仍将沿用传统的风冷的形式进行冷却。液体冷却系统与传统空调系统的联合应用，特别是与带有节能技术的空调系统的有机结合，将进一步提升冷却系统的效率。

5.3 冷却系统的节能和绿色化发展

高性能计算机对环境的影响主要为耗费大量电力能源和产生巨大的热量，进一步降低计算机能耗和有效利用计算机产生的热量将是未来高性能计算机发展追求的目标。随着新冷却技术、新冷却方案的不断涌现，直接液体喷射冷却、液体浸渍冷却等新技术，冷却系统热量回收、自然新风冷却等新应用方案的使用高性能计算机将走上对环境影响更友好、更加绿色化的发展之路。

5.4 冷却系统的商业化发展

高性能计算机除追求运算速度的极限外，其应用范围也在不断地扩大，已不仅局限在研究、军事和实验室等科研机构应用，越来越多的大型商业公司也有对高性能计算机的需求。新的冷却技术要适应商业公司的需求，商业用户更加关注产品的安全可靠，架构的灵活性和成本可控性。

6 结束语

随着高性能计算机技术的发展，电子组装密度的迅速增加带来了高热流密度与冷却技术之间的矛盾成为散热技术发展的主要前进动力。高性能计算机散热冷却技术已成为高性能计算机技术发展的重要组成部分，散热冷却系统的架构会影响到计算机的物理架构设计。冷却技术伴随着计算机技术的发展而不断变化，从最初使用机房空调进行冷却，到定制化的冷却终端空调从计算机内部进行冷却，再到芯片级的液体冷却以及液体喷射蒸发冷却、液体浸没冷却等新技术的应用。性能与成本则是左右这些变化的指挥棒。

冷却技术的进步能够显著提升高性能计算机的运行效率，由于芯片液冷技术在冷却组件与芯片的热阻更小，冷却效率得到显著提高，所以液冷技术能够冷却更高的热密度，芯片液冷技术将会成为未来几年计算机冷却技术中重要的发展方向。

近年来我国在高性能计算机冷却技术的研究和产业化取得了长足进步。建议组织相关科研单位和生产单位，紧密合作大力推进液冷等高效冷却技术的产业化和商业化，有了市场化的推动作用，才能使我国高性能计算机的冷却技术与国外技术处于同一水平或领先水平。

参考文献

- [1] 王普勇, 桂亚东, 王涛. 支撑科技创新的高效能计算[J]. 科学(上海), 2009, 61(5): 5-10.
- [2] 杨学军. 并行计算六十年[J]. 计算机工程与科学, 2012, 34(8): 1-10.
- [3] P Coteus, H R Bickford, T M Cipolla, etc. Packaging the Blue Gene/L Supercomputer [J]. IBM J. RES. & Dev., 2005, 49(2): 213-248.
- [4] Srini Chari. Blue Gene/Q: The Most Energy Efficient Green Solution for High Performance Computing[R/OL]. [2013-06-23] <http://www-03.ibm.com/systems/deepcomputing/bluegene>.
- [5] Klaus Gottschalk. SuperMUC HPC System[R/OL]. [2013-06-23] <http://www.lrz.de/services/compute/supermuc>.
- [6] Dorian Gahm. The Cray ECOPhlex™ Liquid Cooled Supercomputer Offers Energy Advantages to HPC Users [R/OL]. [2013-06-23] <http://www.cray.com>.
- [7] David Tanqueray. The Cray X1 and Supercomputer Roadmap [R/OL]. [2002- 12- 12] <http://www.cray.com/products/systems/x1/>.
- [8] 李利, 陈晓苏. 地球模拟器的硬件技术[J]. 高性能计算技术, 2004, 166(2): 6-10.
- [9] Motoi Okuda. Fujitsu's Approach to Application Centric Petascale Computing: 14th Workshop on Use of High Performance Computing in Meteorology[C].
- [10] 张云泉, 孙家昶, 袁国兴, 等. 2012年中国高性能计算机发展现状分析与展望[J]. 计算机工程与科学, 2012, 34(2): 1-8.

作者简介

陈 旭 国防科学技术大学计算机学院副研究员，主要研究领域为计算机组装技术、数据中心冷却技术等。



王 峰 艾默生网络能源有限公司高级工程师，主要研究领域为机房精密空调、绿色数据中心技术等。



阮 琳 博士，研究员，中国科学院电工研究所蒸发冷却技术研究发展中心主任，博士生导师。主要研究领域包括大电机自循环蒸发冷却技术、IT设备高效蒸发冷却技术等。



SDN 技术发展报告

CCF 互联网专业委员会
CCF 网络与数据通信专业委员会

摘要

软件定义网络（SDN）是一种新型网络架构，将传统网络设备中的网络控制功能与数据转发功能相剥离，把控制层面迁移到可控的计算设备中并采用集中式的控制，采用标准接口协议衔接控制平面和数据平面。控制平面和数据平面分离的思想降低了网络设备功能和架构的复杂性，简化了网络管理和配置操作，实现软件定义网络行为。本文介绍了 SDN 的基本思想，分析了 SDN 技术面临的主要问题和挑战，从控制平面、数据平面和 SDN 应用等方面介绍了国内外的研究进展。最后从控制平面的扩展性、灵活可编程的数据平面机制、高性能转发平面实现技术、基于 SDN 的网络虚拟化和 SDN 创新应用等方面对 SDN 未来发展趋势进行展望和总结。

关键词：软件定义网络，OpenFlow，控制平面，数据平面，集中控制

Abstract

Software Defined Network (SDN) is a novel network architecture, which decouples network control functions from the traditional network devices, and moves them to a centralized cluster to enable centralized control of the network. In SDN, an open and unified control interface is provided as the communication protocol between the centralized control plane and individual data planes. The separation of control planes and data planes facilitates to reduce the functional and architectural complexity of network devices, to simplify network management and configuration, and to enable software defined network behavior. This report introduces the fundamental idea behind SDN, analyzes the major challenges in SDN technologies, and details the SDN research progress in terms of control planes, data planes, and SDN applications. At last, this report shows the SDN development trend and the remaining research challenges in terms of the control plane scalability, data plane flexibility, implementation techniques of high performance data planes, SDN based network virtualization, and SDN based innovative applications.

Keywords: software defined networking, OpenFlow, control plane, data plane, centralized control

1 引言

互联网从诞生至今通过承载各种分布式应用及容纳各种异构网络，取得了巨大的成功，成为推动社会和经济发展的巨大动力^[1]。随着网络新技术、新应用的不断出现，网络服务日趋多样，网络规模不断扩大，网络用户日益增多，当前互联网环境已经发生了

巨大变化，传统互联网的体系结构和设计理念日益暴露出严重不足。一方面，网络服务多样化要求打破传统网络尽力而为服务模式，提高对网络的管控能力，以满足用户多样化的传输需求；另一方面，作为一个庞大的复杂系统，各种协议的决策逻辑交织在一起而产生作用使网络行为呈现出相当的复杂性并且难以预测。这些问题存在的根本原因在于网络核心设备控制逻辑与数据转发之间紧耦合以及核心设备之间的分布式控制。研究人员提出应重新设计网络控制和管理结构，将控制逻辑与数据转发相分离，采用集中控制的方式，从而降低网络核心设备的复杂性，提高网络控制和管理的灵活性，增强对网络新技术、新协议的支持能力。这方面的代表性工作包括：IETF 提出的转发与控制相分离的 ForCES 模型^[2] 及其扩展 OpenRouter^[3]、Greenberg 等人提出的集中式网络控制和管理架构 4D 模型^[4]、Caesar M 使用逻辑中央平台实现的 BGP 路由决策集中管控架构 RCP^[5]，以及美国斯坦福大学 Clean Slate 项目组基于 4D 原则提出的企业网环境下集中管控架构 SANE^[6] 和 Ethane^[7] 等。这些成果多局限于理论性和框架性的设计，虽然对新一代网络体系结构研究有所启发，但由于缺乏业界支持，无法进行较大规模真实网络环境下的验证，导致这些成果未能获得更大范围的认可。

为了解决网络新技术实验验证问题，斯坦福大学 Clean Slate 项目组提出的 OpenFlow 技术^[8] 逐渐引起了人们的关注。该技术将网络的控制逻辑从网络设备中剥离出来，使研究者可以通过一组定义明确的接口对网络设备进行编程，从而方便实现新型的网络协议和拓扑架构，而无需改动网络设备本身，为研究者在真实网络环境下进行网络新技术、新业务的创新实验提供了基础。随后，该项目组采用 OpenFlow 作为实现技术对之前的 SANE、Ethane 等研究工作进行了扩展，提出了软件定义网络（Software-Defined Networking, SDN）^[9] 的概念。SDN 是一种新型网络架构，将传统网络设备中的网络控制功能与数据转发功能相剥离，把控制层面的功能迁移到可控的计算设备中并采用集中式的控制方式；采用标准接口协议衔接控制层面和传输层面，使得网络控制软件化、可编程化，通过开放的可编程软件模式实现网络控制功能的多元化。网络逻辑控制和数据转发分离的思想简化了网络管理和配置操作，保证了控制逻辑的健壮性，而开放式的软件编程模式使得网络控制具备超强的敏捷性，同时可以减少控制开销，集中先进技术实现网络管理的最优化目标。

2009 年，SDN 技术入选美国 MIT 主办的《技术评论》杂志评选出的十大新兴技术之一^[10]，自此获得了学术界和工业界的广泛认可和大力支持，并在许多研究新一代网络体系结构的著名科研项目中得到广泛应用和部署，如美国的 GENI^[11] 和 Internet2^[12]、欧洲的 OFLEIA^[13]、日本的 JGN2 plus^[14]、韩国的 NetOpen^[15] 等。在工业界，SDN 技术也得到了许多知名 IT 企业的大力支持和推广。Cisco、Ericsson、华为等网络设备制造商都认可 SDN 将成为网络演进的方向和手段，并已经发布了有关系统^[16]；谷歌、FaceBook、百度也都开始在其数据中心部署应用了 OpenFlow/SDN 技术^[17]。目前，全球商业界已经普遍认可 SDN 技术的价值，国际著名咨询公司 Gartner 在 2012 年 10 月将 SDN 列为未来五年 IT 领域的十大关键技术之一，著名网站 InfoWorld 在 2011 年年底公布的可能影响未来十年的十项新技术中 SDN 排列第二，IDC 预测到 2016 年 SDN 产业产值将达 20 亿美元。

综上所述，SDN 技术描绘了新一代网络的演进前景，其倡导的传输层面与控制层面相分离、集中式控制体系结构、网络控制器接口开放、网络虚拟化等核心理念不仅促进了新一代网络体系结构的研究工作，同时也为开放、创新的网络技术研究提供了一个有效的实验平台。本报告将从 SDN 的基本思想和架构、SDN 的控制平面与数据平面、SDN 的应用及国内外研究进展三个方面对 SDN 技术的研究进展情况进行探讨。剩余部分安排如下：第 2 节主要阐述了 SDN 主要思想与架构、主要挑战与关键技术问题；第 3 节和第 4 节分别介绍 SDN 的控制平面和数据平面功能、技术问题及研究进展情况；第 5 节介绍了 SDN 技术在流量工程、网络虚拟化、网络节能等方面的应用情况；第 6 节介绍了国内相关研究进展情况；第 7 节对全文进行了总结并对未来的发展进行了展望。

2 软件定义网络的基本思想和架构

2.1 SDN 主要思想和架构

SDN 是一种创新的新型网络架构。其核心思想是：数据与控制平面分离；统一的厂商无关控制与数据平面开放接口；逻辑集中的控制平面；顶层服务或网络控制应用通过物理资源的抽象构建逻辑的全网视图；通过分片和虚拟化底层网络实现资源的优化利用和调度。SDN 的最终目标是无论是研究人员还是网络管理人员或者是第三方基于软件控制方式实现对网络结点和网络集中统一、多抽象层次的灵活控制。

OpenFlow 是 SDN 最重要实现方案，分别由 OpenFlow 控制器和交换机构成，之间通过标准化 OpenFlow（OF）协议通信（如图 1 所示）；其中，控制器以集中方式控制本管理域内的多个交换机，通过提供北向接口来实现业务逻辑，支撑多种网络业务创新研发，诸如流量控制、路由控制、安全控制等功能。现有 SDN/OF 多纳入虚拟化技术，如以 OpenStack 为首的云计算平台也已广泛采用了 SDN 技术为业务提供网络支撑环境，Nicira 以虚拟化技术为基础构建第一个 SDN 操作系统。

上述 SDN 体系结构实现物理设施与功能的分离，带来如下几方面的技术优势：

- 1) 开放网络创新能力：平面间标准化的数据面配置协议（即 OpenFlow 协议）实现控制面与数据面分离，实现廉价、水平可伸缩和开放的网络体系结构（如图 2 所示）替代传统昂贵、垂直集成和封闭的路由器体系结构，避免底层网络的复杂性，使得新的网络业务能够快速构建和测试；
- 2) 网络控制可伸缩性、灵活性和可编程能力：SDN 域内数据面结点和域间结点通过

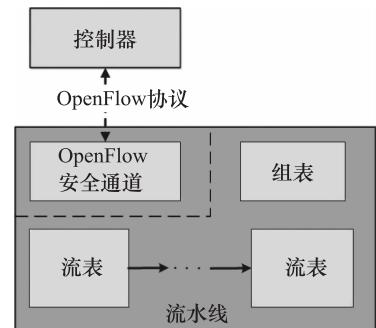


图 1 OpenFlow 结构

集中全局控制避免传统动态路由控制系统所带来的局部性问题，如域内路由振荡、路由环路、路由黑洞、流量控制局部最优以及路由收敛过程路由结点上控制状态不一致等问题；

3) 网络安全性与可管性：现有 BGP、OSPF、IS-IS 和 RIP 等协议的控制的稳定过程难以确保安全性，而 SDN 有助于避免现有网络中出现的 Sybil、RIB Poisoning 和 DoS 攻击等问题，并通过控制策略的实施实现端到端的资源管理；

4) 数据面可扩展性：SDN 允许控制面通过对流表动态编程实现交换机转发行为的动态控制，现有相关研究表明在 SDN 上扩展 IP 协议或新的非 IP 协议更容易；

5) 网络测量感觉能力：OpenFlow 交换机流表项能够快速响应数据面流的高度动态性（通过在交换机中优化其中 Heavy-hitters 项），通过准确测量，快速响应网络的异常问题和动态变化，提高网络控制系统实时性。

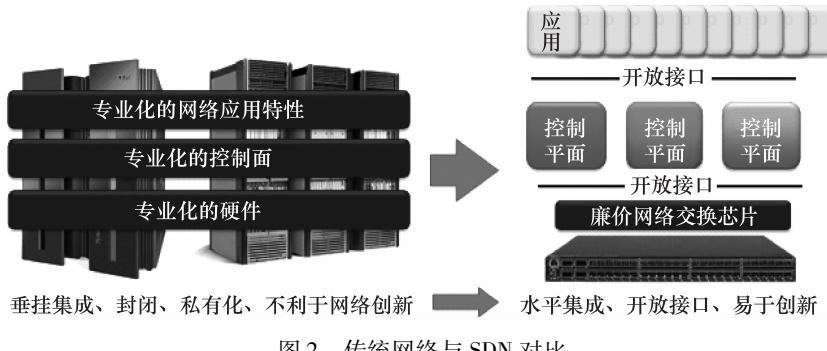


图 2 传统网络与 SDN 对比

2.2 研究问题与主要挑战

在 SDN 体系结构下，需解决其体系结构问题以满足演进性需求，其中 SDN 的核心问题包括：

1) 控制集中化。无论控制层还是应用层之下都有几十个乃至几百个被管理/控制结点，不仅要保证控制层和应用层的正常运行还要维持其高效运行。比如，SDN 的每一次路由/策略决策都涉及路径上的多个结点，需要与多个转发面结点交互下发流表/转发表，系统规模越大，平均转发路径越长，一次决策需要交互的结点就越多。集中统一管理对控制层和应用层带来一些问题，包括可伸缩性、可扩展性和高可用性等。

2) 内容虚拟化。通过虚拟化实现各个逻辑层，建立逻辑层与物理层以及不同逻辑层之间的分离和映射，使网络运营商和使用者可以摆脱网络分布式复杂性、物理设备的具体实现细节、网络协议的部署方式而轻松进行管理和使用。虚拟化包括多虚一和一虚多两块内容，前者是指异构资源如何有效统一和确定抽象程度（颗粒度）的问题，后者是业务的部署及虚拟网络映射，实现不同逻辑层面之间的映射（把上层的需求分解到下层）或重构（用下层来支撑上层的需求）。

3) 接口开放化。“开放”是指网络中各个模块及相互接口的定义和通信协议的对外

公开甚至进一步统一标准化。传统网络体系架构在开放性方面给灵活性带来了最基本最严重的制约。以网络设备为例，传统网络设备在体系结构上都是封闭的，即一个网络设备通常只能由一个厂家生产，包括软件和硬件、网络底层技术等都由一个厂家拥有，用户只有有限的按照厂家提供的软件对设备进行配置的灵活性。就其形式而言，从传统的命令行接口（CLI）和 SNMP 到 XMPP（eXtensible Messaging and Presence Protocol）、Netconf、OpenStack 以及各种虚拟化软件的 API 都是可能的。

4) 网络自动化。这是 SDN 中重要的配套和支撑技术。如果没有必要的网络自动化，由于更频繁的重新配置、业务连续性和灾难恢复风险，以及 IT 人员不堪手动工作的重负等问题的存在，基础架构中的纰漏将会导致网络的不稳定。网络自动化的手段包括数据可视化、性能分析预测、自动代码生成、软硬件模块的可编程性、信息智能化处理。

2.3 SDN 实例

OpenFlow 是在 2008 年 3 月由 Nick McKeown 等人提出并在斯坦福大学成立了 OpenFlow 论坛，它是软件定义网络的一个实例，是第一个遵循软件定义网络架构的控制面/数据面协议，其架构如图 1 所示。控制器位于控制层，由软件实现，支持 OpenFlow 协议的 OpenFlow 交换机位于基础设施层，由网络设备实现。

OpenFlow 协议由控制面和数据面同时提供接口实现。OpenFlow 使用了广义流的概念：符合某些可识别特征的所有的数据包的集合。这些特征主要是数据包的头部字段，如 MAC 层源地址和目的地址、IP 层源地址与目的地址、VLAN 标签以及用户自动字段 TLV（Type Length Value）等。流是对网络数据包特征的高度抽象，流表代替路由表并提供更多功能，以流为单位来指定数据包的处理比以 IP 目的地址指定数据包的处理更具灵活性。

支持 OpenFlow 协议的数据转发设备称为 OpenFlow 交换机，OpenFlow 交换机由一个或多个流表、一个组表以及安全通道组成。流表是交换机支持 OpenFlow 的最关键技术，它由多个流表项组成，每个流表项代表一个流，定义了匹配和处理数据包的规则，这些流表项都由控制器写入。每个流表项由三部分组成：特征域，指令集和计数器。特征域包含用来识别流的所有域，指令集指定了与该流表项匹配的数据包将被如何处理，计数器记录每个流表项的一些统计信息。组表定义了比流表更加复杂的转发语法，组表由多个组表项组成，每个组表项由一系列操作桶组成，可以根据每个组表项中定义的组类型选择一个或多个操作桶执行，从而有效地实现多播或广播、高效的聚合、快速故障恢复等。OpenFlow 安全通道是连接 OpenFlow 交换机与控制器的接口，通常由 TLS 加密，实现控制器对交换机的配置和管理。

OpenFlow 交换机通过流水线来处理数据包，如图 3 所示。当数据包进入 OpenFlow 交换机后，随即进入数据包处理流水线，OpenFlow 交换机为每个数据包关联一个操作集（操作集中的操作等到数据包要退出流水线时执行）。数据包从第一个流表开始匹配，把被匹配流表中某条能满足该数据包所有域的最靠前的流表项称为命中表项。一般会把流

表中最后一个表项配置为能匹配所有的数据包，其优先级最低。若没有配置此表项，未与所有流表项匹配成功的数据包将被丢弃。匹配完成后，命中表项中的指令会被执行：修改数据包，把操作加入到操作集，通过把数据包发到后面的流表或组表来改变流水线处理。当命中表项中的指令没有指定下一个流表或组表，流水线处理就结束，此时，与数据包关联的集中操作会被执行。

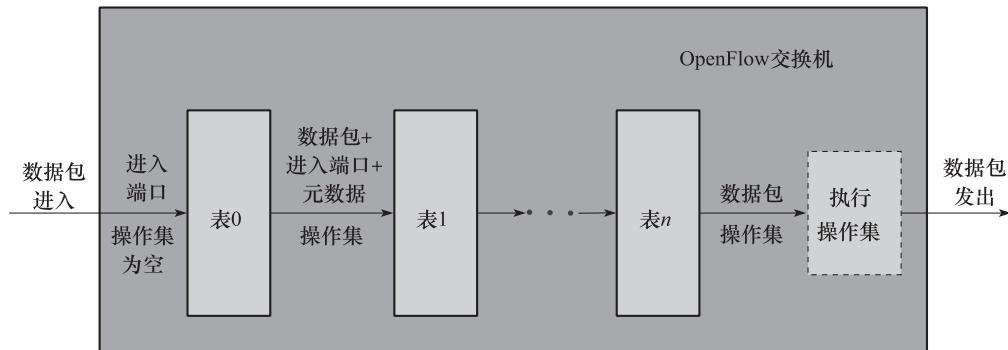


图 3 数据包的流水线处理过程

OpenFlow 协议支持三种消息类型：控制器 – 交换机消息，异步消息，对称消息。

1) 控制器 – 交换机消息：由控制器发起，用来直接管理或检查交换机的状态，添加或删除流表项，交换机根据配置做相应的回复；

2) 异步消息：由交换机发起，交换机用这类消息来向控制器报告网络事件，交换机状态的改变；

3) 对称消息：可以由控制器或交换机发起，用来建立或者保持连接等。

一个控制器可以控制多个 OpenFlow 交换机，一个 OpenFlow 交换机也可以由多个 OpenFlow 控制器控制以解决控制器故障和负载均衡问题。当一个 OpenFlow 交换机由多个控制器控制时，异步消息需要发送给多个控制器；每个控制器有自己的角色，对交换机的读写控制权各不相同。交换机与控制器之间的连接由 OpenFlow 交换机发起，使用 TLS 或 TCP 连接。在维持连接的过程中，控制器和交换机会相互发送 echo 消息来确保连接处于可用状态。当交换机与所有的控制器失去连接后，交换机会立即进入“故障安全模式”或“故障独立模式”。在“故障安全模式”中，发往控制器的数据包和消息会被丢弃；在“故障独立模式”中，交换机相当于传统的以太网交换机和路由器。

3 控制平面：软件定义网络控制器

3.1 控制平面功能

传统的网络控制功能是由分布式设备实现的。SDN 实现了控制平面与数据平面的分

离，将控制平面迁移到一个可控的计算设备之中，使得上层的网络服务和应用程序可以抽象和控制底层的网络设备，并最终通过开放可编程的软件模式实现网络的自动化控制。控制层主要负责集中维护网络拓扑和状态信息，针对用户需求定制数据传输模式并对数据平面的资源编排进行处理等。控制层通过控制平面和数据平面之间的南向接口获取底层网络设施信息，同时为应用层提供可扩展的北向接口。应用层调用控制层的北向接口以实现不同网络功能的应用程序。通过这种调用模式，网络管理者可以动态的配置、管理和优化底层的网络资源，实现灵活、可控的网络功能。

3.2 主要技术挑战

OpenFlow 最初的研究出发点是面向校园网或企业网的创新需求，随着基于 OpenFlow 的 SDN 应用范围逐渐扩大，控制平面面临挑战：

- 1) 流的细粒度处理造成控制器重数据包处理与通信负载^[95]：虽然控制器可以通过主动决策机制提前将控制逻辑部署到数据转发单元以减少数据平面和控制器之间的处理开销，在 OpenFlow 网络中，提前安装流表项也会使大量流表空间无法释放造成浪费资源，但是实际上大部分流的持续时间都很短暂。
- 2) SDN 中控制平面需要维护全局的网络状态信息，网络状态获取的实时性与一致性需求将进一步加重控制器的负载，特别是当网络存在移动结点或拓扑改变时。随着网络规模的不断增大、数据平面转发设备的数量不断增多，单控制器设备已难以满足性能的需求。多控制器存在协同与安全性等问题。
- 3) 网络新型应用需求快速增加，需要将这些新型应用增加到控制平面当中，导致控制器需要对日趋复杂的管控功能进行有效的整合，使得控制平面的处理开销进一步增加。

3.3 主要研究进展

从第一个控制器平台 NOX 出现至今，已逐渐出现一系列基于 OpenFlow 的网络控制器平台。这些控制器平台在向下封装与交换机通信的 OpenFlow 协议的同时，也向上层网络控制应用提供相对更高层的开放编程接口。当前主流的控制器平台主要有 NOX、Onix、Floodlight、Beacon 等，下面对前三种控制器平台进行详细介绍并对比。

(1) NOX

NOX^[18]是作为 OpenFlow 网络中对网络实现可编程控制的中央执行单元而最早引入的。在基于 NOX 的 OpenFlow 网络中，NOX 是控制核心，OpenFlow 交换机是操作实体，其关系如图 4 所示。

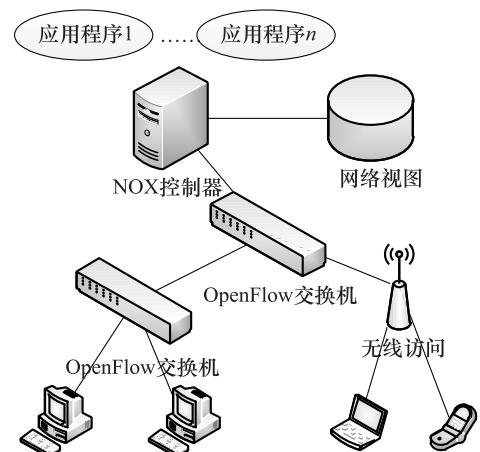


图 4 基于 NOX 的 OpenFlow 网络

示。NOX 需要维护一张网络视图（Network View）来储存整个网络的基本信息，如拓扑、网络单元和提供的服务等。应用程序运行在 NOX 之上，通过调用网络视图中的全局数据，进而操作 OpenFlow 交换机实现对整个网络的管理和控制。NOX 实现了网络基本的管控功能，为 OpenFlow 网络提供了通用 API 的基础控制平台，但未能提供充分的可靠性和灵活性以满足上述控制平面对可扩展的需求。但是，总的来说，NOX 在控制器设计方面实现得最早，目前已成为 OpenFlow 网络控制器平台实现的基础和模板。

NOX 已经发布了多个版本，如 NOX Destiny、NOX Zach 和 POX^[19]等。与 NOX 相比，它们进行了性能优化，并逐渐支持更多的功能，如控制台操作、SNMP 控制等。

(2) Onix

与 NOX 不同，Onix^[20]针对控制平面在可扩展性和通用性等方面不足，提出了一整套面向大规模网络的分布式 SDN 部署方案。Onix 网络架构由网络控制逻辑、Onix、网络连接基础设施和物理网络基础设施四部分组成，其结构如图 5 所示。网络控制逻辑通过 Onix 提供的 API 决策网络的行为；Onix 采用分布式架构向上层提供网络状态的编程接口；网络连接基础设施提供物理网络基础设施和 Onix 之间的通信连接；而物理网络基础设施允许 Onix 读写网络状态。在图 5 中，网络信息库（Network Information Base，NIB）用于维护网络全局的状态信息，Onix 设计的关键就在于实现并维护 NIB 的分发机制，从而保证整个网络状态信息的一致性。

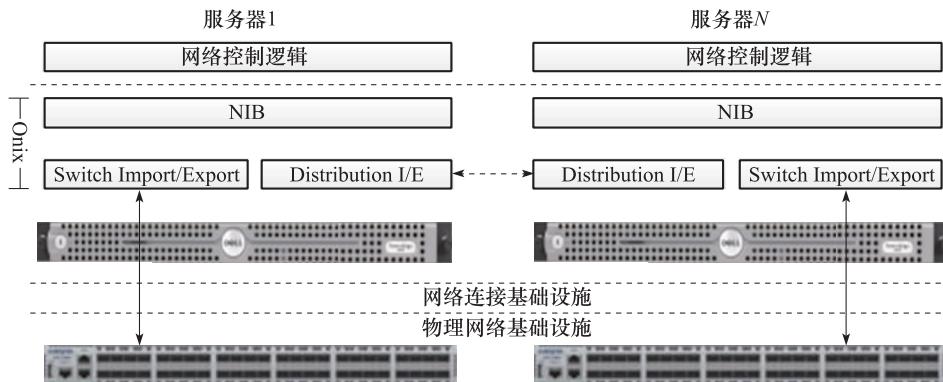


图 5 Onix 网络架构

由于 Onix 目前所采用的 NIB 分发机制实际上都是比较成熟的分布式系统解决方案，因此不可避免地具有这些解决方案自身的某些缺点，如数据库模式自身存在严重的性能限制、多个 Onix 实例对 DHT 更新可能导致状态不一致等。

现在，Onix 已经成为许多组织机构构建商业应用的基础平台。从 Onix 的应用场景来看，每一个 Onix 实例能够管理多达 64 台 OpenFlow 交换机，由 5 个 Onix 实例组成的集群也已经通过了测试。以每台 OpenFlow 交换机能够连接 48 台服务器来计算，Onix 能够应用于具有数万量级主机的较大规模网络。

(3) Floodlight

Floodlight^[21]开放 SDN 控制器是一个企业级的、Apache 许可的、基于 Java 的 Open-

Flow 控制器。它是由一个开源社区支持的，其中包括一些来自 Big Switch Networks 的工程师。Floodlight 是 Big Switch Networks 的商用控制器产品，由专业开发人员的积极测试和改进。与 NOX 相比更贴近于交换机路由器等硬件设备，也更有利于部署网络设备的虚拟化，可以同时和支持 OpenFlow 标准的设备一起工作^[22]。Floodlight 系统示意图如图 6 所示。

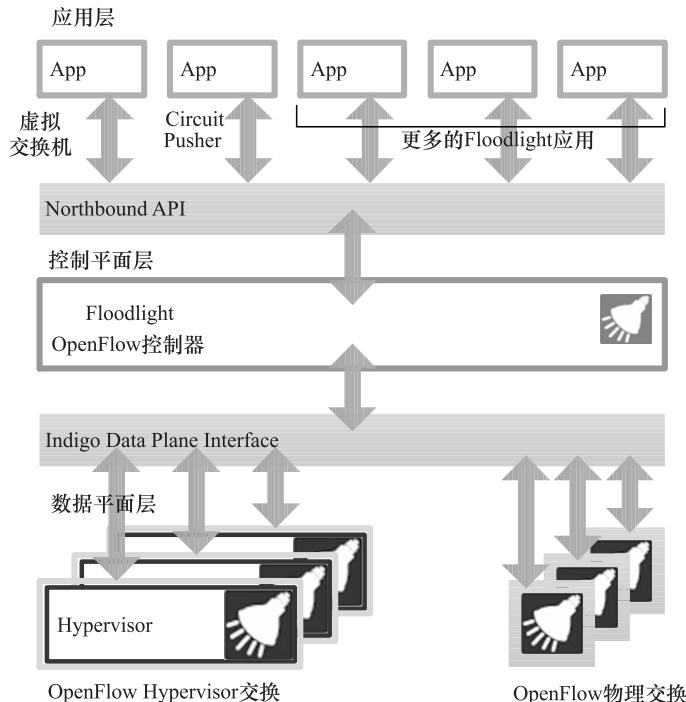


图 6 Floodlight 系统示意图

OpenFlow 通过切换一个远程控制器和明确定义的“转发指令集”来修改网络设备的行为，Floodlight 的设计是用来与支持 OpenFlow 标准的越来越多的交换机、路由器、虚拟交换机和接入点一起工作的。Floodlight 提供了一个模块化加载系统，使得它可以简单地扩展和增强，支持广泛的虚拟和物理的 OpenFlow 交换机，能够处理混合的 OpenFlow 和非 OpenFlow 的网络，支持 OpenStack（链接）的云业务流程平台。

表 1 列举了上述三种 OpenFlow 网络控制器的开发团队、语言及平台，并对比了其主要技术特点。

表 1 OpenFlow 网络控制器软件对比

控制器	开发团队	开发语言	平台	主要技术特点
NOX	Nicira	Python/C++	Linux	最早实现的控制器，当前控制器的模板平台，单线程操作，版本仍在不断更新，性能还在继续完善
Onix	Nicira	Python/C++/Java	Linux	利用分布式系统实现中央控制平面，面向较大规模真实网络的部署方案
Flood-light	Big switch	Java	Win/Mac/Linux	跨平台，基于 Beacon 开发，由开源社区进行维护，遵循 Apache 开源规范，适合推广

4 数据平面：可编程网络转发设备

4.1 数据平面功能

SDN 数据平面主要负责网络数据包查找和转发操作。为了满足 SDN 对数据平面的需求，可编程网络转发设备应该具备以下特点：

- 1) 虚拟化：可编程网络转发设备需要在一台路由器物理设备上同时运行多个虚拟路由器实例，每个虚拟路由器实例可以运行不同的协议，路由器实例之间在功能、性能和故障方面相互隔离，互不干扰，从而支持现有网络业务与未来网络实验和部署的并行运行。
- 2) 可编程：提供从数据平面到控制平面的多层次可编程接口，支持新转发算法和新路由协议，适应创新网络技术的快速实验与部署。
- 3) 高性能：基于有限的物理资源，可编程网络转发设备中的虚拟路由器实例需要达到与商用路由器可比拟的吞吐量和时延等性能指标，使虚拟网络的数据包转发性能接近真实物理网络。

在以上特性中，可编程和虚拟化是松耦合关系：可编程是为了支持自定义的数据包处理，从而支持新协议和新业务的实验与部署；而虚拟化的目的是在此基础之上支持多个不同路由器实例共存，从而支持多个不同用户实验和网络业务的并行隔离运行。可编程和虚拟化会增加系统开销，给路由器性能带来一定的影响（例如，I/O 虚拟化可能导致 I/O 性能下降），因此需要在支持虚拟化和可编程特性的同时，支持高性能数据包处理。

4.2 技术挑战

为了满足上述三个功能需求，可编程网络转发设备在设计与实现上存在着诸多技术挑战，如图 7 所示。

- 1) 可扩展性问题。网络设备虚拟化带来资源可扩展性问题。在可编程网络转发设备中，多个虚拟路由器实例同时运行，每个路由器实例拥有一个转发引擎。随着应用需求的不断增加和网络规模的扩大，虚拟路由器实例个数不断增加，从而导致转发引擎数量不断增加。一方面，转发引擎中的包头协议解析和数据包修改等计算部件，对硬件计算资源（如 FPGA 中的寄存器和逻辑资源）的需求，随着虚拟路由器实例个数的增加而迅速增长；另一方面，转发引擎中的转发表（Forwarding Information Base, FIB）等存储部件，对硬件存储空间（如 TCAM 或 SRAM 芯片）的需求，也随着虚拟路由器实例个数增加而线性增长。如何在硬件计算资源和存储资源有限的情况下，支持尽可能多地转发引

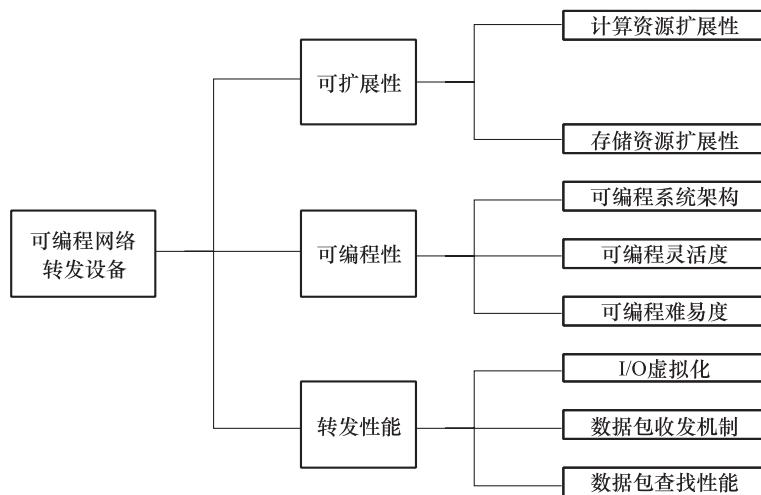


图 7 可编程网络转发设备中的技术挑战

擎，从而支持尽可能多地虚拟路由器实例，成为可编程网络转发设备中面临的可扩展性挑战。

2) 可编程性问题。主要体现在三个方面：可编程系统架构、可编程灵活度和可编程难易度。可编程系统架构要定义数据包处理的基本单元，并提供功能单元之间的统一、友好的编程接口，通过对已有功能单元的选择组合，实现用户自定义的数据包处理流程。如何设计可编程架构，并保证系统实现性能不受影响，是可编程系统架构面临的挑战；此外，在可编程系统架构中，增强单个功能单元（如包头解析单元）的可编程灵活度，可以进一步提高系统的灵活性，满足用户自定义的数据包处理需求，尽量避免用户重新开发新的功能单元。然而，功能单元的可编程灵活度与实现复杂度之间存在矛盾，是功能单元可编程灵活度面临的主要挑战；最后，当需要用户重新开发新功能单元时，如何提供高效的编程方法，以降低用户编程的难度，缩短用户开发时间，构成了可编程难易度的挑战。

3) 转发性能问题。在可编程网络转发设备中，数据包转发性能主要受到 I/O 虚拟化性能，驱动和内核级的数据包收发机制的效率，以及数据包查找算法性能的影响。多个虚拟路由器实例的虚拟网络接口共享物理网络接口，需要 I/O 虚拟化技术的支持。然而，I/O 虚拟化技术由于需要负责虚拟接口和物理接口之间的数据包分发，增加了系统 I/O 开销，降低了数据包转发性能。此外，路由器实例中驱动和内核级的数据包收发机制的效率，也直接影响着数据包采集和发送性能。在传统路由器中，IP 查找和包分类等查找匹配操作，是实现数据包快速转发的瓶颈。在虚拟路由器环境下，上述瓶颈依然存在，而且由于可能承载非 IP 协议，数据包查找匹配问题较传统的 IP 查找和五元组包分类问题更为复杂，面临更为严重的挑战。

为了解决可编程网络转发设备中的上述关键技术挑战，研究人员从可扩展性问题、可编程性问题和转发性能问题等方面展开研究。

4.3 关键技术研究进展

4.3.1 可扩展性

可编程网络转发设备中的可扩展性研究，主要考虑在资源受限的情况下，如何支持尽可能多的虚拟路由器实例。虚拟路由器的控制平面主要由软件实现。由于通用服务器丰富的计算和存储资源，控制平面的可扩展性较好。虚拟路由器的数据平面倾向于硬件实现，从而实现与传统路由器可比拟的性能。高速硬件资源的稀缺性与虚拟路由器实例个数不断增长的应用需求之间存在矛盾。为了提高虚拟路由器数据平面的可扩展性，研究人员提出了数据平面模块共享、转发表合并，以及软硬件虚拟数据平面迁移等方法。

(1) 数据平面模块共享

Anwer 等人^[23]通过分析虚拟路由器数据平面的基本功能模块特性，将数据包处理模块划分为可共享模块和不可共享模块。其中，可共享模块是指无状态模块，不需要维护跟虚拟路由器实例相关的任何状态信息，因此可以在不同虚拟路由器数据平面之间共享，如数据包头部解析、校验和计算、TTL 减一等单元。不可共享模块是指带状态的模块，需要维护跟特定虚拟路由器实例相关的信息，如转发表模块。根据上述分析，Anwer 等人基于可编程硬件 FPGA 设计了虚拟路由器数据平面，通过在各个数据平面之间共享模块，有效提高硬件资源利用率。实验结果表明，在一块 NetFPGA 数据包处理卡上，可以实现 8 个虚拟路由器数据平面。与独立实现 8 个虚拟路由器数据平面相比，节省约 75%~79% 的逻辑资源。上述方法通过共享不同虚拟路由器数据平面之间的计算资源，在一定程度上提高了虚拟路由器数据平面的可扩展性。由于转发表资源并没有在虚拟路由器实例之间共享，总的转发表大小随着虚拟路由器实例个数的增加而线性增长。当虚拟路由器实例个数为 8 时，占用了 FPGA 片内 87% 的存储资源。有限的物理存储资源将进一步导致其支持的虚拟路由器实例个数难以继续增加。

(2) 转发表合并

虽然不同虚拟路由器实例的转发表各不相同，无法直接共享和复用。通过挖掘不同虚拟路由器实例转发表之间的相似性，可以实现多个转发表的合并，从而有效降低多个转发表对存储资源的需求，进一步提高可编程网络转发设备的可扩展性。

Fu 等人^[24]提出了 trie overlap 方法，通过挖掘不同转发表中前缀之间的相似性（其对应的 trie 树也是相似的），在多个 trie 树间共享结点。与多个 trie 树单独隔离存储相比，该方法有效降低了合并后 trie 树的结点个数。合并后的 trie 树结点太大。为了解决结点太大这个问题，trie overlap 方法使用叶推技术。叶推技术会导致增量更新困难，为了避免这一问题，Luo 等人^[25]在合并后的 trie 树结点中引入前缀位图，使结点和下一跳信息分离，每增加一个虚拟路由器实例，trie 树结点大小只增加 1 比特，有效提高了结点大小的可扩展性；在 tire 树合并的过程中，避免了叶推技术的使用，从而在增加可扩展性的同

时，支持快速的增量更新。

虚拟路由器中多个转发表的前缀相似度不一定很高，此时仅仅使用 trie overlap 方法，可能无法取得明显的压缩效果。为了解决不相似 trie 树的合并问题，Song 等人^[26]提出了一种 trie braiding 机制，通过在 trie 树结点中增加一个分支比特，允许 trie 树每个结点的左、右孩子互换位置，从而将不相似的 trie 树调整到相似，以方便使用 trie overlap 方法进行合并压缩。Ganegedara 等人^[27]发现，在 VPN 路由器中，其 FIB 前缀中由于存在不同的公共前缀，即使使用 trie braiding 机制也很难将其对应的 trie 树转换为相似的 trie 树，因此，提出了一种 multiroot 方法，允许从 trie 树的某个中间结点开始合并，从而最大程度上利用子树的相似性。

基于 TCAM 存储 FIB 前缀时，如果每个虚拟路由器实例的转发表单独存储，TCAM 存储空间需求随虚拟路由器实例个数（FIB 个数）的增加而线性增长，有限的 TCAM 存储空间中难以支持大量的虚拟路由器实例的 FIB。为了解决这一问题，Luo 等人^[28]通过挖掘不同 FIB 前缀之间的相似性，在 TCAM 中存储共享前缀，从而降低 TCAM 存储空间需求。

（3）软硬件虚拟数据平面迁移

高效的处理模块共享和转发表合并，可以分别提高硬件计算资源和存储资源的利用率，从而提高硬件虚拟数据平面的可扩展性。高速硬件资源有限是目前存在的客观事实，仅仅依靠提高硬件资源利用率，难以从根本上解决可扩展性问题。要想支持更多的虚拟路由器实例个数，必须借助于可扩展性很好的软件虚拟数据平面。

Unnikrishnan 等人^[29]提出以软件虚拟路由器和硬件虚拟路由器相结合的方式进一步提高可扩展性。少数的、高吞吐的虚拟路由器数据平面在 NetFPGA 中实现，其他大量的、低吞吐的虚拟路由器数据平面在服务器 OpenVZ 中实现。硬件数据平面和软件数据平面支持相同的用户定义的接口，以保证数据平面的位置（在 FPGA 硬件中还是服务器软件中）对用户来讲是透明的。虚拟路由器数据平面可以在软件和硬件中进行动态迁移，以满足动态变化的应用需求。

软硬件虚拟路由器相结合的方式，可以很好地解决虚拟路由器实例的可扩展性问题，但软件路由器往硬件迁移时可能存在较大开销。当一个软件路由器向硬件迁移时，如果硬件中所有的虚拟数据平面均无法实现该软件路由器所需的转发机制，所有的硬件虚拟数据平面都必须迁移到软件中，然后将整个 FPGA 进行重新编程，最后将对应的虚拟路由器实例迁移回硬件。实验结果表明，当硬件中只有一个虚拟数据平面时，上述迁移过程需要超过 10s 的时间。在这段时间里，网络数据包不能在 FPGA 硬件中进行转发。

为了降低迁移时间，Yin 等人^[30]提出将每个硬件虚拟数据平面的转发逻辑和转发表划分为一个局部可重构区域。当硬件数据平面的转发机制需要修改时，不需要对整个 FPGA 进行重新编程，只需利用 Xilinx 公司 FPGA 提供的局部动态可重构特性，对其中某一个局部可重构区域进行动态编程，从而显著缩短了迁移时间，同时也不需要在迁移过程中停止整个 FPGA 硬件数据平面的数据包转发操作。实验结果表明，与整个 FPGA 完全重新编程的迁移方法相比，该方法能将迁移速度提高近 20 倍。

4.3.2 可编程性

软件系统和可编程硬件系统（如 FPGA）本身具备可编程性，可以满足自定义数据包处理的需求。在虚拟路由器中研究可编程性的目的，是为了提出更简单的编程机制，从而降低用户编程的难度；提出更快速的编程机制，从而缩短原型系统设计、实验与部署的时间。为了实现这一目的，研究人员提出了模块化的流水线架构，设计了灵活的可编程模块，实现了易于使用的网络数据包处理编译器。

（1）功能模块的流水线架构

功能模块的流水线架构将路由器中数据包处理的基本单元模块化，通过对这些基本模块进行选择和组合，实现自定义的数据包处理流水线。基于上述模块化的思想，Morris 等人^[31]提出了模块化的路由器数据平面 Click，支持用户自定义的数据包查找和转发处理。然而，Click 数据平面并没有考虑对虚拟化的支持。此外，纯软件实现方式导致其数据包查找和转发的性能较差。为了解决这些问题，Anwer 等人^[32]基于 NetFPGA 硬件实现了模块化的虚拟路由器 SwitchBlade。在数据平面，SwitchBlade 采用模块化的流水线架构，基于 FPGA 实现多个相互隔离的虚拟数据平面，同时兼顾了良好的可编程性和很高的转发性能。

在上述模块化的路由器结构中，可编程性的强弱取决于预先实现的基本模块（基本处理单元）功能是否完备，能否灵活支持模块间的自由选择组合。系统中预先实现的基本模块有限，进一步提高可编程性有两种基本的思路：一种思路是实现灵活的可编程模块，增强单个基本模块的灵活性；另一种思路是降低基本模块的开发难度，以方便用户快速实现自定义模块。

（2）模块的可编程性

在模块化的流水线架构基础之上，增强每个基本模块的灵活度可以进一步提高可编程性。例如，如果包头协议解析模块能够支持自定义偏移、自定义长度的多域关键字抽取，则可以适应 MAC 地址查找、IP 查找、五元组匹配等多种查找操作的包头提取。

对于传统的 TCP/IP 协议，OpenFlow^[33]提出了非常灵活的可编程模块。OpenFlow 将数据包的处理抽象为包头解析、关键字匹配和动作执行三个模块。在包头解析模块，提取数据包的十元组作为查找关键字。在关键字匹配模块，支持基于十元组的通配符匹配和精确匹配。在动作执行模块，定义和实现了路由器中常用的数据包执行动作。上述灵活的可编程模块，可以很好地满足目前 TCP/IP 路由器中常见的各种需求。

OpenFlow 基于预先定义的十元组的匹配方法仍然有其编程的局限性：关键字的提取受到预先定义的十元组的限制；OpenFlow 中元组的定义是基于标准的 TCP/IP 协议，其可编程范围局限在 TCP/IP 体系范围之内。对于非 IP 查找，如 NDN（Named Data Networking）查找等，OpenFlow 预定义的多元组无法满足需求。

CAFE^[34] 和 SwitchBlade^[32] 在其包头解析模块中，设计了任意比特抽取器，其最终目的就是支持数据包头部任意比特域的自由组合，从而支持用户自定义的查找关键字提取方式，支持新的协议类型。这种包头解析模块的灵活性越强，其可编程性越好，对新协

议的支持更好；然而，其复杂程度也越高。复杂的抽取过程给包头解析带来了性能方面的挑战。为了实现高性能的、灵活的包头解析模块，Attig 等人^[35]基于 FPGA 设计了 400Gbps (Gigabit per second, 千兆比特每秒) 包头解析流水线，将每一个头部的解析映射到流水线中的一个流水级，在流水级中通过用户配置的微码，实现灵活的自定义多域抽取。

(3) 数据包处理编译器

如果通过基本模块的自由选择组合，以及可编程模块的灵活配置，仍然无法满足用户的新协议对数据包处理的需求，就需要用户自己编写相应的数据包处理单元，然后以插件的方式，插入到模块化的流水线架构中，或者替换已有模块。

为了简化用户的编程难度，尤其是对硬件模块的编程难度，研究人员设计了网络数据包处理编译器，可以让用户使用自己熟悉的高级语言，如 C/C++ 或脚本语言，编写数据包处理功能，然后，通过编译器将其转换为对应的硬件模块实现。Unnikrishnan 等人^[36]针对基于 FPGA 的模块化虚拟路由器数据平面，设计 Click 风格的数据包处理编译器 ReClick，以降低硬件模块编程的难度。如果用户熟悉 Click 的编程方式，将会很容易使用 ReClick 进行硬件数据包处理。在 Attig 等人^[35]设计的包头解析流水线中，也同时设计了数据包头部解析语言 (Packet Parsing Language, PPL)，以类 C++ 的风格定义了每个头部的格式和处理规则，方便用户对新协议包头解析的自定义处理。

(4) 自定义流表

用户自定义流表可以提高用户可编程的灵活性。不同的 OpenFlow 交换机可能采用异构的硬件芯片，如何在异构芯片上实现用户自定义流表的高效转发，是提高可编程需要解决的重要问题。

实现用户自定义流表可以基于偏移位与长度的 Protocol-Oblivious Forwarding 方法，实现用户自定义字段域^[100]，也可以通过中间层实现流表转换^[92]。

4.3.3 转发性能

在可编程网络转发设备中，I/O 虚拟化带来的开销，数据包收发机制的效率，以及数据包查找匹配操作的性能，成为限制虚拟路由器转发性能的重要因素。为了提高转发性能，工业界和学术界提出大量的虚拟 I/O 加速技术，优化驱动和内核级的数据包收发机制，并加速查找匹配操作的性能。

(1) I/O 虚拟化加速技术

在可编程网络转发设备中，每个虚拟路由器实例都有自己的虚拟网络接口，不同路由器实例之间的虚拟接口逻辑上相互隔离，物理上共享实际的网络接口。为了实现虚拟网络接口的抽象，虚拟机管理程序需要负责完成物理网络接口与各个虚拟网络接口之间的数据包分发，参与各个虚拟路由器实例的每一次 I/O 操作。因此，虚拟机管理程序中的数据包分发工作成为 I/O 虚拟化带来的瓶颈。

Russell 等人^[37]给 Linux 操作系统中的各种虚拟化软件设计了一套标准化的半虚拟 I/O 接口 Virtio。半虚拟化 I/O 能够让客户操作系统意识到自身运行在虚拟化环境中，并在这

个基础上采用一定的优化措施，提升了客户操作系统的 I/O 性能。目前，KVM 虚拟化技术支持 Virtio。

尽管半虚拟化 I/O 技术能够提升客户操作系统的 I/O 性能，但仍然存在上下文开销过高等问题。解决问题的关键在于取消虚拟机管理程序的代理，让客户操作系统能够直接访问到物理设备。具体说来，直接访问物理设备可以划分为两个问题：1) 对客户操作系统的设备缓冲区直接 DMA (Direct Memory Access) 访问。2) 中断重映射。Intel VT-d 技术 (Intel Virtualization Technology for Directed I/O) 解决了以上两个问题。VT-d 技术提供了数据由底层硬件到客户操作系统的直接通路，SR-IOV (Single-Root I/O Virtualization) 则提供了将单个 PCI 设备虚拟成多个设备的方法。SR-IOV 为每个虚拟机提供独立内存空间、中断和 DMA 流，SR-IOV 架构的设计允许一个 I/O 设备支持多个虚拟功能，从而提供了一种不需要软件模拟就可以共享 I/O 设备和 I/O 端口的物理功能的方法。Xen 开发人员开发的 PCI Passthrough 技术支持直接访问 PCI 设备，使得客户操作系统能越过虚拟机管理程序访问设备，从而获得接近原生 Linux 环境下的 I/O 性能。

(2) 数据包收发机制优化

I/O 虚拟化加速技术降低了虚拟化引入的额外 I/O 开销，数据包收发机制的优化，能够进一步提高数据包的转发性能。

RouteBricks^[38] 使用轮询方式代替传统硬件的中断方式，极大地提高了通用硬件对 64B 小包的转发性能；采用批处理的方式，每次总线事务 (Bus Transaction) 都传递多个数据包的地址，充分利用了 PCIE 的总线带宽；利用网卡的多队列特性，将不同队列与不同核进行绑定，极大地提升了数据处理的并行性。使用以上三种手段，可以使得单台服务器的 64B 小包转发性能达到 9.7Gbps。PacketShader^[39] 对 Linux 的内核协议栈进行了优化，将内核中 sk_buff 结构体由原来的 208B 减少为 8B，采用批处理的方式，极大地提升了数据包处理性能。实验表明，仅采用这种优化手段，可以将数据包吞吐量提升到 10.5Gbps，性能提升 13.5 倍。PacketShader 同时也利用了 GPU 多核多队列的特性，并且考虑到其采用的架构为 NUMA (非一致性内存访问)，针对数据包的分发均衡性和提升数据局部性做了精心优化，使其整体转发性能达到 40Gbps。

(3) 查找算法优化

数据包查找操作一直是软件虚拟路由器设计中的一个主要的性能瓶颈。传统的 IP 查找算法也能够应用于可编程网络转发设备中，用于实现快速的 IP 查找。可编程网络转发设备不仅仅需要支持快速的 IP 查找，还需要对非 IP 协议进行支持，如支持 NDN 协议中的基于内容名字的查找等。

NDN 中的内容名字具有类似 HTTP URL 的层次结构。与传统的 IP 地址相比，NDN 数据名字长度不定长，且无上界，导致线速最长名字前缀匹配成为 NDN 查找中一个重要的挑战。为了解决这一问题，Wang 等人^[40] 提出了高效的名字子串编码方法，并应用改进的状态转移数组加速最长名字前缀匹配。Wang 等人^[41] 提出了 MATA 以压缩名字前缀的存储空间，并基于通用服务器和 GPU 实现了线速的名字查找流水线。Yuan 等人^[42] 评估了 NDN 的原型系统实现 CCNx，分析了其转发平面的性能瓶颈，并找到影响 NDN 转

发性能的三个主要问题：支持快速更新的精确字符串匹配问题，基于变长和无上界名字的最长前缀匹配问题，以及大规模流表维护问题。针对这些问题，提出了五大设计原则，用于指导性能可扩展的 NDN 转发引擎设计。

4.4 原型系统实现进展

在上述关键技术研究成果基础之上，国内外研究机构设计实现了多个可编程网络转发设备。

vRouter。 兰开斯特大学的 Laurent Mathy 等人^[43]提出了典型的虚拟路由器架构——vRouter。vRouter 是一种基于通用计算机平台的纯软件虚拟路由器。vRouter 利用普通网卡的硬件多队列特性，提供不同数据平面的数据隔离；采用 Xen 虚拟技术，提供不同的控制平面隔离性；利用内核态 Click 的轮询技术提高数据包转发吞吐量。由于采用纯软件的实现方式，尽管对数据包处理队列做了大量的优化，但 vRouter 性能较差。

OpenFlow 交换机。 斯坦福大学研发的基于 OpenFlow 的新型数据包转发设备^[33]，其采用数据平面与控制平面完全分离的结构。数据平面基于 NetFPGA 实现，采用十元组规则（即 MAC 地址对、IP 对、端口对等）作为数据平面的数据包转发规则，从而实现不同于传统 IP 协议的数据包转发策略。OpenFlow 较好地体现了可编程性，但是 OpenFlow 在数据平面不支持虚拟化，OpenFlow 的多元组查找也面临严重的性能挑战。

SwitchBlade。 佐治亚理工大学的 Anwer 等人^[32]设计了一套基于 FPGA 的可编程网络转发设备框架 SwitchBlade。SwitchBlade 利用硬件中的流水线解耦包处理中的各个模块，方便协议设计人员对特定服务的解析查找模块进行编程，并“插入”到包处理流水线中，加速新协议的实验和部署。SwitchBlade 原型采用 NetFPGA 实现，能够同时支持 4 个不同的硬件虚拟数据平面，硬件虚拟数据平面可以实现线速数据包转发。FPGA 中硬件计算和存储资源非常有限，限制了 SwitchBlade 数据平面的可扩展性。

PEARL。 PEARL^[45]是中科院计算所研制的基于 FPGA 加速板卡和通用服务器的虚拟路由器。采用软件数据平面和硬件数据平面相结合的方式，有效解决了 SwitchBlade 中数据平面可扩展性不足的问题。硬件中支持少数高优先级的虚拟数据平面，采用 FPGA 和 TCAM 加速路由查找和转发等功能；基于通用服务器中丰富的计算和存储资源，实现大量低优先级的虚拟数据平面，显著扩展了支持的虚拟数据平面的数量。基于普通服务器和 PCIe 接口卡的实现方式，限制了 PEARL I 平台的整体吞吐量。为了解决这一问题，中科院计算所正在研制 PEARL II，基于异构芯片与数据包处理功能抽象实现支持虚拟化、可编程和高吞吐量的新一代 SDN 网络转发设备。

ServerSwitch。 微软亚洲研究院^[44]采用 Switch 芯片和普通 Server 相结合，设计了针对数据中心网络应用的高性能可编程路由器 ServerSwitch。通过 Switch 芯片本身提供的可编程性，满足数据中心网络中新应用的可编程需求；通过普通 Server 提供的丰富计算和存储资源，在网络结点中提供复杂的数据包处理能力。ServerSwitch 的可编程性受到商用 Switch 芯片的限制。

表 2 综合比较了上述可编程网络转发设备原型系统研究进展。基于通用服务器平台实现的可编程网络转发设备（如 vRouter）具有很好的可扩展性和可编程性，然而其转发性能较差；基于可编程硬件 FPGA 实现的可编程网络转发设备（如 OpenFlow、Switch-Blade、PEARL 等）具有很好的转发性能。受 FPGA 硬件资源以及编程灵活度的限制，其可扩展性和可编程性有待进一步提高；基于 ASIC 芯片（如 Switch）设计的可编程网络转发设备（如 ServerSwitch）具有比 FPGA 平台更高的转发性能，其可编程灵活性和功能的可扩展性非常有限。总体来讲，单一硬件平台难以同时满足 SDN 中可编程网络转发设备的虚拟化（资源隔离性和可扩展性）、可编程性和高性能三大特性需求，异构的硬件平台逐渐得到广泛关注（如 Server 和 FPGA 相结合、Server 和 Switch 相结合）。如何在异构平台上充分发挥各种硬件平台的优势以弥补对方的不足，设计灵活高性能的可编程网络转发设备，仍然值得进一步研究。

表 2 可编程网络转发设备研究进展

原型系统	实现平台	可扩展性	可编程性	性能
vRouter	Server	基于服务器的丰富资源实现较好的可扩展性	软件灵活可编程	7Mpps (受访存延迟的限制)
OpenFlow	Server, NetFPGA	通过 OpenFlow 控制器扩展复杂的数据包处理功能	标准、开放的可编程接口	OpenFlow 查找转发面临严重的性能挑战
SwitchBlade	Server, NetFPGA	受硬件资源限制，虚拟数据平面数量可扩展性较差	模块化流水线架构，降低 FPGA 编程难度	1.4Mpps (受发包仪发包速率限制)
PEARL	Server, FPGA, TCAM	使用软件数据平面弥补硬件数据平面数量不足的问题	硬件数据平面可编程性有限，软件数据平面灵活可编程	4Gbps (受 4 个 GE 网络接口带宽限制)
ServerSwitch	Server, Switch	使用 CPU 作为协处理器，扩展复杂的数据包处理功能	Switch 芯片可编程性有限，只支持标准协议	4Gbps (受 4 个 GE 网络接口带宽限制)

5 软件定义网络的应用

5.1 基于 SDN 的流量工程

流量工程一直以来是互联网的热点问题。具备控制平面智能化处理特性的 SDN 为流量工程，尤其是数据中心内部和数据中心之间的流量工程提供了良好的平台。无论是数据中心内部流量控制还是数据中心之间的流量调度，SDN 的集中式控制机制和可快速部署新型路由策略的特性都可以提供灵活的解决方案。

第一，SDN 将数据平面与转发平面分离。软件层根据流量工程需求进行定制，而不是根据路由协议定制。第二，集中式的全局控制可以完成更加高效的路由策略，保证数据流传输需求。第三，由于 SDN 周期性收集设备信息并监控链路状况，SDN 可以快速地

完成容错重定向。第四，良好的统一编程接口使得各种特定需求的优化方案或路由策略可以简单快速地完成部署。谷歌公司在 2012 年宣称已经在 G-Scale 中完整部署了基于 OpenFlow 交换机的 SDN 方案，用以解决流量工程中遇到的问题，其中充分体现了 SDN 的良好实施特性^[49]。针对 SDN 在流量工程中的调度算法，谷歌在 2012 年的网络顶级会议 INFOCOM 上发表 3 篇论文^[59~61]。文献[59]针对如何为一个流分配最少的传输路径问题进行研究。在流量工程中，每条传输路径要占据沿途路由器的路由表项，并且软件控制层需要周期性地检测各分配路径的可用带宽情况。因此，最小化传输路径数目既可以减少路由表大小，又能降低软件控制层开销。传统的流分配最优解是 NP 问题，文献[59]通过改进的宽度贪心算法和长度贪心算法在多项式时间内保证了大多数流(1/3)的分配路径数目接近最优(最少)数目。长度贪心算法根据测得的 RTT 作为依据，每次选取当前最短(RTT 最小)可用路径；宽度贪心算法以能当前能容纳流的数目作为选取路径的依据。文献[60]则提出一种改进的 Max-Min Fair 方案，Upward Max-Min Fair。若对于第 N 大的流无法增加分配资源，则可以通过对小于该流的其他流重定向路由使得该流获得更多资源。这种方式对于前 N 个大流仍然保持 Max-Min Fair，但对于其余的小流则不一定保证，因此是 Upward Max-Min Fair。文献[61]分别从用户和网络管理员的角度出发，对流量工程的带宽分配提出一种均衡 Max-Min Fairness(用户期望)和吞吐率(网络管理员期望)的算法。由于谷歌的 SDN 架构拥有对全局资源的控制和应用程序需求函数，文献[61]的均衡式线性规划根据当前网络情况，如瓶颈链路数量和平均链路利用率等，选择不同的均衡参数，从而得出不同的路径分配，提高资源利用率。

凭借 SDN 在 G-scale 上的三年部署经验，谷歌在 2013 年的网络会议 SIGCOMM 上发表了论文 B4^[58]，详细介绍了谷歌在 G-scale 上的集中式流量工程策略。B4 的部署初衷有两方面：链路/设备失效是频繁的事件，交换机硬件应该具备良好的编程接口以便快速部署新型路由策略。相较于传统的流量工程策略，B4 面临着一些 Google 独有的挑战：第一，大部分数据传输是由同步大数据集造成的，这种应用需要高带宽，但不要求低延迟与高容错；第二，谷歌全球部署的数据中心数目有限(目前为止全球共 12 个)，拥有对所有数据中心的交换机、服务器和应用程序的集中控制能力；第三，传统广域网为了保证传输可靠性，链路利用率控制在 30%~40%，这对带宽密集型的 G-Scale 是一种资源浪费。因此，B4 的设计针对上述特点，采用低成本的交换机，用带宽换取容错能力，提高可扩展性；交换机软硬件分离，软件层可快速定制满足 B4 需求的路由策略；集中式流量工程，利用集中式控制，根据应用程序优先级进行多路径转发，保证可用带宽和负载均衡，并快速容错。

利用 SDN 技术可以快速实现集中式流量工程。集中式流量工程通过利用网络拓扑特性、当前网络资源的使用情况以及数据流自身特性，为数据流选择最优的转发路径，既可以提高网络资源利用率，也能节省数据中心的能耗开销。当前云数据中心网络拓扑以多根树为主，如典型的 Fat-Tree^[3]结构等，任意两个服务器之间具备多条路径。传统的基于 spanning tree 的路由协议只能发现单根树形拓扑，无法充分利用多路径，造成链路

资源利用率低。工业界针对多路径特性采用了 ECMP (Equal Cost Multi Path) 作为转发策略。然而，ECMP 没有考虑网络使用情况（如链路带宽）和数据流本身（如吞吐率需求）的特点，使得数据中心的网络资源利用率仍然低下。Hedera^[52]指出，ECMP 存在对大数据流调度不均匀的缺陷，导致多个大数据流哈希到相同路径造成拥塞。因此，Hedera 提出对大数据流（实际占用带宽大于链路带宽的 10%），由控制器根据当前网络状态重新计算转发路径并写入交换机转发表，对于小数据流则仍然采用 ECMP。ElasticTree^[53]则从节能的角度通过控制器监测数据中心交换机状态，选择保持活跃交换机的最小子集，关闭不需要的交换机和链路，从而降低数据中心能耗开销。

SDN 与云数据中心大数据特性的结合为集中式流量工程提供了更加灵活的策略。文献[54]提出将 SDN 中的网络控制器与云数据中心的应用控制器（如 Hadoop 中的 Master）进行结合。由于一个计算任务通常包括多条数据流，因此流量工程将任务完成时间作为参数之一加入调度算法中，更加优化了网络资源使用率。

当前 SDN 的概念仍然没有统一的规范，学术界和工业界的研究者大多以 OpenFlow 作为实现平台来部署 SDN。然而，OpenFlow 的作者 Casado 在文献[55]中指出，SDN 的概念和 OpenFlow 的实现方式还不够清晰和完善，尤其是针对流量工程的需求。当前互联网的骨干网通过 MPLS 来完成流量工程，最突出的特点在于由标签决定传输路径和基于标签的快速转发。复杂的加解标签工作由 LER（边缘路由器）完成，LSR（中间交换路由器）只是完成快速转发。然而，OpenFlow 要对每个数据包包头的多个字段进行解析（几百比特），处理速度远低于 MPLS 的标签处理（几十比特）。另外，若数据流由 IPv4 变为 IPv6，则 OpenFlow 交换机/路由器的匹配规则也要随之更改，而 MPLS 则是将这种协议解析的工作放在边缘路由器来完成。因此，Casado^[55]提出一种类似 MPLS 方式的 SDN 流量工程策略（针对广域网），即边缘交换机/路由器由边缘控制器完成复杂的网络服务，如地址过滤等。核心交换机/路由器由核心控制器完成数据包的快速转发，核心交换机/路由器的连接方式是一种 Fabric，目的在于保证为所有数据流提供最大可用带宽。由于两种控制器分工不同，保证了核心网络（核心控制器）可以专注于流量工程而不需要处理其他复杂的网络服务（类比于 MPLS），提高骨干网络的网络资源利用率。

Lithium^[57]是由佐治亚理工大学 Nick Feamster 教授领导的实验室的 SDN 研究项目。该项目主要以事件驱动的控制器模型来改进当前 OpenFlow 语义匮乏所造成的网络配置问题。当前 OpenFlow 控制器只是针每条流进行操作，缺乏其他网络信息原语，如时间特征、历史信息和用户信息。Lithium 通过提供丰富的网络特征抽象原语，让管理员可以更加快速方便地根据应用程序对传输质量的需求进行（通过控制器）路由选择。

5.2 基于 SDN 的网络虚拟化

数据中心拥有海量的计算资源与存储资源。为了提高这些资源的利用率，实现资源的统计复用，需要使用虚拟化技术。虚拟化技术实现了资源的隔离，多租户可以安全地共同使用数据中心硬件资源。随着技术的进步，计算虚拟化与存储虚拟化已经逐步成熟。

计算虚拟化实现了数据中心计算资源的隔离与复用，用户可以使用虚拟化的计算资源而无需关心物理计算机的管理、维护与升级。存储虚拟化既为用户提供海量的存储空间，又保障用户数据的备份与安全。计算虚拟化与存储虚拟化技术的成熟，按需使用、按需付费的理念正在变成现实，并将成为未来人们使用数据中心的主要方式。

当用户租用多个虚拟机时，虚拟机之间需要进行网络互连和通信，组成了虚拟网络。相对于计算与存储系统虚拟化，网络虚拟化技术的发展相对滞后。由于网络资源的共享特性，为虚拟数据中心网络用户提供安全隔离、智能迁移、带宽保障和灵活调度，是实现数据中心资源复用的必然要求。

SDN 有利于推动数据中心网络虚拟化技术的进步。在 SDN 中，集中式的控制层负责维护全网络的状态信息，网络管理者可以在控制层上编写程序，实现对网络资源的灵活配置与管理，保障网络资源安全。SDN 中的网络控制器能实现全网范围的虚拟机管理与智能迁移调度，并可以与存储资源管理器紧密结合，通过集中式的控制与调度，提高服务器与带宽资源利用率。在 SDN 中，网络管理者可以实现对不同网络设备商提供的网络设备的统一编程，这种方式有利于网络虚拟化机制在全网范围的部署。

为了实现可扩展的虚拟网络隔离和虚拟机的灵活迁移，IETF 先后制定了 NVGRE^[62] 和 VXLAN^[63] 两个网络虚拟化报文格式标准。VXLAN 通过将用户虚拟机发出的以太网数据分组封装在其自定义的 VXLAN 数据分组头部中，再将封装后的数据分组封装在 UDP/IP 数据分组内，实现多租户虚拟网络隔离。NVGRE 则利用标准的 GRE 封装格式对虚拟机发出的以太网数据分组进行封装，实现虚拟网络隔离。在两种报文格式标准中，分组头部都包含了一个 24 比特的用户虚拟网络标识，设计目标都是在三层网络架构上部署虚拟化的二层网络。它们都极大地提高了数据中心可以支持的虚拟网络数量，并实现了三层网络上的虚拟机动态实时迁移。在这两种报文格式标准的部署上，它们都要求服务器上部署一个负责对虚拟机发出的数据分组进行封装的代理，这个代理需要维护对数据分组进行封装所需的信息，在 SDN 中，网络控制器维护着网络中设备的标识以及标识之间的映射，它可以通过与服务器上的代理交互的方式将代理对数据分组进行封装所需的信息发送给代理。

NetLord^[64] 通过将用户虚拟机发出的二层数据分组封装在三层 IP 分组中实现对多用户虚拟网络的支持。通过封装，它允许用户使用任意的二层与三层通信协议，并通过在外层的 IP 分组中包含一个 24 比特的用户虚拟网络标识提高数据中心可以支持的虚拟网络数量。NetLord 利用封装的方式防止用户虚拟机的二层地址信息出现在交换机的转发表中，极大地减小了交换机的转发表空间。NetLord 同样要求在服务器上部署对虚拟机发出的数据分组进行封装的代理，服务器上虚拟机的配置信息会在将该虚拟机分配给该服务器时由 SDN 控制器发送给服务器代理。同时，服务器上的 NetLord 代理可以通过 SDN 控制器与远端的其他 NetLord 代理进行交互，获取对虚拟机发出的数据分组进行封装所需的信息。

SEC2^[65] 通过将用户虚拟机发出的以太网数据分组封装进二层以太网分组中以实现多用户虚拟网络的安全隔离。它将网络分为核心域与边缘域。边缘域由服务器组成，通过

一台或多台交换机与核心域相连，核心域中的交换机负责维护不同边缘域中的虚拟机通信。SEC2 在边缘域中使用 VLAN 实现对用户虚拟网络的隔离，同一个用户在不同的边缘域中租用的虚拟机可能被分配不同的 VLAN 子网标识。在 SEC2 中，SDN 集中式控制器负责维护虚拟机的地址信息和位置信息，当位于不同的边缘域中的虚拟机需要进行通信时，控制器会将对虚拟机发出的以太网分组进行 VLAN 子网标识改写以及对分组进行封装所需的信息发送给连接发起通信的虚拟机所在的边缘域与核心域的交换机，由这些交换机完成将数据分组发送给目的虚拟机的任务。SEC2 突破了 VLAN 子网标识对数据中心可以支持的虚拟网络数量的限制，并通过封装减小了核心域中交换机转发表空间。

在 SDN 中，集中式的控制平面可以实时监控网络中各链路的带宽资源利用情况，并根据用户的需求给用户分配合适的带宽资源。在数据中心中，带宽共享机制主要基于两种思路，一种是基于竞争，另一种是基于分配。由于传统的基于 TCP 流的带宽竞争机制是不公平的，现有的基于竞争的带宽共享方案往往是在虚拟机或用户级别实现带宽竞争，以防止应用程序通过增加流的数目抢占别的应用程序的带宽资源，保证了一定的公平性。在基于竞争的带宽共享机制中，应用了 SDN 架构的包括 Netshare^[68] 和 Faircloud^[69] 等工作。在这两个工作中，SDN 集中式控制器负责根据每个服务的不同需求为其分配相应的权值，最后根据权重对全网的带宽资源以 max-min fairness 的方式进行竞争。这可以在提高带宽利用率的同时，保证需求更大、优先级更高的应用程序可以获得更多的带宽。在 Netshare 中，集中式控制器负责为服务分配相应的权值，而在服务内部不同的虚拟机之间，则使用基于 TCP 流的带宽竞争机制。Faircloud 则是深入研究分析了各种基于权值进行带宽分配的机制的特性。

为了给应用程序提供严格的带宽保障并实现应用程序性能可预测，需要使用基于分配的带宽共享方案。在基于分配的带宽共享方案中，SDN 集中控制层会确切定义每台虚拟机或者每个用户对网络带宽的需求，直接给虚拟机分配足量的带宽，并使用限速机制来确保每台虚拟机或者每个租户使用的带宽不会超过分配的限额。为了对用户的带宽需求进行准确的定义，常用的模型包括流量矩阵模型和“软管”模型。在“软管”模型中，每个用户的虚拟网络拓扑被定义为多个虚拟机连接到一个交换能力无限大的虚拟交换机上，并限制每台虚拟机连接到虚拟交换机的网络带宽。使用“软管”模型的典型分配机制是 Oktopus^[70]，它针对同构的虚拟机带宽请求情况设计了相应的虚拟机分配算法。TIVC^[71] 沿用了 Oktopus 的基本模型，在此基础上，建立了定义应用程序随时间变化的带宽需求的模型，并利用这个模型，设计了相应的虚拟机分配算法，与 Oktopus 相比，提高了带宽的利用率。文献^[72] 则针对虚拟机之间异构的带宽请求情况设计了启发式的虚拟机分配算法，改进了 Oktopus 的分配算法，提高了带宽的利用率。使用流量矩阵模型的典型分配机制是 SecondNet^[73]，它定义了 3 种基本的服务等级，并根据每个应用程序对应的不同服务等级给运行该应用程序的服务器分配虚拟机与网络带宽资源。Bazaar^[74] 设计的资源分配机制通过灵活地调整分配给服务器的虚拟机数量和网络带宽，在保证应用程序服务质量的同时，实现对数据中心虚拟机资源与网络带宽资源的高效利用。

CloudNaaS^[75] 则综合了对虚拟网络隔离与网络带宽共享机制两个问题的研究，基于

SDN 设计了一个虚拟网络架构，实现了用户的应用程序之间的安全隔离，同时保障应用程序的服务质量。在 CloudNaaS 中，用户通过 SDN 控制层提供的接口定义他们对租用的虚拟网络的要求，其中，对服务质量的要求会被转化成虚拟网络中虚拟机之间相互通信的流量矩阵，并根据这个流量矩阵在分配虚拟机时为虚拟机分配足量的带宽。CloudNaaS 通过将属于同一应用程序的虚拟机尽可能地分配到网络拓扑中邻近的位置上，实现网络带宽利用率的提高。通过将虚拟机发出的数据分组中的网络地址修改为服务器的网络地址，CloudNaaS 解决了不同用户之间虚拟机地址空间重叠的问题，并减小了交换机的转发表空间。

5.3 基于 SDN 的网络节能

当前大规模的云数据中心在支持日益丰富的应用和众多租户资源需求的同时，也带来了巨额的能耗成本问题。据统计，2011 年美国数据中心和服务器的总能耗达到了 1000 亿千瓦时^[76]，其中网络设备的能耗占数据中心总能耗的 20% ~ 30%^[77]，并且该比例值还将继续升高。数据中心的巨大能耗问题不但导致了巨额的网络营运成本，而且对社会和环境造成了严重的负面影响，已经成为当今云服务和数据中心可持续发展的重大障碍。

由于当前硬件节能技术发展缓慢且受物理工艺极限的限制，同时云数据中心网络流量呈现较大的随时间波动的特点，因此可以通过软件方式对云数据中心网络进行能耗感知的流量聚合，从而使更多的结点处于空闲状态并休眠，以节省能耗。在 SDN 中，网络智能被抽取到一个集中式的控制器中，数据流的路由、调度等功能都由该控制器来完成，而交换机只是按该控制器所设定的规则进行数据分组的转发。将软件定义网络的思想应用于数据中心当中，能够通过集中式控制提高数据中心网络的可管理性，优化网络的性能和能耗，并能够加快新网络功能的引入。在软件定义的可定制网络框架下，网络控制器可以通过带外（Out-of-Band）获取云数据中心网络的能耗信息、网络流量信息等，并以软件控制的方式进行流量整形（Traffic Shaping）和聚合、网络的节能流量工程、结点休眠和唤醒状态转换等自动能耗控制。因此，当前学术界和工业界开始关注于如何在软件定义的可定制网络中实现网络能耗节省，研究如何在不影响网络性能和可靠性的前提下设计节能的路由和流调度机制。对于 SDN 的技术思路是否可用于广域网，当前学术界和工业界还存在不少争议。但一般认为，SDN 技术更加适用于云数据中心网络环境。

近年来，硬件功耗的降低使得路由器在能耗优化方面已经取得了很大的进展，但研究表明，数据中心流量的增长对网络设备性能和能耗需求的增长远远快于硬件工艺的发展。此外，基于单设备的节能技术依赖于网络流量分布，所带来的节能效果非常有限。最近研究者开始关注于网络级的节能技术，从网络协议设计的角度出发，研究面向节能目标的路由和流调度机制，通过对网络流进行合理、有效的调度，在低网络负载时使更多空闲的网络设备进入休眠状态，从而实现数据中心网络的能耗节省。在新型数据中心网络拓扑（如 Fat-Tree^[78]、BCube^[79]、DCell^[80]）中，每对服务器之间存在多条路由路

径，这使得数据流的路由调度具有较大的灵活性。目前，等价多路径路由策略被广泛应用于数据中心网络的路由选择中，以提高网络吞吐率和改善网络性能，然而等价多路径方案在实现网络流负载均衡的同时，没有考虑到网络能耗因素，造成巨大的能耗浪费。在研究数据中心网络节能路由时，最基本的思路都是在网络低负载时通过关闭或者休眠部分的网络设备来实现能耗的节省。然而节能技术的实施可能会造成网络性能的下降，同时影响网络的可靠性。因此，研究重点在于如何权衡网络节能效果与网络性能之间的关系，以及网络节能效果与网络可靠性之间的关系。另一个研究点是如何权衡节能效果和节能路由计算复杂度之间的关系。高复杂度的路由算法能够获得更好的节能效果，然而不能满足实时性的要求；反之，通过牺牲部分节能效果，能够获得较好的实时性。因此，在研究过程中除了追求节能目标之外，还需要考虑节能路由技术的可行性和实时性。

按照节能路由计算与网络流量负载变化的相关性，可以将当前基于 SDN 技术的数据中心网络节能路由协议分为流量感知的节能路由和拓扑感知的节能路由两类。下面将依次对这两类节能路由的典型方案进行介绍。

流量感知节能路由根据网络流量的实时变化来动态调整路由，使得该路由在网络连通性和网络性能等约束条件下所使用的网络资源最少。通常可以将流量感知节能路由的最优化问题规约成一个多物品流问题（Multi-Commodity Flow, MCF）^[81]，其求解复杂度为 NP 难。一种直接的求解思路是使用最优化问题求解工具来获得该问题的最优解，Chabarek 等人^[82] 使用通用代数建模系统 GAMS 和 CPLEX 工具对该问题进行求解。然而当网络规模较大时该方法很难保证计算的时效性，因此当前常见的方案大多采用近似算法来解决上述问题。

节能的网络资源删除方案基本思想是在网络服务质量约束下尽可能将网络流量聚集起来进行路由传输，使其所占用的网络资源最少。该方案中节能路由的计算过程如下：首先基于给定的网络拓扑和流量矩阵，使用当前缺省的常规路由协议为流量矩阵中每一条流计算初始路由路径；然后在网络连通性和网络性能等约束下从初始路由中按照特定策略迭代地删除网络结点和链路，每一次迭代过程中都需要进行拓扑更新、路由重新计算以及约束条件检查等工作；当网络资源删除后无法继续满足上述约束条件时停止删除操作，最终获得网络流的节能路由路径。该类节能路由方案的研究重点是网络资源的迭代删除策略，当前方案通常根据拓扑和流量信息采用贪婪删除算法。文献[83]提出根据网络结点的连接度和结点流量大小两个参数值递增顺序删除网络结点；文献[84, 85]提出结点与链路配合删除策略，首先按照结点流量与能耗比值递增的顺序删除网络结点，之后按照网络链路利用率递增顺序删除相应的链路；文献[86]针对骨干网路由器之间的逻辑链路包含多条物理传输链路的场景，提出分别按照链路可用带宽的递减顺序和链路删除后对路径拉长影响度的递增顺序删除物理链路的策略。

Shang 等人^[87] 提出一种能量感知的数据中心路由算法。该方案基于数据中心网络拓扑和流量信息计算最优节能路由 R^* ，使得 $N(R^*) = \text{Min } N(R)$ ， $R \in R^+$ ，且 $T(R) \geq K$ 。其中 $N(R)$ 表示在网络路由 R 下所使用的网络设备数量， R^+ 表示所有路由的集合， $T(R)$ 和 K 分别为路由 R 下的网络吞吐率和预定义的网络性能阈值。该方案的目标是使用最少

数量的网络设备提供路由服务，同时满足数据中心网络吞吐量约束。

Zhang 等人^[88]提出一种适用于当前链路状态路由协议和 MPLS 协议的节能流量控制技术。首先使用集中控制器通过 LSA 报文和 TE-LSA 报文收集网络拓扑和链路利用率信息，接着使用最优化问题求解方法在网络流与性能约束下计算节能的传输路径集合，并在链路利用率约束下将网络流均衡分配到该路径集合上，最后使用集中控制器按照其计算结果配置网络结点实现网络节能和负载均衡的联合优化目标。

Vasic 等人^[89]提出一种节能的流量聚集和转移策略：首先计算每个网络结点的流量，在网络链路利用率约束下优先将最小流量结点上的网络流转移到其他传输路径上，并将空闲路径中的网络结点和链路置于休眠模式；接着设定链路速率调整阈值，依次计算每条非空闲链路降低一级速率后与当前负载大小的差值，若差值低于该阈值则将此链路上部分流量转移到其他链路上；最后降低相应链路的配置速率，同时确保在网络链路利用率约束下不升高其他链路的速率。该方案综合利用了设备休眠及链路速率调整等技术实现节能的流量控制。

拓扑感知的节能路由基本思想是使用集中控制器收集网络拓扑信息，分析网络拓扑的自身特性，并在网络拓扑连通性和鲁棒性等约束条件下计算节能的网络路由策略。文献^[91]根据网络拓扑自身特点及与能耗、性能的关系，通过更新网络拓扑和调整路由路径来实现网络能耗节省。由图论可知，若给定网络拓扑 G 的结点互连关系则能够计算出该拓扑的拉普拉斯矩阵 $L(G) = D(G) - A(G)$ 和相应的特征值集合，其中 $D(G)$ 为拓扑 G 中结点连通度对角矩阵， $A(G)$ 为网络结点的邻接矩阵。将 $L(G)$ 矩阵的特征值集合进行递增排序后，其第二个特征值 $\lambda_2(L(G))$ 在一定程度上能够反映网络拓扑的连通性、鲁棒性以及网络结点之间平均距离等特性，因此将其作为网络拓扑特性的评估参数。节能的拓扑连通性调整方案基于上述理论给出连通性调整策略，首先定义两个参数：一是链路 l 对拓扑连通度的影响参数 $I(l) = \lambda_2(L(G)) - \lambda_2(L(G - \{l\}))$ ，反映从拓扑 G 中删除链路 l 后网络连通度的下降程度；另一个是网络连通度降低百分比阈值，表示删除链路后网络连通性与初始拓扑网络连通性的比值下限；接着在该阈值约束下按照参数值 $I(l)$ 递增的次序从拓扑图 G 中删除对应的网络链路并更新网络拓扑；最后根据更新后的拓扑调整网络路由，休眠已删除的网络链路实现网络能耗节省。

在使用拓扑感知的节能路由时，由于节能的路由路径选择不随网络流量负载的变化而变化，因此能够有效避免网络流量频繁变化所引起的路由振荡。

6 国内研究进展

我国研究界和产业界已经在 SDN 的体系结构、网络操作系统、SDN 网络设备、SDN 测量技术和 SDN 应用等方面进行了相关研究，并取得了一些成果。科技部、国家基金委等也设立了相关的研究课题。

浙江大学在 SDN 体系结构和网络操作系统两方面进行了研究。在 2012 年 4 月第二届

全球开放网络峰会（2nd Open Networking Summit）上演示了基于 SDN 架构的可重构网络体系结构 XFlow。该体系结构基于功能构件化的实现思路并提供了动态重构机制。通过标准化的异构网络构件模型采用基于 XML 的技术实现方式，实现对网络的控制和转发功能的抽象封装，同时通过构件的更新、升级、加载、卸载及构件间的组合实现更灵活多样的网络功能，并实现基于工作流的构件协同机制，提供可扩展和可重组的网络构建模型，从而降低了网络功能实现的复杂性，并满足多样化的网络业务应用需求。XFlow 采用管理面、控制面和数据面三层松耦合的网络架构方式，XFlow 动态重构机制允许运行时构件的动态加载、卸载、组合和调整，使得在运行时能够通过对网络业务的提供和形变来满足多样化的应用需求，如 VLAN 和 QoS 保证等。此外，通过标准化的 NetStore 服务和协议提供第三方的开放构件库和重构功能提供服务。分布式 SDN 网络操作系统（DNOS）直接管控底层物理网络，以提高网络操作系统的可伸缩性、可靠性和响应能力。实验室研究目前紧密结合产业链，研制兼容 OpenFlow 协议的分布式 SDN 控制层网络设备。该系统自下而上采用如下四层架构：1) 数据转发层，由 DNOS 实现南向接口来对接标准化 SDN 数据转发层，兼容现有 OpenFlow 协议，并拓展安全可靠、可伸缩、高效的域间数据面控制机制；2) 网络操作系统层管控底层物理网络资源，通过 DNOS 间分布式协作实现水平可伸缩、业务可编程和标准化南北向接口等技术特性，该层向上层提供一致性的网络视图作为北向接口，在视图模块上实现底层网络资源状态的全局可管控；3) 网络业务控制层，为网络具体应用实现诸如路由、安全、接入等网络控制功能，各个功能根据用户需求可重构，并根据运营商等需求可重用，提供标准化的第三方接口；4) 网络业务应用层，为高级网络管理和应用，如 BOSS、计费系统以及增值服务等，提供 OSA/Parlay、Parlay/X、OneAPI 和 IMS 等接口。

中科院计算所研制了构建 SDN 网络的基础路由交换设备。所研制的可编程虚拟路由器 PEARL 具备网络虚拟化功能并提供多种编程方法，能满足未来网络协议创新和业务创新的需要。PEARL 能够在单一物理路由器上虚拟出 128 个逻辑路由器，支持逻辑路由器对网络端口的共享与隔离。不同逻辑路由器能够加载运行不同的协议与业务，支持对不同虚拟网络用户区分服务。PEARL 提供多种编程方法和 API 接口，除具备 IPv4/IPv6 路由转发功能外，还支持 OpenFlow 及 NDN、SOFIA、SERVAL 等未来网络协议架构。相比 OpenFlow 交换机，PEARL 能够在多个虚拟平面上并行运行 OpenFlow，实现更完备的网络虚拟化功能。为实现虚拟化与灵活可编程功能，PEARL I 采用了高性能多核服务器配合专用网络接口卡实现^[45]。其中网络接口卡内建 FPGA 和 TCAM 芯片实现高速复杂数据包处理。设备实测具备 4Gbps 线速处理的能力，能满足接入网络要求。华为公司在 PEARL I 平台基础上与中科院计算所合作开发了 OpenFlow1.2 交换机，采用 OpenFlow 软硬件中间层多级流表转换算法，解决软硬件流表灵活性的差异，最大化利用硬件处理资源提升系统性能^[92]。为解决 FPGA 编程灵活性局限及服务器架构性能局限，中科院计算所在 PEARL I 基础上研制了 PEARL II 系统，它采用中间板交换体系结构，利用众核处理器作为主要处理单元，提供了更好的编程灵活性和更高的性能。PEARL II 提供 10GE 和 GE 网络接口，系统最大吞吐率能力达 320Gbps，可以满足 SDN 核心网络的要求。在可编程

虚拟路由器的实现过程中，数据包查找是最重要的技术问题之一。该问题存在查找速率、更新开销和存储空间可扩展性三大挑战，SRAM 与 TCAM 相结合可能是一种实现方案：文献[93]提出了一种支持快速更新的混合 IP 查找方法，结合 TCAM 和 SRAM 的特点分别优化不相交前缀集和重叠前缀集的更新机制，实验结果表明，该架构能够实现的吞吐量达 250 兆次查找/秒，远远超过 100Gbps 网络链路对 IP 查找的性能需求，且将 TCAM 最坏情况下的更新开销降至 1 次写操作/更新；文献[25]提出了一种基于 SRAM 的可扩展 IP 查找方法，在合并的 trie 树结点中引入前缀位图，使结点和下一跳分离，实验结果表明，存储 14 个 IPv4 核心路由器的 FIB，只需要 10MB 的 SRAM 存储空间，与传统的隔离方法相比降低了 87% 的 SRAM 存储空间需求，更新开销为 1 次 write bubble/更新；文献[28]提出了一种基于 TCAM 的可扩展 IP 查找方法，设计 FIB 填充和 FIB 分割方法解决共享前缀带来的前缀掩盖问题，实验结果表明，与传统的非共享方法相比，使用上述两种方法存储 14 个 IPv4 核心路由器的 FIB 时，能够分别降低 92% 和 82% 的 TCAM 存储空间需求。

清华大学在 SDN 的系统架构以及其在数据中心网络中的应用方面展开了深入研究，主要研究成果包括：1) 以数据为中心的软件定义网络架构 SODA (Software Defined Data Centric Networking)。与 OpenFlow 相比，SODA 大大增强了数据层面的处理能力，可以对数据流进行更加灵活的处理和转发，为 SDN 控制平面的功能定制提供了更开放的接口。2) 基于 SDN 的数据中心网络虚拟化机制。设计了轻负载的数据中心网络隔离机制以及虚拟网络带宽保障算法^[97]。3) 基于 SDN 的数据中心网络组播协议。通过集中式的组播组分配和管理，实现可扩展数据中心网络组播，通过低端交换机支持大规模的组播组会话^[98,99]。4) 基于 SDN 的数据中心网络内容转发协议。通过灵活定义分组字段和转发规则，实现在 IP 转发基础上、利用内容名字进行转发加速。相关研究成果发表在 ToN、INFOCOM、ICNP 等会议和期刊上。此外，在 SDN 技术应用方面，清华大学部署在该校信息楼内的网络安全系统 Livesec^[94]，该系统在传统以太网之上，通过无线接入技术和虚拟化技术引入了基于 OpenFlow 协议的控制层，显著降低了构建成本。清华大学还实现了面向 SDN 的网络操作系统 TUNOS；在 SDN 网络操作系统之上构建了一个虚拟云平台 VCP；对 OpenFlow 进行扩展，实现了支持 OpenFlow 的商用路由器 OpenRouter；NOX 上开发的应用程序来获取全局网络视图，以解决因特网源地址验证标准（SAVI）的不足。

解放军理工大学全面综述了 OpenFlow 技术研究的发展现状^[95]。针对 SDN 测量问题进行了研究，研制了一种基于 OpenFlow 的未来互联网测量平台 OpenTrace。该平台目标是为定量分析各种创新应用或机制提供测量分析手段。OpenTrace 采用了分布式测量与集中式分析结合的体系结构。分布式测量功能是指每个测量实体的本地测量功能：所有设备内置被动网络测量的功能，记录下经过本实体的所有控制平面和数据平面信息；以最小侵扰的方式进行测量，记录两个平面运行轨迹的日志仅存放在本地，直至试验结束为止。其中集中式分析功能是指汇集分散在各个实体上的本地日志，以形成完整的网络运行视图。

针对目前 OpenFlow 交换机多表结构对硬件资源的开销，国防科技大学提出了三层的

OpenFlow 交换机实现模型：硬件层、内核层和用户层。硬件层直接采用 TCAM 实现多级流表的报文转发功能，内核层实现流表的精确匹配功能并实现相关的规则管理，用户层则负责规则逻辑的保存。目前尚未看到实际的硬件实现平台，但这种分层的实现方法符合 OpenFlow 交换平台转发规则的软硬件实现规范。

在广域网中部署 SDN 需要考虑多控制器的部署位置问题，如果考虑多个优化指标则将导致 NP 问题。北京邮电大学的 Hu Yan-nan 等人针对特定的指标提出了几个多控制的部署算法，仿真结果表明新的部署算法能够响应的提高这些指标的性能。

华为、中兴通信等也设立专门的研发团队研发 SDN 交换机与控制器系统，盛科网络研制了 OpenFlow 芯片。

7 总结与展望

自从 2008 年 OpenFlow 诞生以来，软件定义网络（SDN）的技术理念已经引起了计算机网络学术界和工业界的高度重视。各国政府资助了大量的 SDN 项目，产业界成立了 ONF 联盟开展 SDN 的标准化和研发工作，国际学术会议如 SIGCOMM、NSDI 等也有数量众多的 SDN 领域的论文发表。

SDN 的核心要义是“开放”，即要把网络功能以可灵活编程的方式开放给上层应用，实现网络功能的灵活定制，打破网络设备“封闭”的现状。因此，SDN 将对网络架构和协议的实现产生重大的革命性影响。当前，SDN 最具代表性的协议架构是 OpenFlow。围绕 OpenFlow 框架，学术界和工业界对 SDN 南向接口的协议、北向接口的高级编程语言和网络操作系统都进行了大量的研究工作。但由于 OpenFlow 主要提供了控制器与网络结点之间的通信协议并且其数据层面的抽象方式过于简单，业界仍然在对 SDN 体系结构进行深入探讨：

1) 控制平面的扩展性与协同。SDN 体系结构中逻辑集中的控制管理需要控制平面维护全局网络信息。随着网络规模不断扩大，网络结点和链路动态性的增加，集中式控制平面处理负担急剧增加，面临可扩展性挑战。因此，在保持控制平面“逻辑集中”的基础之上，需要实现分布式多控制器协同处理，从而降低单一控制器的处理负担，提高控制平面可扩展性。

2) 更灵活可编程的数据平面抽象机制和实现方式。SDN 对网络的可控性归根到底取决于数据平面提供的编程灵活性。目前以 OpenFlow 为代表的 SDN 数据平面抽象机制仅以“Match-Action”方式对数据包进行转发，无法支持需求更灵活的数据平面操作，比如基于逻辑运算的包转发机制、有状态的分组计算等。如果进行更加合理、更加灵活的数据平面抽象，是 SDN 值得研究的重要问题。基于 SDN 的网络编程语言。SDN 为网络管理员和用户提供了易用的可编程接口，但需要通过相应的网络编程语言来实现。目前并没有专门针对开放网络功能的可编程高级语言，因此这也将是 SDN 发展的重要方面。

3) 高性能转发平面技术与实现。对于转发平面而言，不同高速硬件平台具有不同优势，如FPGA适合于简单、高性能的转发计算，TCAM支持高速查找，而多核处理器具有灵活性强，支持复杂转发处理等特点。在SDN网络转发设备中，为了更好地满足不同网络应用的性能需求，硬件平台趋于异构化。如何利用异构芯片的性能实现高效灵活的数据平面转发，既需要支持高效查找、快速更新、低存储开销的算法，也需要基于流表特征提升实现技术。

4) 基于SDN的网络虚拟化。虚拟网络需要满足不同租户之间的网络隔离、网络带宽保障等需求。SDN为网络虚拟化提供了新的使能技术，将大大促进网络虚拟化技术的发展。如何在大规模网络环境下保证虚拟网络之间流量不泄露、地址空间重叠的情况下实现正确路由、海量虚拟机通信的网络中实现可扩展转发以及各个租户网络之间保障可用带宽，都是值得研究的问题。

5) SDN创新应用。SDN在数据中心网络中的应用已经有不少成功案例，而对于SDN在广域网环境下的应用前景到底如何，还需假以时日。我国已经有不少高校和科研院所开展SDN相关研究，这也是我国信息网络产业在国际新一轮技术革命浪潮中抢占技术制高点的大好机会。

参考文献

- [1] Barry M Leiner, Vinton G Cerf, David D Clark, Robert E Kahn, Leonard Kleinrock, Daniel C Lynch, Jon Postel, Larry G Roberts, Stephen Wolff. Brief History of the Internet, <http://www.internetsociety.org/internet/what-internet/history-internet/brief-history-internet>.
- [2] Yang L, Dantu R, Anderson T, Gopal R. Forwarding and Control Element Separation (ForCES) Framework. RFC 3746, 2004.
- [3] Wang B, Xia Y, Chen X, et al. Research and Implementation of ForCES-based IPv6 Router[J]. JOURNAL-NATIONAL UNIVERSITY OF DEFENSE TECHNOLOGY, 2006, 28(3): 44.
- [4] Greenberg A, Hjalmtysson G, Maltz DA, Myers A, Rexford J, Xie G, Yan H, Zhan J, Zhang H. A Clean Slate 4D Approach to Network Control and Management. ACM SIGCOMM Computer Communication Review, 2005, 35(5): 41-54.
- [5] Caesar M, Caldwell D, Feamster N, Rexford J, Shaikh A, Merwe J. Design and Implementation of a Routing Control Platform. Proceedings of the 2th USENIX Symp. on Networked Systems Design and Implementation(NSDI). Boston: USENIX Association, 2005. 15-28.
- [6] Casado M, Garfinkel T, Akella A, Freedman MJ, Boneh D, McKeown N, Shenker S. SANE: A Protection Architecture for Enterprise Networks. Proceedings of the 15th Conf. on USENIX Security Symp. Vancouver: USENIX Association, 2006. 137-151.
- [7] Casado M, Freedman MJ, Pettit J, Luo J, McKeown N, Shenker S. Ethane: Taking Control of the Enterprise. Proceedings of the SIGCOMM 2007. Kyoto: ACM Press, 2007. 1-12.
- [8] McKeown N, Anderson T, Balakrishnan H, et al. OpenFlow: Enabling Innovation in Campus Networks. ACM SIGCOMM Computer Communication Review, 2008, 38(2): 69-74.

- [9] <http://www.im2013.org/program/keynotes>.
- [10] MIT technology review. 2009. <http://www2.technologyreview.com/article/412194/tr10-software-defined-networking/>.
- [11] Elliott C. GENI: Opening up new classes of experiments in global networking. *IEEE Internet Computing*, 2010, 14(1): 39-42.
- [12] Internet 2. 2012. <http://www.internet2.edu>.
- [13] OFELIA. 2012. <http://www.fp7-ofelia.eu>.
- [14] JGN2 plus. 2012. <http://www.jgn.nict.go.jp/english/index.html>.
- [15] Kim N, Kim J. NetOpen Networking Service: Software-defined Networking Service on Programmable Network Substrates. *Proceedings of ASIA FI*, 2011.
- [16] CISCO. 2012. http://www.cisco.com/web/CN/aboutcisco/news_info/corporate_news/2012/06_14.html.
- [17] GOOGLE. 2012. <http://research.google.com/pubs/Networking.html>.
- [18] Gude N, Koponen T, Pettit J, Pfaff B, Casado M, McKeown N, Shenker S. Nox: Towards an Operating System for Networks. *ACM SIGCOMM Computer Communication Review*, 2008, 38(3): 105-110.
- [19] NOX. 2012. <http://noxrepo.org>.
- [20] Koponen T, Casado M, Gude N, Stribling J, Poutievski L, Zhu M, Ramanathan R, Iwata Y, Inoue H, Hama T, Shenker S. Onix: A Distributed Control Platform for Large-Scale Production Networks. *Proceedings of the 9th USENIX Conf. on Operating Systems Design and Implementation (OSDI)*. Vancouver: USENIX Association, 2010.
- [21] Floodlight. 2012. <http://floodlight.openflowhub.org>.
- [22] Floodlight is an Open SDN Controller [EB/OL]. [2012-10-19]. <http://www.projectfloodlight.org/floodlight/>.
- [23] Anwer M B and Feamster N. Building a Fast, Virtualized Data Plane with Programmable Hardware. *Proceedings of the 1st ACM SIGCOMM Workshop on Virtualized Infrastructure Systems and Architectures (VI- SA)*. Barcelona, Spain, 2009: 1-8.
- [24] Fu J and Rexford J. Efficient IP-address Look Up with a Shared Forwarding Table for Multiple Virtual Routers. *Proceedings of the 4th ACM International Conference on Emerging Networking Experiments and Technologies (CoNEXT)*. New York, NY, USA, 2008: 1-12.
- [25] Luo L, Xie G, Salamatian K, Uhlig S, Mathy L, and Xie Y. A Trie Merging Approach with Incremental Updates for Virtual Routers. *Proceedings of the IEEE INFOCOM*. 2013.
- [26] Song H, Kodialam M, Hao F, and Lakshman T V. Building Scalable Virtual Routers with Trie Braiding. *Proceedings of the IEEE INFOCOM*. 2010: 1442-1450.
- [27] Ganegedara T, Jiang W, and Prasanna V. Multiroot: Towards Memory-Efficient Router Virtualization. *Proceedings of the International Conference on Communications (ICC)*. 2011.
- [28] Luo L, Xie G, Uhlig S, Mathy L, Salamatian K, Xie Y. Towards TCAM-based Scalable Virtual Routers. *Proceedings of the 8th International Conference on Emerging Networking Experiments and Technologies (CoNEXT)*. New York, NY, USA, 2012: 73-84.
- [29] Unnikrishnan D, Vadlamani R, Liao Y, Dwaraki A, Crenne J, Gao L, and Tessier R. Scalable Network Virtualization Using FPGAs. *Proceedings of the 18th Annual ACM/SIGDA International Symposium on Field Programmable Gate Arrays (FPGA)*. Monterey, California, USA, 2010: 219-228.
- [30] Yin D, Unnikrishnan D, Liao Y, Gao L, Tessier R. Customizing Virtual Networks with Partial FPGA

- Reconfiguration. Proceedings of the Second ACM SIGCOMM Workshop on Virtualized Infrastructure Systems and Architectures(VISA). New Delhi , India , 2010: 57-64.
- [31] Morris R, Kohler E, Jannotti J, and Kaashoek M F. The Click modular router. Proceedings of the Seventeenth ACM Symposium on Operating Systems Principles (SOSP) . Charleston, South Carolina, United States, 1999: 217-231.
- [32] Anwer M B, Motiwala M, Tariq M B, Feamster N. SwitchBlade: a Platform for Rapid Deployment of Network Protocols on Programmable Hardware. ACM SIGCOMM Computer Communication Review, 2010, 140 (4) : 183-194.
- [33] McKeown N, Anderson T, Balakrishnan H, Parulkar G, Peterson L, and Turner J. OpenFlow: Enabling Innovation in Campus Networks. ACM SIGCOMM Computer Communication Review, 2008, 38 (2) : 69-74.
- [34] Lu G, Shi Y, Guo C, and Zhang Y. CAFE: a Configurable Packet forwarding Engine for Data Center Networks. Proceedings of the 2nd ACM SIGCOMM Workshop on Programmable Routers for Extensible Services of Tomorrow(PRESTO). Barcelona, Spain, 2009: 25-30.
- [35] Attig M and Brebner G. 400 Gb/s Programmable Packet Parsing on a Single FPGA. Proceedings of the Seventh ACM/IEEE Symposium on Architectures for Networking and Communications Systems(ANCS) . 2011: 12-23.
- [36] Unnikrishnan D, Lu J, Gao L, Tessier R. ReClick-A Modular Dataplane Design Framework for FPGA-Based Network Virtualization. Proceedings of the Seventh ACM/IEEE Symposium on Architectures for Networking and Communications Systems(ANCS) . 2011: 145-155.
- [37] Russel R. Virtio: towards a De-facto Standard for Virtual I/O Devices. SIGOPS Operating Systems Review, 2008, 42(5) : 95-103.
- [38] Dobrescu M, Egi N, Argyraki K, Chun B G, Fall K, Iannaccone G, Knies A, Manesh M, and Ratnamsamy S. RouteBricks: Exploiting Parallelism to Scale Software Routers. Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles. Big Sky, Montana, USA, 2009: 15-28.
- [39] Han S, Jang K, Park K, and Moon S. PacketShader: a GPU- Accelerated Software Router. ACM SIGCOMM Computer Communication Review, 2010, 40(4) : 195-206.
- [40] Wang Y, He K, Dai H, Meng W, Jiang J, Liu B, and Chen Y. Scalable Name Lookup in NDN Using Effective Name Component Encoding. Proceedings of the 32nd International Conference on Distributed Computing Systems(ICDCS). Washington, DC, USA, 2012: 688-697.
- [41] Wang Y, Zu Y, Zhang T, et al, Wire Speed Name Lookup: A GPU-based Approach. Proceedings of the 10th USENIX Symposium on Networked Systems Design and Implementation(NSDI). April 2-5 , 2013.
- [42] Yuan H, Song T, Crowley P. Scalable NDN Forwarding: Concepts, Issues and Principles, in Proceedings of the 21st International Conference on Computer Communications and Networks(ICCCN) , 2012.
- [43] Egi N, Greenhalgh A, Handley M, Hoerdt M, Huici F, Mathy L. Towards High Performance Virtual Routers on Commodity Hardware. Proceedings of ACM CoNEXT, 2008.
- [44] Lu G, Guo C, Li Y, Zhou Z, Yuan T, Wu H, Xiong Y, Gao R, and Zhang Y. Serverswitch: A Programmable and High Performance Platform for Data Center Networks. Proceedings of the 8th USENIX Conference on Networked Systems Design and Implementation(NSDI) , 2011.
- [45] G Xie, P He, H Guan, Z Li, Y Xie, L Luo, J Zhang, Y Wang, K Salamatian. PEARL: A Programmable Virtual Router Platform, IEEE Communications Magazine, 2011.

- [46] NetMagic, <http://www.netmagic.org/>.
- [47] T Feng, J Bi, H Hu. OpenRouter: OpenFlow Extension and Implementation Based on a Commercial Router. Proceedings of the IEEE ICNP, 2011.
- [48] Y Li, L Su, D Jin, L Zeng. TUNIE: A Virtualized Platform for Network Experiment on Programmable Infrastructure. Proceedings of the IEEE ICNP, 2011.
- [49] <https://www.opennetworking.org/images/stories/downloads/sdn-resources/customer-case-studies/csgooglesdn.pdf>.
- [50] <http://www.ietf.org/proceedings/84/slides/slides-84-sdnrg-4.pdf>.
- [51] Mohammad Al-Fares, Alexander Loukissas, Amin Vahdat. A Scalable, Commodity Data Center Network Architecture. Proceedings of the ACM SIGCOMM, 2008.
- [52] M Al-Fares, S Radhakrishnan, B Raghavan, N Huang, A Vahdat. Hedera: Dynamic Flow Scheduling for Data Center Networks. Proceedings of the NSDI, Apr. 2010.
- [53] B Heller, S Seetharaman, P Mahadevan, Y Yiakoumis, P Sharma, S Banerjee, and N McKeown. ElasticTree: Saving Energy in Data Center Networks. Proceedings of the NSDI, 2010.
- [54] Wang G, Ng T S, Shaikh A. Programming Your Network at Run-time for Big Data Applications[C]. Proceedings of the First Workshop on Hot Topics in Software Defined Networks. ACM, 2012: 103-108.
- [55] Casado M, Koponen T, Shenker S, et al. Fabric: a Retrospective on Evolving SDN[C]. Proceedings of the First Workshop on Hot Topics in Software Defined Networks. ACM, 2012: 85-90.
- [56] <http://openflowswitch.org/wk/index.php/PAC.C>.
- [57] <http://projectbismark.net/lithium/>.
- [58] Sushant Jain, Alok Kumar, Subhasree Mandal, et al. B4: Experience with a Globally-Deployed Software Defined WAN. Proceedings of the ACM SIGCOMM, 2013.
- [59] Hartman T, Hassidim A, Kaplan H, et al. How to Split a Flow. Proceedings of the IEEE INFOCOM, 2012: 828-836.
- [60] Danna E, Hassidim A, Kaplan H, et al. Upward Max Min Fairness. Proceedings of the IEEE INFOCOM, 2012: 837-845.
- [61] Danna E, Mandal S, Singh A. A Practical Algorithm for Balancing the Max-min Fairness and Throughput Objectives in Rraffic Engineering. Proceedings of the IEEE INFOCOM, 2012: 846-854.
- [62] M Sridharan, K Duda, I Ganga, etc. NVGRE: Network Virtualization Using Generic Routing Encapsulation. IETF draft, 2011.
- [63] M Mahalingam, D Dutt, K Duda, etc. VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks. IETF draft, 2012.
- [64] J Mudigonda, P Yalagandula, B Stiekes, etc. NetLord: A Scalable Multi-tenant Network Architecture for Virtualized Datacenters. Proceedings of the ACM SIGCOMM 2011, Toronto, Ontario, Canada.
- [65] F Hao, T Lakshman, S Mukherjee, etc. Secure Cloud Computing with a Virtualized Network Infrastructure. Proceedings of the USENIXHotCloud, June, 2010.
- [66] A Greenberg, J Hamilton, N Jain, etc. VL2: A Scalable and Flexible Data Center Network. Proceedings of the SIGCOMM, 2009.
- [67] R Mysore, A Pamboris, N Farrington, etc. Port Land: A Scalable Fault-Tolerant Layer 2 Data Center Network Fabric. Proceedings of the ACM SIGCOMM, 2009.
- [68] T Lam, S Radhakrishnan, A Vahdat. NetShare: Virtualizing Data Center Networks across Services, Tech-

- nical Report, UCSD, 2010.
- [69] L Popa, A Krishnamurthy, S Ratnasamy, etc. Fair Cloud: Sharing the Network in Cloud Computing. Proceedings of the HotNets, 2011.
- [70] H Ballani, P Costa, T Karagiannis, etc. Towards Predictable Datacenter Networks. Proceedings of the SIGCOMM, 2011.
- [71] D Xie, N Ding, Y C Hu, etc. The Only Constant is Change: Incorporating Time-varying Network Reservations in Data Centers. Proceedings of the SIGCOMM, 2012.
- [72] J Zhu, D Li, J Wu, etc. Towards Bandwidth Guarantee in Multi-tenancy Cloud Computing Networks. Proceedings of the ICNP, 2012.
- [73] C Guo, G Lu, H J Wang, etc. Second Net: a Data Center Network Virtualization Architecture with Bandwidth Guarantees. Proceedings of the CoNext, 2010.
- [74] V Jalaparti, H Ballani, P Costa, etc. Bridging the Tenant-provider Gap in Cloud services. Proceedings of the SoCC, 2012.
- [75] T Benson, A Akella, A Shaikh, etc. Cloud NaaS: A CloudNetworking Platform for Enterprise Applications. Proceedings of the ACM SOCC, June 2011.
- [76] U. S. Environmental Protection Agency. Data Center Report to Congress. <http://www.energystar.gov>.
- [77] B Heller, S Seetharaman, P Mahadevan. ElasticTree: Saving Energy in Data Center Networks. Proceedings of the NSDI, Apr 2010.
- [78] M Al-Fares, A Loukissas, and A Vahdat. A Scalable, Commodity Data Center Network Architecture. Proceedings of the SIGCOMM, 2008.
- [79] C Guo, H Wu, K Tan, L Shiy, Y Zhang, S Lu. BCube: A High Performance, Server-centric Network Architecture for Modular Data Centers. Proceedings of the SIGCOMM, 2009.
- [80] C Guo, H Wu, K Tan, L Shiy, Y Zhang, S Lu. DCCell: A Scalable and Fault-tolerant Network Structure for Data Centers. Proceedings of the SIGCOMM, 2008.
- [81] Even S, Itai A, Shamir A. On the Complexity of Time Table and Multi-commodity Flow Problems. Proceedings of the 16th Annual Symposium on Foundations of Computer Science (SFCS '75). Berkeley: IEEE, 1975: 184-193.
- [82] Chabarek J, Sommers J, Barford P, Estan C, Tsiang D, Wright S. Power Awareness in Network Design and Routing. Proceedings of the IEEE INFOCOM 2008. Phoenix: IEEE, 2008: 457-465.
- [83] Chiaravaggio L, Mellia M, Neri F. Reducing Power Consumption in Backbone Networks. Proceedings of the IEEE International Conference on Communications. (ICC '09). Dresden: IEEE, 2009: 1-6.
- [84] Chiaravaggio L, Mellia M, Neri F. Energy-aware Backbone Networks: a Case Study. Proceedings of the IEEE International Conference on Communications Workshops (ICC Workshops '09). Dresden: IEEE, 2009: 1-5.
- [85] Gelenbe E, Silvestri S. Optimisation of Power Consumption in Wired Packet Networks. Quality of Service in Heterogeneous Networks, 2009, 22(12): 717-729. [doi: 10.1007/978-3-642-10625-5_45].
- [86] Fisher W, Suchara M, Rexford J. Greening Backbone Networks: Reducing Energy Consumption by Shutting Off Cables in Bundled Links. Proceedings of the 1st ACM SIGCOMM Workshop on Green Networking. New Delhi: ACM, 2010: 29-34.
- [87] Y Shang, D Li, M Xu. Energy-aware Routing in Data Center Network. Proceedings of the 1st ACM SIGCOMM Workshop on Green Networking. New Delhi: ACM, 2010: 1-8.

- [88] Zhang MG, Yi C, Liu B, Zhang B C, Green T E. Power-Aware Traffic Engineering. Proceedings of the 18th IEEE International Conference on Network Protocols (ICNP'10). Kyoto: IEEE, 2010: 21-30.
- [89] Vasic N, Kostic D. Energy-Aware Traffic Engineering. Proceedings of the 1st International Conference on Energy-Efficient Computing and Networking (e-Energy'10). Passau: ACM, 2010: 169-178.
- [90] Heller B, Seetharaman S, Mahadevan P. Elastic Tree: Saving Energy in Data Center Networks. Proceedings of the 7th USENIX Symposium on Networked Systems Design and Implementation (NSDI'10). San Jose: USENIX Association, 2010: 249-264.
- [91] Cuomo F, Abbagnale A, Cianfrani A, Polverini M. Keeping the Connectivity and Saving the Energy in the Internet. Proceedings of the IEEE INFOCOM 2011 Workshops. Shanghai: IEEE, 2011: 319-324.
- [92] Heng Pan, Hongtao Guan, Junjie Liu, Wanfu Ding, Chengyong Lin, Gaogang Xie. The Flow Adapter: Enable Flexible Multi-Table Processing on Legacy Hardware. Proceedings of the ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN), 2013.
- [93] Layong Luo, Gaogang Xie, Yingke Xie, Laurent Mathy, Kavé Salamatian. A Hybrid IP Lookup Architecture with Fast Updates. Proceedings of the INFOCOM, Orlando, Florida, USA, Mar. 25-30, 2012.
- [94] Qi YX, He F, Wang K, Chen X M, Fong J, Xie F, Shao Y Y, Gao Y, Xue Y B, Li J. Live Sec: Open Flow-based Security Management for Production Networks. Proceedings of the 30th IEEE International Conference on Computer Communications (INFOCOM) Demo, 2011.
- [95] 左青云, 陈鸣, 赵广松, 邢长友, 张国敏, 蒋培成. 基于 OpenFlow 的 SDN 技术研究 [J]. 软件学报, 2013, 24(5): 1078-1097.
- [96] Qi Y X, He F, Wang K, Chen X M, Fong J, Xie F, Shao Y Y, Gao Y, Xue Y B, Li J. Live Sec: OpenFlow-based Security Management for Production Networks. Proceedings of the 30th IEEE International Conference on Computer Communications (INFOCOM) Demo, 2011.
- [97] Jing Zhu, Dan Li, Jianping Wu, Hongnan Liu, Ying Zhang, Jincheng Zhang. Towards Bandwidth Guarantee in Multi-tenancy Cloud Computing Networks. Proceedings of the IEEE ICNP 2012, Austin, TX.
- [98] Dan Li, Henggang Cui, Yan Hu, Yong Xia, Xin Wang. Scalable Data Center Multicast using Multi-class Bloom Filter. Proceedings of the IEEE ICNP 2011, Vancouver, BC Canada.
- [99] Dan Li, Ying Liu, Mingwei Xu, Xia Xie, Yong Cui, Guihai Chen. Reliable Multicast in Data Center Networks, Accepted by IEEE Transactions on Computers.
- [100] Haoyu Song. Protocol Oblivious Forwarding: Unleash the Power of SDN through a Future-Proof Forwarding Plane. Proceedings of the ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking (HotSDN), 2013.

作者简介

谢高岗 现任中国科学院计算技术研究所研究员，博士生导师，网络技术研究中心主任。主要从事可编程虚拟路由器、互联网体系结构与协议、网络测量的研究工作。



陈 鸣 现为解放军理工大学指挥信息系统学院网络工程教研中心教授，技术三级，博士生导师，军用网络技术实验室（军队重点实验室）主任，全军网络技术研究中心主任。担任中国计算机学会网络与数据通信专委会副主任和其他多个学术团体常务理事或委员。负责国家863计划、自然科学基金、国家科技攻关及军队重大项目多项，获省部级科技进步二、三等奖10多项，获国家教学成果二等奖1项，军队教学成果一、二等奖各1项，有国家发明专利8项。出版教材著作6部（一部获教育部2007年精品教材，一部填补网络工程专业教材空白）、译著8部，发表被三大检索收录的论文80余篇。获得全军优秀教师（两次）、军队院校育才奖“金奖”、军队优秀专业技术人才一类岗位津贴、总参优秀青年专家、江苏省优秀研究生导师等荣誉。研究方向是分布式计算、网络测量和网络管理等。



兰巨龙 现为国家数字交换系统工程技术研究中心教授，总工程师，博士生导师，国家下一代广播网（NGB）总体专家委员会委员，国家重点基础研究发展计划（973计划）项目首席科学家。一直从事IP网络理论与技术研究、核心设备开发工作，先后主持了3项国家863计划重大课题，1项国家973计划重大项目，1项国家支撑计划重大项目，并于2005年主持研制成功我国第一台全部核心技术拥有自主知识产权的高性能IPv6路由器。目前作为首席科学家主持973重大项目“可重构信息通信基础网络体系结构研究”，同时主持国家科技支撑计划项目“面向融合网络的大规模接入汇聚路由器关键技术研究与产业化应用”。先后获得军队科技进步一、二等奖、河南省科技进步一等奖、全军育才奖银奖等。



吴春明 浙江大学计算机系统结构与网络安全研究所教授、博士生导师。主要从事可重构柔性网络构建技术、未来网络体系结构、网络虚拟化等方向的科学的研究工作。近年来承担了多项863课题、973课题、国家自然科学基金、国家科技基础条件平台建设重点项目等的研究与开发工作。目前重点关注基于SDN架构的开放可重构网络试验平台技术。已发表学术论文60余篇，申请国家发明专利16项，出版著作2部。现担任《电子学报》编委会委员，“十二五”国家863信息技术领域网络与通信技术主题专家组专家。



李 丹 清华大学计算机系副教授，ACM、IEEE、CCF会员。2007年从清华大学毕业获得博士学位，主要从事计算机网络方向的研究工作。在ACM SIGCOMM、IEEE INFOCOM、IEEE ICNP、IEEE/ACM Transactions on Networking、IEEE Journal on Selected Areas in Communications、IEEE Transactions on Computers、IEEE Transactions on Parallel and Distributed Sys-



tems 等网络领域著名会议和期刊上发表论文 40 余篇。担任 IEEE LANMAN 2014 程序委员会主席、IEEE LANMAN 2015 大会主席、IEEE ICNP 2013 Poster/Demo 程序委员会主席、IEEE INFOCOM 2010-2014 程序委员会委员。国家 973 计划项目（青年科学家专题）“软件定义的云数据中心网络基础理论与关键技术”首席科学家，国家“十二五”863 信息技术领域网络与通信技术主题专家组秘书。

罗腊咏 2013 年在中科院计算所获工学博士学位，现为微软亚洲研究院副研究员。主要研究方向包括：可编程虚拟路由器、高性能网络数据包处理算法、数据中心网络等。



罗军舟 博士，东南大学计算机科学与工程学院教授，CCF 理事，CCF 网络与数据通信专委会主任，主要研究方向包括：下一代网络体系结构、网络安全和管理、云计算、无线局域网等。



李 星 清华大学电子工程系教授，博士生导师，清华大学信息网络工程研究中心副主任，中国教育和科研计算机网网络中心副主任。目前担任洲际研究型网络协调委员会 CCIRN 副主席，美国 Sigma Xi 科学学会会员。



李 伟 博士，副教授，现任东南大学计算机科学与工程学院/软件学院院长助理，CCF 高级会员、CCF 网络与数据通信专业委员会秘书长，主要研究方向包括：下一代网络体系结构、服务计算等。



徐明伟 分别于 1994 年和 1998 年在清华大学计算机科学与技术系获工学学士和工学博士学位。现为清华大学计算机系教授，博士生导师，网络所所长，863 计划“三网融合演进技术与系统研究”重大项目总体专家组成员，中国通信标准化协会技术管理委员会专家，研究领域为计算机网络体系结构、互联网路由和高性能路由器。



计算机辅助设计与图形学研究进展与趋势

CCF 计算机辅助设计与图形学专业委员会

摘要

计算机辅助设计与图形学是计算机科学的一个重要分支领域与应用方向。基于计算机辅助设计与图形学的理论和方法的一系列高技术产品和系统，在一定程度上改变了人们的工作和生活方式。产业和市场需求的驱动、硬件的发展，促使计算机辅助设计与图形学的研究人员不断探索新的应用领域、新的理论和方法。本文介绍了近两年来计算辅助设计图形学领域的研究热点和难点，重点介绍三维几何建模、数字城市建模、体网格生成与编辑，以及可视媒体四个领域的国内外研究动态及其应用，并进行了综合分析和对比，最后对中国的计算机辅助设计与图形学发展进行了展望。

关键词：计算机图形学，计算机辅助设计，基于知识的建模，体网格生成与编辑，数字城市建模，可视媒体

Abstract

Computer aided-design and computer graphics is one of the most active research streams in the development of computer science and technology. Various high- tech products and systems based on CAD&CG come into the daily life and impact people’ s work and life style. On the other hand, industrial and market demands and graphics hardware developments facilitate the researchers to explore new application area, novel theory and methodology. This report briefly introduces research hot topics and difficulties in computer-aided design and computer graphics recently. It mainly focuses on domestic and overseas research progresses of semantics-aware shape modeling, Digital City modeling, volume mesh generation and editing, visual media and their applications meanwhile gives a comprehensive analysis and comparison. The state of the art in the above fields is summarized and the future R&D directions are suggested.

Keywords: computer graphics, computer- aided design, knowledge- driven shape modeling, volume mesh generation and editing, digital city modeling, visual media

1 引言

计算机辅助设计与图形学经过 40 多年的发展，已经进入较为成熟的发展期。目前，其主要研究领域包括曲线曲面造型、产品建模、协同设计、面向领域的辅助设计方法与系统、图形绘制、计算机动画、可视化与可视分析、虚拟现实、图形硬件等。由于这些研究领域的推动，计算机辅助设计与图形学也有了新的发展机遇与挑战。

三维模型建模方法是计算机图形学的重要基础，而随着计算机图形技术的普及和发

展，各类用户都提出了高效的三维建模需求，其中三维内容生成（3D content creation）是当前计算机图形学的一个关键瓶颈问题^[1]。如何让专业人员甚至是普通用户快速构建复杂三维模型，是制约三维图形发展和普及的重要因素之一。随着三维建模与处理研究的深入发展，人们逐渐认识到，为更加有效地构建三维模型，需要面向特定应用需求，有效地分析、提取和利用三维对象所蕴含的丰富语义知识，为自动或交互式建模提供语义约束或智能辅助，即所谓知识驱动的三维几何建模（knowledge-driven shape modeling）。

知识驱动建模大致可分为两种方式。一是直接基于领域相关的先验知识，形成用于辅助和引导几何建模的语法规则^[2]，与功能相关的结构约束^[3]，或模板及相应参数^[4]。二是以现有三维模型作为输入，通过对模型的结构和功能分析，得到面向各种建模方法的不同语义约束。因此，该方法又称为基于样例的建模或数据驱动建模。本文将该方法称为感知语义的三维建模（semantics-aware shape modeling）。其核心是面向三维模型结构、功能及其相互联系的知识感知和发现，即对三维模型进行高层语义分析（high-level shape analysis）^[5]。传统几何处理与建模主要研究几何属性与度量，而感知语义关注面向功能和语义的三维模型结构分析。功能部件分析是模型结构分析的重要内容，包括三维模型的功能部件的识别和提取^[6,7]、模型间功能部件的对应关系^[8,9]，以及模型集中对应功能部件的形变参数等^[10,11]。

近年来，随着 Google Earth 3D、Microsoft Bing Map 3D 以及查查吧、E 都市等产品的推出，三维数字城市建模工作得到了更加广泛的关注，基于移动车载激光扫描和视频摄像的大规模城市场景数字化工作成为研究的热点，使得大规模城市建模等存在迫切需求。“数字城市”的概念是从“数字地球”发展而来的，是数字地球在城市环境中的应用。总体来讲，数字城市是现实城市的数字化及其延伸，但又有别于虚拟城市：数字城市以数据为中心，强调数据的采集、分析与应用，是现代城市数字化、信息化的集中体现，反映了一个城市现代化水平的程度。数字城市经过多年发展，无论在理论研究还是在应用实践上都得到了长足的进展，其突出特点在于：从二维到三维（2D Map to 3D World）、从静态到动态（Static to Dynamic）、从批处理到实时（Batch to Real-Time）逐步迈进的发展过程。目前，大规模城市场景三维建模有 4 类方法，分别是基于图片的单视点三维重建、多视点重建、基于语法的重建和基于激光扫描三维点云的重建。

作为工程科学与计算科学相交叉的一个重要研究领域，网格生成研究至今依然很活跃。网格生成是将连续的形体（曲面或者实体）离散成简单的基本网格单元（例如三角形或者四面体）。这些简单的网格单元是对原始形体的逼近，它们可以有效地支持偏微分方程（PDE）或者差分微分方程（ODE）的数值求解。由于网格质量和生成速度直接影响有限元分析的精度和效率，因此网格生成对于有限元分析而言至关重要。

根据离散对象的不同，网格总体上分为面网格和体网格。相比于面网格，体网格的生成要困难得多，这主要因为模型维度的增加导致模型复杂度大幅上升，造成体网格生成困难，同时随着模型变得复杂，结构化、各向异性、自适应等网格质量需求满足起来非常困难。体网格生成又分为四面体网格生成和六面体网格生成两类。由于四面体网格的拓扑限制较小，因此四面体网格的自动生成比较可靠稳定；而全六面体网格的自动生成

成因其严格的拓扑限制，虽然研究人员围绕体网格的生成与编辑开展了大量工作，取得了不少进展，但至今仍存在很多极具挑战性的问题，并在近年再度成为网格生成研究的热点。

视觉是人类认知世界最重要的手段之一，在人脑所处理的全部感觉信息中，超过70%的信息与视觉有关^[12]。承载视觉信息的媒体类型称为可视媒体，包括图像、视频、数字几何等。可视媒体的来源广泛，物理空间中数字化设备，如照相机、摄像机、激光测量、遥感卫星、传感器、监控系统等产生了大量的可视媒体。网络互联互通（互联网、传感网、物联网、通信网络、社会网络）的飞速发展，使得可视媒体随处可见，深入渗透到社会生活的各个领域。可视媒体已成为我们研究现实世界的重要手段，蕴含着无数影响人们认知方式和生活方式的重要应用。网络海量可视媒体的智能处理已经成为互联网信息高效利用与智能服务、面向三网融合的数字内容产业和泛在网络视觉智能感知等信息网络战略性新兴产业国家重大需求的共性基础。通过揭示人类感知、记忆和判断的认知机理，实现符合人类视觉认知机理和神经系统原理的认知计算模型，建立符合人类感知的可视媒体交互、融合与呈现的基础理论，提出新的可视媒体数据组织和分析方法，对建立网络环境下海量信息智能处理的理论体系，为海量信息分析整合与高效利用提供科学支撑，具有重要的科学意义。

2 国际研究现状

2.1 基于知识的建模/高层语义分析

知识驱动的三维建模和面向知识感知的高层语义分析，是当前图形学几何建模与处理领域的前沿和热点。在近几年的 SIGGRAPH 和 SIGGRAPH Asia 上，每年都有 2~3 个相关 Session。最近，Mitra 等人在欧洲图形学大会上发表了题为“感知结构的几何处理”的前沿进展报告，从结构分析与感知的角度介绍了高层语义分析与应用的最新进展^[13]。

（1）基于先验知识的建模

用以约束和引导三维建模的先验知识主要来源于人类对特定领域几何对象的结构认知，包括部件构成、组织结构及其与对象功能之间的关系等。由于不同领域三维对象的结构差异，先验知识及其在几何建模中的应用方式不尽相同。

过程式建模是应用最广的基于先验知识的建模方法之一。过程式建模将关于对象结构知识转化为基于语法规则的产生式模型，通过反复的规则替换生成结构各异的模型^[2,14]。有效的形状语法是过程式建模的基础。为某一类物体编写形状语法对用户的领域知识和编程技能有较高要求，因而过程式建模很难面向大众用户。

另外一种方法是利用先验知识形成显式的结构约束用以引导几何建模。Gal 等人通过分析三维模型特征线之间的结构关系（如对称、共面等），实现了人造物体模型的结构

保持编辑^[3]。基于同类物体的不同变种具有相似结构这一先验知识，将上述结构关系转化为结构约束，用于快速生成大量结构合理的模型变种。类似地，Zheng 等人基于部件结构关系实现了结构保持编辑^[15]。

(2) 感知语义的建模

普林斯顿大学 Funkhouser 等人于 2004 年发表的开创性工作“Modeling by Example（基于样例的建模）”首次提出了“内容重用”的三维建模思想^[16]。该工作对数据驱动建模，特别是三维模型自动合成起到了重要的启发作用。基于内容重用思想，结合对输入样例模型的结构分析，实现模型自动合成是快速生成大量三维模型的有效途径。

在单个物体建模方面，斯坦福大学的 Koltun 教授研究组基于大量有标注信息的三维模型，提出数据驱动的部件建议生成，以支持交互式的创造性三维建模^[17]。该方法可根据正在构建中的模型，自动从数据库中搜索相关部件作为建议，以引导和启发用户的设计和建模过程。随后，该小组基于同样的数据训练了描述模型功能部件上下文关系概率图模型^[18]，用于生成相关性更好的部件建议。最近，该方法被进一步扩展形成用于自动合成三维模型的产生式模型^[19]。

在三维场景建模方面，斯坦福大学 Hanrahan 教授小组基于 Trimble 3D Warehouse 提供的大量室内场景模型研究了场景物体间的上下文关系^[20,21]，用于指导基于样例的室内场景自动生成^[22]。基于场景分析得到的语义信息，人们还研究了室内场景点云的识别与重建^[23~25]，以及基于草图室内场景重建^[26]。

感知语义建模的另外一个方式是逆向过程式建模。不同于过程式建模，逆向过程式建模不需要用户来指定形状语法，而是通过对样例模型进行结构分析，自动提取出可以重现样例模型的语法规则。基于该形状语法，可以自动生成与样例模型风格相似的模型变种。逆向过程式建模的基础是对三维模型结构的精确分析，技术上具有很大难度。目前，逆向过程式建模还仅是围绕一些简单形体展开，如二维矢量图形^[27]和具有重复结构的规则物体^[28]。

(3) 高层语义分析

高层语义分析重点关注功能相关的结构提取与分析，包括三维模型功能部件的识别和提取，模型间部件的对应关系的计算，以及同类物体模型集合中对应部件的变化参数的学习。这三个方面构成了高层语义分析研究的主要内容。在研究方法上，基于模型集合的联合分析已成为主流，因为一组同类物体的三维模型比单个模型包含更丰富的语义信息。

三维模型功能部件的识别与提取最早源于单个模型的有意义分割。单个模型的分割一般基于先验知识构成的分割规则（如最小割规则）。2010 年，多伦多大学的 Singh 和 Hertzmann 教授小组首次运用监督学习方法得到了三维模型的有标注分割^[6]。2009 年，普林斯顿大学的 Funkhouser 教授研究组首次提出对一组同类物体三维模型几何的协同分割问题^[29]。受此启发，Xu 等人于 2010 年提出了联合分析的概念，并提出了基于无监督学习的三维模型集协同分割^[8]。上述两项工作掀起了联合分析的研究热潮^[7,9,30~33]。

在联合分析过程中，部件对应关系的计算往往与部件的分割同时进行。为避免部件分割的困难，有些研究人员采取直接计算模型表面点对应关系的研究思路。代表性工作有斯坦福大学 Guibas 教授研究组研究的基于环路增强的模型对应性计算^[34]。点对应关系虽然不是直接面向部件结构分析，但其结果可以间接地用于模型集中对应部件的变化参数学习，用以实现对大规模三维模型集的快速浏览^[11,35]。

综上所述，较之于传统的低层次形状分析，高层语义分析主要有以下特点：1) 研究内容由面向低层次应用的形状表示研究发展为面向更高层次应用的结构和功能分析；2) 分析对象由原来的面向单个物体的分析发展为面向一组模型的联合分析；3) 分析方法由原来的直接基于先验知识和启发式规则的过程式分析转向基于机器学习的数据驱动分析；4) 形状分析的应用已不仅局限于模型的检索与分类，而是广泛应用于几何处理和建模。

2.2 数字城市建模

数字城市的核心是城市数字化（Digitization），其任务是对城市实体进行几何、空间信息的测量和三维重建。近年来，随着 Google Earth 3D、Microsoft Bing Map 3D 以及查查吧、E 都市等产品的推出，三维数字城市建模工作得到了更加广泛的关注，基于移动车载激光扫描和视频摄像的大规模城市市场景数字化工作成为研究的热点。下面简要总结城市市场景三维建模方面的研究进展和重大应用成果。

(1) 基于图像的单视点三维场景重构

在三维城市市场景重构以及基于图像的城市建模方面近年来已经有许多研究成果。Müller 等人^[36]根据建筑物的重复样式得到建筑物表面结构，该方法是基于手动输入得到深度信息。通常，这些方法都需要较强的用户交互来生成合适的三维模型。Hoiem^[37]基于单幅图像提出了一种地面、竖直立面、天空等的不同布局的估计方法，用于粗略建模。Saxena^[38]直接从图像特征得到深度信息。Barinova^[39]利用建筑物的曼哈顿结构特点，将三维模型拟合重建问题转换为链图推理。但是，这些方法生成的建筑物模型比较粗糙，缺少细节。

(2) 基于图像的多视点三维场景重构

Hengel 等人^[40]在单幅或多幅图片上使用手绘的方法进行普通物体建模，该方法的缺点是即使加入很多交互操作，也难以进行精细建模。Xiao 等人^[41]基于近似垂直视点的视频图像序列，提出了 facade 细节自动重构和用户交互细化的方法，他们的方法需要对每个建筑物墙面进行精确的初始化和切割划分，甚至需要人工干预。Sinha 等人^[42]进行多视点的配准，根据灭点提取出主方向。这两种方法都需要大量的用户交互才能生成较好的结果，因此难以应用在大规模的城市市场景建模中。文献[43]给出了一个自动的街景逼真模型生成方法。一些团队提出了更加系统的方法，利用摄像机对城市环境进行建模^[44,45]。他们在实时视频配准方面取得了很多成果，重点研究基于稠密立体匹配的全局重构。此外，文献[46]介绍了谷歌在城市市场景重建方面的最新成果，文献[47]直接从十万多张社区照片中重建罗马。

(3) 基于文法规则的场景建模

基于文法的建模方法 (grammar-based modeling) 在植物和人工制品的建模方面逐渐流行。近年来, 图形学领域的学者们一直关注基于文法的建模, 因为它可以模拟真实对象的衍变过程。基于文法的建模方式有许多分支, 例如形状文法 (shape grammar) 和图文法 (graph grammar) 等。文法建模, 特别是形状文法建模, 非常适用于建筑物模型的生成, 它可以脱离 CAD 设计生成建筑物。形状文法可应用于建筑物设计的生成和分析。Müller^[48]将其扩展成为 CGA 文法, 增加了随机、遮挡查询等规则, 可让建筑物模型生成更加丰富自然的细节。道路生成方面, Alexander 等人提出了一种样式语言 (pattern language), 包括了超过 250 种的相关样式, 用以生成城市和建筑物。这个样式库不仅包含了环形道路之类通用的样式, 还提供了“石头与缝隙铺平道路”之类特殊的样式。Prish 和 Müller 发现路网是生成大型城市的关键, 他们提出了基于 L-systems 的路网建模方案, 该方案由一段初始道路开始, 逐步添加扩张成一个完整的城市路网, 其过程与树木生长的过程类似。为了利用更少的数据生成新的城市, 一些方法利用了城市中的关键因素, 例如道路样式、人口密度^[49] 和用户设计^[50] 等。

(4) 基于激光扫描数据的场景重建

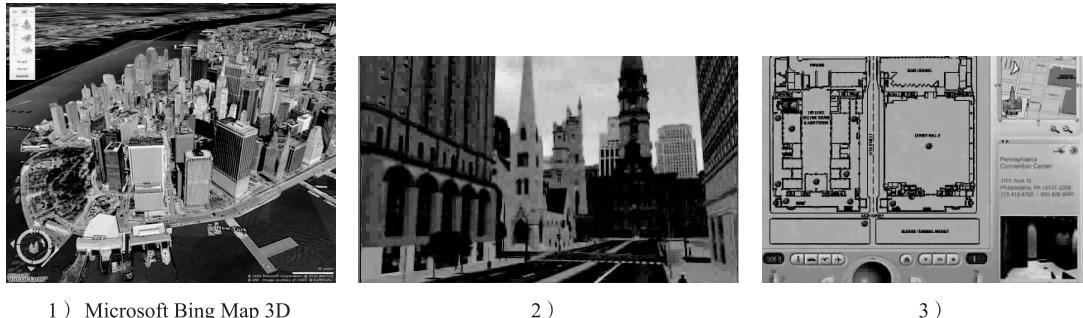
目前, 已经有很多项目研究如何自动采集和处理城市场景的三维数据。建筑物通常包含很多基本几何元素 (如平面、球面、圆柱和锥面等), 并且在结构上具有一定的规则性, 例如一栋楼各层的阳台、门窗等通常都是重复的。在对点云中的这些结构和特征进行识别方面, 已经出现了一些成果, 其中最具代表性的是 Pauly 等人^[51] 将识别三维点云中规则结构的问题转换为在变换空间拟合平面网格, 变换参数由任意对应面片的变换矩阵给出。文献[52]通过在点云中利用邻域信息拟合一些基本几何形状, 以此作为理想形体对点云的法向量、尖锐棱角等进行增强, 用于辅助重建。在基于激光扫描的城市场景重建方面, 陈宝权博士领导的团队进行了持续深入的研究, 提出了一系列解决方案。由于室外环境的限制, 完整而充分地对每一个建筑的每一个墙面进行扫描往往是不可能的。文献[53]提出了自动重建多棵重叠树木的方法, 不需要事先对重叠树木点云进行分割, 后续工作包括基于 Lobe 表示的树木三维建模方法, 该方法的核心思想在于树冠形状取决于外形轮廓, 而其内部的枝叶结构具有相似性, 可以基于树木种类从有限的模板库通过纹理合成的方法近似生成^[54]。

数字城市在国土测绘、城市规划、建筑设计、智能交通、生活娱乐、虚拟旅游、无线电通信、社会安全等众多领域具有大量的潜在用户, 市场应用前景广阔, 能产生很大的经济效益, 发挥重要的社会作用。下面就其在数字城市信息平台、文化遗产保护、城市规划、决策、应急指挥等方面的应用做简要介绍。

谷歌公司的 Google Earth 3D (见图 1), 微软公司的 Bing Map 3D、Skyline 和 Worldwind 等以及“美国费城虚拟城市”(见图 2)都是典型的数字城市平台。其中, 美国费城虚拟城市通过俯视、沉浸的方式, 游客能够身临其境般地在城市内行走观光, 不仅可以看到沿途的景点、商店和广告, 还可以进入建筑内部观光, 了解室内布局, 或在固定视点环视室内场景 (图 2)。



图 1 Google Earth 3D



1) Microsoft Bing Map 3D

2)

3)

图 2 费城虚拟城市及其室内浏览

除浏览漫游之外，三维数字城市建模在城市规划、应急指挥、文化遗产保护和决策等方面具有重要指导作用和意义。比如说，在应急指挥方面，2006 年联合国建立了“联合国灾害管理和应急控制信息平台”（UN-SPIDER），旨在普及各类与灾害管理有关的空间信息和服务；在文化遗产保护方面，既有希腊亚里士多德大学的古希腊 Polyphemus 巨人岩洞的三维建模项目，也有我国龙门石窟研究院与北京大学等合作的龙门石窟三维数字图书馆、武汉大学的敦煌莫高窟数字化项目等；在对社会行为分析决策方面，美国华盛顿大学在美国国家科学基金会（National Science Foundation，NSF）支持下研发的城市仿真模型 UrbenSim [<http://www.urbansim.org/>]，把人口、交通网络和土地状况融合经济趋势对住房、商用地、城市扩张等提供决策，已在休斯敦、盐城等城市进行了应用。

2.3 体网格生成与编辑

在国际上，有关体网格生成和编辑的研究近年来主要集中在六面体网格的自动生成与优化、四面体网格的优化和编辑，以及体网格生成的并行加速等几个方面。

为了攻克六面体网格自动生成与优化这一挑战性问题，人们进行着长期不懈的艰辛努力。有效识别扫成特征是实现基于体分解的六面体网格生成的关键一步，但这一问题目前还没有完全解决。对此，Lu、Jean Hsiang-Chun、Inho Song、William Roshan Quadros

和 Kenji Shimada 近来提出了基于中面分组的方法来识别扫成特征，并提供了友好的交互界面来进行手动分解^[55~57]。该方法具有一般性，其缺点则在于中面生成本身就是一项非常复杂且耗时的工作。同时，Jonathan E. Makem、Cecil G. Armstrong 和 Trevor T. Robinson 提出了一种有效识别并分解出模型中狭长区域的方法，这部分区域可以生成结构化的各向异性的六面体网格^[58]，但是该算法的针对性太强，无法用于生成模型整体的六面体网格。子部分网格化后需要重新组成完整网格，这时就需要调整子部分之间的网格，使之变成拓扑一致。对此，Matthew L. Staten、Jason F. Shepherd、Franck Ledoux 和 Kenji Shimada 提出了 Hexahedral Mesh Matching 算法^[59]，该算法利用对偶操作，能够将相邻面上的网格逐步调整成拓扑一致。之后，Matthew L. Staten 在其博士论文中提出了对该算法实现自动化的改进方案^[60]。另外，James Gregson、Alla Sheffer 和 Eugene Zhang 提出了一种基于八叉树的 PolyCube 方法^[61]，能够大幅度减少体网格中扭曲变形的单元数量，但是该方法仍然无法避免八叉树方法的本质缺陷，即对模型空间摆放位置的敏感性以及会在网格表面生成奇异点。Eloi Ruiz-Gironés、Xevi Roca 和 Josep Sarrate 提出了一种将八叉树和波前法结合的生成算法，能够在模型内部和表面都生成高质量的网格单元^[62]。然而该方法仍然不够鲁棒，例如在处理剩下的小空间的时候，它会调用有限的模板，无法处理所有情况。由于六面体的空间复杂性，其无法直接使用四面体优化中很多常用的方法。T. J. Wilson、J. Sarrate、X. Roca、R. Montenegro 和 J. M. Escobar 在提出新的优化目标函数的基础上将四面体网格优化算法推广到了六面体网格^[63]。此外，针对六面体网格特有的内在结构，Jason F. Shepherd 和 Matthew L. Staten 等学者提出了用 sheet 操作来对网格进行拓扑优化^[60,64]。Engin Yalcin、Asim Egemen Yilmaz 和 Mustafa Kuzuoglu 对当前几种主要的六面体网格优化算法进行了综合比较，并分析了它们各自的优缺点^[65]。

在四面体网格的优化与编辑方面，针对传统网格优化算法无法有效作用于高次体网格的问题，Jean-François Remacle、Thomas Toulorge 和 Jonathan Lambrechts 提出了一种鲁棒的高次体网格优化算法^[66]，能够有效地去除扭曲网格。但是该方法对于网格其他方面的提高，例如网格密度、点和边的度数等，缺少优化效果。Hiroki Maehama、Hiroaki Date 和 Satoshi Kanai 则针对在设计阶段 CAD 模型不断变动的情况下网格模型快速生成的问题，提出了基于最优 Delaunay 三角化的网格调整优化算法^[67]。在四面体网格编辑方面，Mirela Ben-Chen、Ofir Weber 和 Craig Gotsman 将变分调和映射的方法应用于四面体网格的自由形变^[68]，在变形的过程中能够很好地保持模型的刚性。然而这类方法难以定义用户约束，也很难保证编辑后网格的质量。M. Sypkens Smit 和 W. F. Bronsvoort 利用边界表示模型的变动指导网格编辑修改，能够实现网格模型的局部重新生成，并可以保证网格编辑精度^[69]，然而其需要边界表示模型的支持，并且难以进行网格的自由变形。Suzanne M. Shontz 和 Stephen A. Vavasis 改进了基于有限元网格的 FEMWARP 方法，利用有限元单元的形函数进行内部结点的移动^[70]，Daniel Sieger、Stefan Menzel 和 Mario Botsch 使用了基于三次调和径向基函数的方法对四面体和六面体进行网格变形，这两种方法都可以有效地保证网格质量并易于实现^[71]。然而这些方法仅考虑了结点的移动，没有考虑网格单元数量的变化，对于大幅度的网格变化，则难以保证其网格质量。

2.4 可视媒体

可视媒体本身具有数据量大、非结构性、维数高、语义多样性的特点。网络环境下，可视媒体突显新的重要特性，呈现出无序、涌现、可靠性低、多态共生和多维度关联等新特征，使得网络可视媒体分析与处理的复杂性急剧增加。在当前网络普及和 Web 2.0 的互联网应用潮流下，每个人都成为可视媒体内容生产与获取者，例如，Facebook 上已分享了超过 100 亿张照片，YouTube 每天上传视频量超过 2 万小时，中国国家网络电视台（CNTV）每年将通过网络播出 27 万小时的电视节目。然而面对极大丰富的网络海量可视媒体资源，目前仍然缺乏对其进行高效利用的有效途径。网络海量可视媒体的智能处理，从信息处理的各个环节着手，广泛地模拟人的智能和社会行为，建立可实现的认知计算模型和海量数据组织处理方法并发展应用，有可能带来未来信息处理技术突破性的进展。

作为计算机图形学、数字媒体处理、机器学习和认知分析等交叉学科的前沿，可视媒体的智能处理在多个国际著名会议和期刊上都得到了广泛的关注，会议包括 ACM SIGGRAPH、ACM Multimedia、CVPR、ICCV 和 ACM CHI 等，期刊包括《ACM Trans. Graphics》、《IEEE Trans. Pattern Analysis and Machine Intelligence》、《IEEE Trans. Visualization and Computer Graphics》、《IEEE Trans. Multimedia》和《ACM Transactions on Computer-Human Interaction》等。以 ACM SIGGRAPH、ACM Multimedia 为例，在 2012 年 ACM SIGGRAPH 的 93 篇论文中，除了 4 篇与声音有关的论文外，其余全部与形状、动画、渲染、图像视频处理和显示有关。在 ACM Multimedia 2012 的 163 篇长文中，除了 4 篇关于声音和音乐的应用、3 篇关于触觉的应用，其余全部是关于图像视频的搜索、大数据分析、数字媒体语义标注与建模、社会媒体分析以及数字媒体的各种新颖应用等。

3 国内研究进展

3.1 基于知识的建模/高层语义分析

本节分别介绍国内学者近年来在知识驱动建模和高层语义分析方面的研究工作和主要贡献。

(1) 知识驱动的建模

三维几何建模主要包含 4 种方式：交互式建模、点云获取与三维重建、过程式建模和模型自动合成。

1) 知识驱动的交互式建模。交互式建模既可以直根据领域知识形成建模约束，也可以通过感知语义的方式从样例中学习结构和语义信息用于约束和引导建模。微软亚洲研究院（以下简称“亚研院”）的许威威等人提出了关节敏感的形变技术^[72]，发表于

SIGGRAPH 2008。该方法对输入模型的关节属性进行分析，形成关节约束，实现了对带关节人造物体的交互式智能编辑。在交互式建筑物建模方面，中国科技大学的陈雪锦与亚研院合作提出了一种基于草图的建筑物建模方法 Sketching Reality^[73]，发表于 ACM TOG。该方法由矢量化的建筑物二维草图出发，以基元几何、细节几何和纹理样本构成的数据驱动为驱动分析和解释二维草图的几何和结构信息，生成可用于真实感绘制和交互浏览的 2.5D 建筑结构。

2) 知识驱动的三维重建。基于激光扫描点云数据的三维重建是几何建模领域的经典问题。对于一些特定物体，由于其几何形状和拓扑结构的特殊性，通用的重建方法往往不能奏效，需要借助一些先验知识实现高精度建模。例如，中国科技大学的刘利刚教授团队研究了基于局部形状先验知识的三维管状物体建模^[74]，发表于 SIGGRAPH Asia 2010。中科院深圳先进技术研究院（以下简称“先进院”）的李扬彦等人针对机械部件的低质量点云，采用同时进行基元拟合和基元全局关系优化的方法，实现了高精度建模^[75]，发表于 SIGGRAPH 2011。最近，刘利刚教授与美国肯塔基大学合作研究了住宅场景点云的语义分解与重建^[76]，发表于 SIGGRAPH 2013。该方法首先采用监督学习方法对输入场景进行分类，然后根据特定类别的对象，基于领域知识进行三维重建。

近年来，LiDAR 技术的发展成熟使城市级大规模三维场景的数据获取成为可能。基于 LiDAR 数据的大规模场景建模成为三维重建的重要子方向。城市建筑物的结构往往非常复杂，同时由于数据获取设备和环境的限制，建筑物点云的分辨率一般较低且数据缺失较严重，这些为传统三维重建方法带来了新的挑战。但同时，建筑物的结构组成蕴含着丰富的结构知识和语义信息，因此非常适合采用知识驱动方法进行建模。知识驱动的三维重建成为大规模点云三维重建的重要方法。近年来，中科院先进技术研究院的陈宝权教授团队在知识驱动的三维城市建模方面的研究成果瞩目，有多篇文章发表于 SIGGRAPH 和 SIGGRAPH Asia。代表工作有：南亮亮等人提出的基于立方体基元拟合的交互式建筑物点云重建^[77]；郑倩等人研究的基于规则性检测的建筑物点云增强处理^[78]；Livny 等人基于预先训练的树种知识库实现的树木点云的结构分析和三维重建^[79]。

清华大学胡事民教授小组研究了基于单视点深度摄像机捕获的低质量点云的单个人造物体三维重建^[80]，发表于 SIGGRAPH Asia 2012。该方法以一个小规模样例模型集作为基础，通过基于图像的部件检索和自底向上的部件组装完成建模。在同一会议上，先进院和亚研院分别发表了基于低质量点云的室内场景三维重建工作^[24,25]。二者都属于数据驱动的重建方法。所不同的是，前者采用点云聚类的方法进行物体识别，而后者采用交互式场景分割。

3) 知识驱动的过程式建模。过程式建模的基础是精简有效的形状语法。逆向过程式可自动生成语法规则，但依赖于对输入样例模型结构的精准分析。这对于建筑物模型具有很大挑战性。最近，中科院先进院的徐凯等人研究了建筑物立面图像的层次化结构分析^[81]，发表于 SIGGRAPH 2013。该方法结合 Gestalt 规则（分解过程中的对称最大化原则）和最简解释原则定义层次结构的优化目标函数，通过全局优化方法得到最优层次分解。该层次分解可用于结构保持的建筑立面图像编辑和自由缩放，以及结构相关的建筑

立面图像检索。

4) 知识驱动的模型自动合成。国防科技大学的徐凯等人基于同类人造物体三维模型集合的结构分析,提出了一种图像启发的数据驱动三维建模方法,发表于 SIGGRAPH 2011。基于单幅图像的人造物体三维建模是传统性难题。该文提出一种数据驱动的方法,借助一个与图像中物体同类的三维模型库,通过模型集的协同分割和单个模型的结构分析,提取模型库中的语义信息。由于模型库与图像中物体属同类,其结构具有共性,因此可以利用三维模型库语义信息,实现语义驱动的三维建模。基于三维模型集合的联合语义分析,徐凯等人进一步提出了三维模型集合的自动演化^[82],发表于 SIGGRAPH 2012。该方法将遗传算法的思想引入三维几何建模:将单个三维模型看做生物个体,模型集合看做生物种群;定义了模型个体的交叉和变异两种基本遗传操作;采用用户交互的方式确定新生个体(合成三维模型)的适应性。该方法可用于快速生成大批结构新颖且符合用户需求的三维模型。

清华大学的徐昆等人提出了一种基于草图的数据驱动室内场景建模方法^[26],发表于 SIGGRAPH 2013。该方法由室内场景三维模型库出发,通过对场景中物体的空间结构关系进行联合分析,得到表示场景集中稳定出现的、描述物体功能与空间关系的结构组。基于此,他们提出了基于草图的场景物体协同检索与协同放置,实现了功能-空间关系合理、物体形状风格一致的场景布局。

(2) 高层语义分析

高层语义分析包括三维模型结构功能分析和三维模型集联合语义分析。

1) 三维模型结构功能分析。对三维模型的结构分析主要涉及对模型的组成部分的分析,包括部件分割,以及对应、对称等部件关系的提取。在三维模型中,对称性与模型功能和语义密切关联:模型中自对称的部分往往是具有独立语义的部件;对称部件之间往往具有相同的功能。而对称性可完全依赖于几何方法度量和检测。可以说,对称是连接低层次几何属性与高层次功能和语义的桥梁。

徐凯等人研究了三维模型局部内蕴对称检测,并基于检测出的对称得到了传统方法不能达到的有意义分割结果^[83],发表于 SIGGRAPH Asia 2009。随后,他们进一步研究了三维模型的多尺度局部内蕴对称检测,并实现了多尺度对称驱动的三维模型多层次分割^[84],发表于 SIGGRAPH Asia 2012。国防科技大学的王彦臻、徐凯等人对人造物体的三维模型提出了对称层次的概念^[85]。该方法以对称性作为指导,基于人类关于人造物体部件组装和对称相关的结构知识,建立了一系列分析规则,得到了模型部件的组织和有意义层次分割。基于对称层次得到的层次分割很好地反映了人造物体的功能部件分解。

2) 三维模型集联合语义分析。一组同类物体的三维模型比单个模型包含更丰富的语义信息:同类物体往往共享相同的功能部件。因此,对同类物体模型集合的一致分割可以反映该组模型的功能部件分解与对应,这对于分析和理解模型的结构和功能具有重要意义。

徐凯等人研究了同类人造物体的三维模型集合的联合语义分析,提出了一种基于形状风格分析的联合分析框架^[8],发表于 SIGGRAPH Asia 2010。他们将形状风格定义为模型部件的各向异性尺度,基于无监督学习方法对输入模型集合进行形状风格聚类,再通

过类内模型的一致分割以及类间模型的部件对应，实现了模型集的一致分割。先进院的汪云海等人研究了基于半监督学习的模型集一致分割^[86]，发表于 SIGGRAPH Asia 2012。该方法从 Diffusion Map 聚类得到的初始分割出发，通过用户交互指定分割片是否属于同一对应部件，构成特征空间中的连接约束，基于这些约束不断更新特征空间中的聚类形成新的分割结果。此外，该方法还实现了不确定分割片自动建议，以方便用户更快地完成交互分割。类似地，浙江大学黄劲副教授小组基于条件随机场定义了分割目标函数，同时考虑了有标注和无标注的训练数据，在有标注训练集规模较小的情况下得到了较好的分割结果^[32]。刘利刚教授小组提出了基于子空间距离的模型集一致分割方法^[31]。该方法将一致分割描述成多特征空间的子空间聚类问题，并通过稀疏编码的方法实现特征选择。汪云海等人提出了基于多特征融合的三维模型协同分割方法。该方法对每个特征进行加权，显式地构造相似度量矩阵，将特征选择和谱聚类纳入了统一框架^[33]。

2013 年，徐凯与西蒙弗雷泽大学等合作，针对同类人造物体三维模型集合，提出了一种迭代式的聚类 - 代表树选择算法，用以计算一致性对称层次树，实现了三维模型集的层次化联合分析^[87]，发表于 SIGGRAPH 2013。

3.2 数字城市建模

在基于规则的过程式建模研究领域，国内研究学者也取得了出色的成绩。Liu 等人^[88]做了一系列基于图像的对称检测的工作，这对城市建筑物表面建模非常有用。Xiao 等人^[41]基于近似垂直视点的视频图像序列，提出了 façade 细节自动重构和用户交互细化的方法，他们的方法需要对每个建筑物墙面进行精确的初始化和切割划分，甚至需要人工干预。文献[43]提出了一个自动的街景逼真模型生成方法。

在大规模场景数据获取、高精度实体重建方面，中国科学院深圳先进研究院可视计算研究团队从 2008 年创建以来取得了丰硕的研究成果，处于国内领先地位。获取城市场景的三维描述和漫游是一个新颖并充满挑战的课题，也是构建智慧城市的数据基础。

在基于激光扫描的城市场景重建方面，陈宝权博士领导的团队进行了持续深入的研究，提出了一系列解决方案。由于室外环境的限制，完整而充分地对每一个建筑的每一个墙面进行扫描往往是不可能的。陈宝权等在文献[73]中采用自顶向下的建模理念，从缺失的扫描数据中实现建筑物三维建模。激光扫描技术的最新进展使得可以对城市场景进行快速扫描，但是也会造成大区域的数据丢失、采样密度的变化、噪声和异常数据的存在等。通过检测，提取和利用建筑物的重复性和自相似性，可以增强不完整扫描的城市建筑物数据^[89]。陈宝权等人还开发了一套交互式工具 SmartBoxes^[90]，使用户能够从大规模的城市场景扫描得到的点云出发，通过简单勾画和交互将高层次的语义描述来引导建模，快速组装建筑模型。植物在城市环境中无处不在，文献[91]开发了一种半自动的高效的方法来重建树木模型，建好的模型能够恢复点云中丢失的细节。在文化遗产保护领域，有我国龙门石窟研究院与北京大学等合作的龙门石窟三维数字图书馆、武汉大学的敦煌莫高窟数字化项目等。

3.3 体网格生成与编辑

近年来，国内相关研究人员围绕体网格生成与编辑开展了大量工作，取得了很多成果，限于篇幅难以系统全面地进行综述。这里仅重点对六面体网格生成、四面体网格编辑优化等方面的最新研究成果进行简介。

在针对自由形体的高质量六面体网格生成算法研究方面，国内研究人员取得了可喜的突破。鉴于对称方向场指导下的全局参数化技术在曲面四边化中取得的成功，Huang Jin、Tong Yiyang 等学者在 2011 年发表的两篇重要论文“Boundary Aligned Smooth 3D Cross-Frame Field”^[92] 和“CUBECOVER-Parameterization of 3D Volumes”^[93] 中分别首次提出了关于体内对称方向场的构造方法和整体光顺的六面体参数化方法，开始尝试将这种策略推广到三维实体模型，并取得了初步成功。随着研究的深入，人们发现方向场中奇异线结构可能存在矛盾，常常导致参数化中的退化问题，使得生成的六面体网格中有缺失单元。针对这一重要问题，2012 年国内学者 Yufei Li、Yang Liu、Wewei Xu、Wenping Wang 和 Baining Guo 在“All-Hex Meshing using Singularity-Restricted Field”^[94] 一文中提出了局部检测和矫正内部奇异线矛盾的方法，可生成满足六面体化要求的体方向场，大大提高了这类六面体网格生成方法的实用性。但是，该方法局部修正奇异线的方式无法保证奇异线全局拓扑与六面体网格同构。最优拓扑结构的六面体网格化仍然存在很多的困难，比如全局拓扑一致性等。因此，Mengfei Li、Ruofeng Tong 的论文“All-Hexahedral Mesh Generation via Inside-Out Advancing Front based on Harmonic Fields”^[95] 提出了一种扫描体结构六面体网格半自动生成算法。该方法首先构造梯度互相垂直的轴向和径向调和场，然后从中心开始沿径向往外逐步扩张直到填满整个模型。该方法虽然不适用于任意形状的模型，但其简单可靠，在一定的人工交互干预下能生成高质量的六面体网格。此外，Jinming Chen、Hua Zhu、Shuming Gao 和 Haiyan Wu 在基于体分解的六面体网格生成算法研究方面，对现有的六面体网格匹配算法提出了多项改进^[96]，改进后的算法能够处理包含内孔的情况，并具有更高的性能。

在体网格优化算法研究方面，针对优化过程非常耗时，优化效率亟待提高的问题，国内研究人员大胆尝试，在提高四面体网格优化效率和六面体网格优化质量方面取得了较大进展。Xiaoshen Chen、Dewen Peng 和 Shuming Gao 首次将机器学习技术引入有限元网格拓扑优化之中，提出基于样本学习的有限元网格拓扑优化方法^[97]，通过机器学习大幅度减少网格拓扑优化过程中进行无效拓扑操作的数量，从而显著提高了有限元网格拓扑优化效率。该方法为提高有限元网格拓扑优化效率提供了一条新途径。Dewen Peng、Shuming Gao、Hang Si 和 Qiqiang Zheng 首次将 GPU 并行计算引入四面体网格优化，分别提出了基于 GPU 加速的四面体网格几何并行优化算法以及面翻转、边收缩、点插入等并行拓扑优化算法^[98]，通过 GPU 并行计算显著提高了有限元网格优化效率。与传统算法相比，基于 GPU 加速的四面体网格并行优化算法可以将复杂网格模型的优化效率提高数十倍以上，并且网格规模越大，效率提高越明显。针对基于栅格法生成六面体网格所存

在的质量问题, Lu Sun、Guoqun Zhao 和 Xinwu Ma 提出了基于改进的拉普拉斯算法的网格优化算法^[99,100], 对栅格法生成六面体网格质量有较大的提高。

为了克服基于空间编辑技术的体网格变形质量难以保证的问题, Wenhao Song 和 Ligang Liu 提出了一种非线性的最优化编辑方法, 以用于进行四面体网格的变形操作^[101], 通过综合考虑四面体变形中的扭曲程度和体积, 提高了网格变形的质量, 但同时也会耗费大量的计算时间。Shenghui Liao、Ruofeng Tong、Jinxiang Dong 和 Fudong Zhu 也提出了基于拉普拉斯坐标和梯度场的变形方法^[102]。Chuhua Xian、Shuming Gao 等学者提出了一种体网格特征生成方法和基于特征的体网格编辑方法^[103,104]。在给出体网格特征表示的基础上, 从表面网格特征出发, 提出使用最小通量原则生成体网格特征的方法。这个方法能高效生成高质量的体网格特征之间的分割面, 并且能同时处理四面体网格和六面体网格。基于体网格特征, 进一步提出了基于插值的体网格编辑方法, 支持对局部特征进行高效编辑。Hua Zhu、Shuming Gao、Ming Li 和 Wanbin Pan 提出了一种基于网格重用的变动模型四面体网格快速生成方法^[105], 支持基于设计变动的四面体网格高效编辑。该方法根据网格密度的变化, 确定网格重划分的范围和可重用网格的范围, 能够对变动模型实现高效高质量的网格重划分, 同时能够处理复杂的网格剔除情况。

3.4 可视媒体

在国内的高等院校和科研院所中, 2012 年浙江大学和清华大学在可视媒体研究中占有重要地位。以 ACM SIGGRAPH 2012 为例, 浙江大学参与了 5 篇论文的工作, 清华大学参与了 3 篇论文的工作, 这些工作涉及液体模拟、变形体网格的碰撞检测、碰撞检测中的连续作用力、交互式图像编辑和头发建模等。在 ACM SIGGRAPH Asia 2012 中, 国内参与单位广泛, 包括清华大学、北京大学、中科院深圳先进技术研究院、中国科学技术大学、杭州师范大学、国防科技大学、山东大学、北京航空航天大学、中科院软件所等, 参与的论文有 13 篇, 另外有 3 篇来自中国台湾。

面向网络海量可视媒体智能处理的理论与方法, 国内清华大学、北京大学、上海交通大学、中山大学和中科院心理研究所等有关单位, 依托清华大学信息科学与技术国家实验室, 在 2012 年从以下 6 个方面开展了研究: 1) 可视媒体的认知计算模型; 2) 可视媒体的结构分析与机器学习; 3) 网络可视媒体的语义分析与信息整合; 4) 网络可视媒体安全; 5) 网络可视媒体的交互与合成; 6) 网络可视媒体的有效搜索与服务。取得的突出进展包括:

1) 中科院心理研究所开展人类视觉感知、记忆与判断研究, 发现视觉显著特征影响注意的分布和转移, 并影响记忆的编码和提取^[106]; 记忆和判断过程相对分离^[107]; 动态视觉信息的感知影响决策, 并构建了人类速度感知决策模型^[108]。

2) 在可视媒体结构分析计算研究中, 清华大学研究了二维流形上的 Voronoi 图结构, 并探讨了精确测地线度量下 Voronoi 图的代数结构和组合复杂度^[109], 设计了测地路径查找的快速算法^[110], 并应用到可视媒体分类、检索与识别等任务中; 提出一种新的超限

插值（Transfinite Interpolation）方法泊松坐标，是著名的均值坐标的推广形式^[111]。

3) 清华大学开展了可视媒体结构分析研究，提出了一种基于样例的 Kinect 点云数据的结构分析方法，可以快速地将低质量的 Kinect 扫描数据转换成高质量的三维几何模型^[80]；从几何的角度对 popup 进行研究，对其独特的几何结构进行分析，研究了生成 v-style popup 的基本结构，研发了一个交互式的 v-style popup 几何造型工具^[112]。

4) 中山大学开展了可视媒体内容保护研究，提出了针对不同媒体的一系列高鲁棒性的数字水印算法，能有效地提高水印的鲁棒性，尤其是抗几何变换的鲁棒性和不可察觉性^[113]；在安全的信息隐藏方法上，已经提出 JPEG 图像通道选择方案，它和同类算法相比有更高的安全性^[114]。

5) 清华大学开展了网络可视媒体的交互与合成研究，提出了基于全局对比度分析的图像视觉重要性区域检测^[115]，构建了基于互联网的人像数据库^[116]；提出了基于语义的可视媒体一致性编辑与融合方法^[117]，提出一种基于内容的鱼眼视频/图像矫正方法^[118]和基于运动对象时空关系的可视媒体编辑^[119]，与浙江大学合作研究了基于几何信息的可视媒体交互编辑技术^[120]，与中科院软件所合作提出了基于草图的视频表示与组织方法^[121]。

6) 清华大学开展了网络可视媒体的有效搜索与服务研究，提出了一种全新的基于社会网络用户传播的视频内容分发机理，突破了传统的基于内容流行度的视频内容分发方法。该方法对于社会网络环境下视频分享机制具有极其重要的理论意义和应用价值，可以提高目前流行视频网络系统性能 2~4 倍^[122]。研究论文获得 ACM Multimedia 2012 的最佳论文奖。

4 国内外研究进展比较

4.1 基于知识的建模/高层语义分析

虽然知识驱动建模已成为三维几何建模的主流方法和必然趋势，但整体上看现有研究仍处于初期阶段，一些重要的理论和方法（如三维模型语义知识表示、高效的感知语义方法、知识驱动建模的数学描述等）尚不完善。

知识驱动建模面临的一个非常重要的问题是，如何构建适合于知识感知、表示与运用的三维几何表示方法。传统的三维几何模型表示（如三角网格、边界表示、隐式曲面等）主要关注模型表面几何外形和拓扑结构。尽管这些表示非常适合于渲染、几何处理等任务，但并不适合于知识的表示和编码。以最常见的三角网格为例，三角形单元并不具有独立的语义。因此，需要提出可高效、自然地融合语义信息的几何表示方法。理论上，这种表示应既具有精确的几何编码能力和灵活的拓扑处理能力，又具有对模型的结构信息、领域知识进行高效编解码的能力。

近几年，一些研究人员在结构相关的模型表示方面做出了一些有益尝试。例如，Gal 等人对人造物体模型提出了基于特征线的结构编码方式，将模型结构信息表示为特征线的相互关系^[3]。Zheng 等人将模型部件抽象为若干种基本几何元，用基元之间的相互关系编码部件级的结构信息^[15]。这些方法已被成功应用于人造物体模型的结构保持编辑。然而，其缺点也显而易见：首先，在模型编辑过程中，结构关系作为离散的约束条件和启发式规则，并没有统一的数学描述，因此“结构保持”缺乏明确定义的优化目标；其次，这些离散表示方式缺乏灵活的拓扑处理能力，使得模型编辑无法生成新颖的结构；最后，这些表示方式仅能处理人造物体模型，很难做到通用。因此，如何高效、形式化地组织结构信息，并形成统一的知识驱动建模框架，是新型几何模型表示方法亟待解决的问题。

三维建模很大程度上体现了建模人员的创造性，因而创造性建模是几何建模的重要方面。国际上，斯坦福大学的 Koltun 教授小组研究了开放式建模（即建模过程既没有起始也没有目标，完全依赖于用户的创造性导向），采用数据驱动的方法自动为用户提供部件建议，以启发用户的创造性思维^[17]。中科院先进院的陈宝权教授团队研究了基于草图输入的交互式创造性建模^[123]。该方法与斯坦福工作的相同点是都采用数据驱动方法生成部件建议。不同的是，该方法结合了用户草图输入和当前模型的上下文进行候选部件检索，使用户不但可以通过模糊草图检索得到的候选部件来探索设计空间，也可以在设计确定后利用明确草图表达精确设计意图。

如何有效地从样例中学习和感知语义信息是感知语义建模的关键。当前，语义感知面临的主要困难是用于结构和语义学习的训练数据十分缺乏。三维建模本身十分困难，生成可用于训练数据的、具有语义标注信息的三维模型更为困难。尽管已有一些模型库提供了标注信息（如 3D Warehouse），但由于标注过程繁琐且易错，而不同人的标注结果也很难做到统一，现有数据库的标注信息尚且很难直接用于语义学习。因此，训练数据缺乏是制约高层语义分析的重要因素。针对这一问题的主要解决途径主要有两种：一是基于监督学习的三维模型标注分割^[6]；二是采用半监督学习，同时利用有标注和无标注数据进行标注分割^[7]。另外一个可能的思路是利用大量存在的二维图像作为训练数据，即结合二维和三维数据的联合语义分析。

尽管三维内容生成问题仍尚待解决，三维模型获取和建立的效率也远未达到大数据阶段，但是三维模型的数量呈爆炸式增长已是一个不争的事实。如何高效地组织、维护、浏览和使用大规模三维模型数据库已经成为几何处理领域的一个重要课题。斯坦福大学的 Guibas 小组提出了基于模板形变的三维模型集变化建模^[124]，用于感知和浏览模型数据库的部件级变化。类似工作还有普林斯顿大学 Funkhouser 小组的基于模糊点对应^[35]和基于模板拟合^[11]的模型集浏览。在国内，胡事民教授小组与国际学者合作借鉴生物信息学中基于局部四元组构建种系树的方法，研究了大规模、异构三维模型集的可视化组织和摘要浏览^[125]。

感知语义建模的核心是语义增强的数据库。构建大规模三维模型库，特别是带有语义标注信息的三维模型库对于感知语义建模具有十分重要的意义。一些欧美发达国家构

建的商业或公开三维模型数据库已经初具规模。例如，美国的大型商业三维模型库 TurboSquid 维护了 30 万个高质量三维模型（统计截至 2013 年）；Trimble 3D Warehouse 为专业和非专业用户提供免费的三维模型共享平台，维护了 3000 万个由 SketchUp 构建的三维模型（统计截至 2011 年）。欧洲维护了全球最大的扫描重建三维模型数据集 AIM@SHAPE，并面向全球免费开放。与之相关联的是三维模型语义知识库的构建。目前，3D Warehouse 已经为大多数三维模型提供了标注信息，但尚未形成完整的语义知识库。类似的大规模数据集在我国尚未出现。中科院先进院的陈宝权教授团队建立了第一个用于评测协同分割的三维模型数据集 COSEG^[7]。

4.2 数字城市建模

据 2012 年《中国城市发展报告》的数据显示，全国城镇人口达到 71182 万，城市总数为 658 个，城镇化水平已经达到 52.57%。城市的快速发展给城市人口、交通、环境等日常生活等方面带来了巨大挑战。为了应对这种挑战，目前国内已有几十个大中城市提出或正在实施“数字城市”工程。国内的高校、科研机构也纷纷开始数字城市研究工作。与国外相比，国内的数字城市建设目前主要面向政府管理，公众参与度相对较低。以数字纽约为例，Rachel Sterne 作为该项目领导在他的《数字城市蓝图》中特别强调从四个方面建设数字纽约：更广泛的数字信息享用权、政府应用接口平台、提高城市吸引力和促进城市科技工业发展。数字纽约在实施过程中非常注重公众的参与，倡导建立政府和公众充分互动的数字城市公共平台。

数字城市的建设是一个综合性的、复杂的、长远的系统工程，目前国内外已经在运用数字城市提升政府管理能力方面做出了可喜的研究与实践，但如何在城市数字化过程中有效融合与利用现有城市资源，如果使数字城市得到广大公众的认可与逐步普及仍然是非常具有挑战性的问题。

4.3 体网格生成与编辑

随着四面体网格自动生成技术相对成熟，同时鉴于六面体网格在计算精度、变形特性、划分网格数量、抗畸变程度及再划分次数等方面具有明显的优势，甚至对有些工程计算问题，只能采用六面体网格进行有限元分析，因此近年来全六面体网格自动生成这一难题重新成为国内外的研究热点。研究工作主要集中在复杂形体的整体优化六面体网格自动生成和基于体分解的六面体网格自动生成两个方面。

复杂形体的高质量六面体网格自动生成研究因受到近年来进展很大的复杂形体整体优化四边形网格生成研究的推动而引起了人们的极大重视。人们希望将现有的整体优化四边形网格生成方法推广到六面体网格自动生成。然而，整体优化的六面体网格生成越来越多地涉及对称群、整体微分几何、图论等领域，成为了理论上很深入的一个研究课题。尽管当前这一技术已有一些初步工作，然而大都缺乏理论性分析，可靠性不足。采

用各种工程的技巧来解决其中的各种问题在一定程度上能推动其实用化，然而在一定的理论框架下系统地解决这些问题更具有吸引力。目前美国加州理工、法国 Inria 研究院和德国亚琛工大等计算机应用几何著名研究机构都在深入开展这方面的研究并取得了一定进展，我国浙江大学近几年在这方面也有突出的成果，具有国际水平。

对于基于体分解的六面体网格自动生成研究，国际上很重视，美国 Sandia 国家实验室、卡内基 - 梅隆等著名研究机构在这方面的研究工作不断深入，提出了基于中面分组的扫成特征识别方法、自动的 Hexahedral Mesh Matching 算法等。近年来，国内也开始重视这方面的研究，并提出了改进的 Hexahedral Mesh Matching 算法，但总体上这方面的研究成果还不多，相关研究有待进一步加强。

虽然四面体网格自动生成技术已经相对成熟，但直接生成的网格质量往往不能很好地满足实际需要。因此，四面体网格优化研究仍然受到国内外学术界的重视。比较而言，国际上更重视提高网格优化的效果，提出了一些新的优化准则和方法；国内则在如何有效提高网格优化效率尤其是拓扑优化效率方面做出了具有特色的工作，先后提出了两种新方法，一种是利用机器学习进行网格拓扑优化的方法，另一种是基于 GPU 加速的四面体网格并行优化算法，这两种方法都能显著提高四面体网格优化特别是拓扑优化的效率。

4.4 可视媒体

可视媒体是人类接受信息的主要来源，可视媒体的构建是人类从海量信息中高效获取和表达信息的重要手段。交互的过程就是人与计算机之间不断进行信息交换、理解、再交换的过程，自然和谐的交互可以提高可视媒体构建的效率，有效减轻用户的负担。为增强信息表现力，可视媒体的无缝合成是一个重要前提。因此，可视媒体的合成与交互是构建高质量可视媒体，进行信息高效获取与传达的关键问题。

交互性是未来数字媒体应用的重要特征，和谐自然的交互是可视媒体构建的必要手段。非精确交互（草图）用户界面及相关技术的研究是当前国际上很多著名的大学和研究机构的研究热点之一，如麻省理工学院（MIT）人工智能实验室、加州大学伯克利分校、华盛顿大学、东京大学以及微软研究院等。在国内，中科院软件所、北京大学、清华大学等多家科研机构也都在草图支撑技术和草图界面方面展开了积极的研究。

一致性的融合是决定可视媒体合成质量的关键。目前的研究工作主要有针对颜色域 alpha 融合和梯度域的图像融合方法。此类融合算法主要从图像空间像素颜色融合角度，融合过程中较多考虑的问题是边缘的无缝、自然的效果。重光照问题一直是可视媒体处理的重要研究内容，当前提出的重光照理论和方法与可视媒体融合的实际应用还有较大差距，传统的基于图像的重光照方法需要提供光照图像库，且人工交互复杂、难以实现，对于场景光照的条件表示和传播的研究较少，需要从新的视角提出基于可视媒体合成方法。

由于目前非精确交互模式的模糊性和二义性，利用网络海量的信息作为支撑提高非精确标注识别的准确性是一种可能的途径。清华大学提出的 Sketch2 photo 和 Sketch2

Scene 技术^[126,127]对此问题进行一次成功的探索，该技术充分利用了网络的海量可视媒体数据作为支持，根据用户绘制的草图交互进行有效的图像合成和三维场景构建。

5 发展趋势及展望

5.1 基于知识的建模/高层语义分析

3D 游戏、3D 电影等娱乐产业的飞速发展，以及智慧城市建设、大数据可视化等大规模战略需求，为我国图形学的发展带来了新的机遇和挑战。在三维内容生成方面，如何面向大规模应用提供智能化、高质量的三维可视内容是十分重要的课题，而知识驱动建模必将在其中担任重要角色。

首先，为适应数字化城市、数字化文物典藏、安防生产监控等需求，需要进一步研究面向激光扫描的大规模点云三维建模，以及结合领域先验知识的专用、高精度三维重建技术。例如，针对数字化城市和建筑物，需研究建筑物三维模型的结构分析，不同风格建筑结构的分类与识别，不同风格建筑物结构知识库的构建，以及城市级点云的增强、重建和浏览。

其次，需要在面向 3D 游戏和电影产业的高效三维建模技术方面有所突破。对应于游戏和电影生产过程中单个角色建模、批量角色建模和场景建模三个层次，需研究面向精确、高质量要求的智能化交互式建模；面向大规模批量建模的数据驱动模型自动合成；以及大规模场景自动合成。知识驱动建模在这三个方面都有非常重要的作用。

在我国，3D 打印已经受到工业界和学术界的广泛关注。构建可打印的三维模型是 3D 打印的关键步骤。“可打印性”包括对三维模型结构稳定性、材料使用、空间占用等方面的要求。面向 3D 打印的几何建模正成为图形学领域所关注的热点问题。如何结合 3D 打印专业领域知识，研究智能化的可打印三维模型构建，具有很好的应用前景。

数据是感知语义的核心。需要进一步推动三维模型库，特别是有语义信息的三维模型库的构建。研究如何结合有标注三维模型，以及互联网上存在的大量知识库（如亚马逊、淘宝等提供的庞大的物体分类知识），借助机器学习方法，为更多无标注模型提供分割、标注、分类等语义信息，并挖掘结构与语义的关联，以更好地组织、维护这些三维模型库，并利用模型库为高层次、智能化的应用提供语义支持。

在高层语义分析方面，一个重要发展趋势是三维模型的功能分析。针对模型集合的一致分割是迈向功能分析的重要途径之一。当前，一致分割主要依靠监督或无监督学习方法寻找三维物体几何属性与部件语义的对应关系。在模型几何外形和拓扑结构差异较大时，则需要借助用户输入。因此，如何将机器学习中的特征学习机制引入三维模型结构分析，以挖掘更深层次的语义相关结构特征，是值得研究的问题。高层语义分析的另外一个发展趋势是分析对象的复杂化，如由单个物体向复杂场景过渡，由单个模型向大

规模、异构模型集合发展。

5.2 数字城市建模

目前数字城市的三维空间数据基础建设中，传统的遥感、测绘方法仍是主要的数据获取方式，构建的主要信息交流平台仍然是相对独立的地理信息应用，例如在平台上显示遥感影像、数字地图和简单的社会信息、提供简单的交互查询方式等。大部分应用是针对某个单独领域的，而没有整合为涉及城市方方面面应用的巨系统，交互方式也以机器为主体。未来数字城市的发展方向应该是以生活在城市中的人为中心，是融合建筑、交通、电气、通信等实体和基础网络设施以及植物、气象、环境等自然要素的空间数据虚拟现实世界，也是社会经济等非空间信息的载体；数字城市基于空间、非空间信息而建立，同时在此框架下能够跨越时间空间进行复杂的自然和社会现象的模拟。比如美国 Linden 实验室开发的第二人生（second life）游戏就可以由玩家构建一个全三维的虚拟世界，在这个世界中玩家可以制造物品、买卖交易、旅行休闲等，可以选择不同的交通工具去任何想去的地方，去商场、夜总会、体育馆、教堂、图书馆、赌场等，玩家在这里可以换一种身份、性别和角色展开另外一种逼真的生活，经历不同的人生，体验不同的感受。如果将这里的三维场景换成基于真实世界的数字城市，建立起现实世界的一种数字化再现，人们就可以在这个数字化的世界中体验真实的人生，我们称之为“第一人生”。第一人生表现的是一个高度模仿人类现实世界的虚拟世界，它复制了现实世界的场景及其具有的各种信息属性，可以看做是现实世界的虚拟镜像，很多原来需要在现实世界中完成的工作都可以转化到这个虚拟世界中来实现，成为社会、经济、政治、军事等活动的试验场。比如传统的旅游、教育、电子商务、游戏娱乐、军事演习等活动都可以在这个虚拟世界中完成。随着它的进一步发展，它与现实世界的分界越来越模糊，两个世界就融合在一起，成为人类生存的一个新空间，实现人类生活的飞跃。

5.3 体网格生成与编辑

近年来，体网格生成与编辑研究虽然取得了较大进步，但现有方法在质量、效率、可靠性、智能性、自动化等方面都还存在许多问题，需要进一步研究解决。体网格生成与编辑研究的发展趋势包括以下方面：

（1）几何容错的体网格生成

目前的体网格生成方法都对几何模型的几何拓扑有效性具有严格的要求，由于目前的 CAD 建模及其几何模型交换技术无法完全满足其要求，因此十分耗时的人工几何清理不可避免，这严重影响了体网格生成的效率和自动化程度。为了能够从根本上解决这一难题，研究能够直接对带有少量几何拓扑瑕疵的几何模型进行体网格生成的新方法势在必行。

（2）基于先进并行计算技术的并行体网格生成和优化

随着产品模型越来越复杂，分析精度要求越来越高，工业界对体网格生成与优化的

效率提出了更高的要求。为此，有必要研究基于 GPU、云计算等先进并行计算技术的并行体网格生成和优化方法，以大幅度提高体网格生成与优化的效率。

(3) 基于领域知识的智能化体网格生成

现有的体网格生成方法主要基于几何生成体网格，所生成的体网格常常难以很好地满足实际需要，因为不同的领域和不同的问题通常需要不同特质的体网格，如各向同性或各向异性的体网格。为此，分析人员往往需要基于领域知识对体网格生成过程进行人工干预，以生成能够满足实际需要的体网格。为了能够取消体网格生成过程的人工干预，研究基于领域知识的智能化体网格生成方法十分必要。

(4) 体网格生成与 CAD 的无缝集成

目前的 CAD 建模与体网格生成一般独立进行。设计修改后，通常需要首先修改 CAD 模型，然后对变动 CAD 模型进行完整的网格划分，生成新的体网格模型，过程很漫长。为了有效提高产品设计分析效率，研究以 CAD/CAE 集成模型为基础，能够有效支持 CAD 与体网格编辑关联互动的体网格生成与 CAD 无缝集成方法是大势所趋。

(5) 全六面体网格的自动生成与高质量编辑

六面体网格具有许多四面体网格无法比拟的优异特性和计算精度，甚至对有些工程计算问题，只能采用六面体网格进行有限元分析。然而由于六面体网格生成的巨大难度，虽历经几十年的研究，复杂三维实体的全六面体网格的自动生成问题至今未能真正解决。鉴于全六面体网格的不可替代性，研究全六面体网格的自动生成与高质量编辑方法必将是人们长期的任务。

5.4 可视媒体

可视媒体作为新兴交叉学科，对于未来工作的展望包括以网络可视媒体交互与合成为主线建立网络可视媒体内容构建和呈现的研究体系。具体反映在借鉴人类视觉信息自动抽取机制和自适应记忆系统工作机理，以可视媒体认知计算模型、网络可视媒体语义分析与整合、网络可视媒体有效搜索与服务为主线建立可视媒体知识获取与智能服务的研究体系，并通过可视媒体的内容认证、真伪鉴别等研究为网络可视媒体提供内容安全保障，形成满足数字内容产业和网络可视媒体搜索及服务等国家重大需求的完整的有特色的研究体系。

6 结束语

早在计算机图形学发展初期，几何建模已经是该领域的一个核心问题。有趣的是，在图形学取得惊人成就的今天，几何建模仍是图形学最重要也是最具挑战性的问题之一。

从长达几十年的发展来看，几何建模的变革往往首先从三维模型的表示开始，随之带动相应的处理、编辑等方面的发展。例如，20世纪90年代提出的细分曲面表示，一举

改变了参数曲面表示统治几何建模数十年的局面，对工业界产生了巨大影响。随后，激光扫描和数字几何的兴起带动了包括重建、处理、编辑等内容的数字几何处理的发展，并在参数化表示下形成了统一的处理框架^[128]。如今，随着输入设备和交互技术的飞速发展，三维模型的建立和获取手段越来越丰富。对于绝大多数建模方式而言，以知识为驱动、以数据为中心已经成为建模智能化、规模化的重要支持。而如何有效地运用领域知识，并从数据中感知特定的语义信息，形成语义增强的三维几何表示，是解决上述问题的关键。

知识驱动建模和高层语义分析共同构成了以感知语义为中心的几何建模与处理新方向，众多技术上的新挑战、新趋势必将引出诸多新的科学问题，为图形学发展带来新的机遇，也为学科交叉融合带来新的方向，进而为整个图形学学科发展注入新的活力。

建立城市场景的三维模型是数字城市建设的重要内容和必要步骤。调查研究显示：越是经济发达的城市，三维城市模型应用的规模越大。建立城市场景的三维模型在城市规划、三维地图、智能交通、城市监控、应急指挥等领域发挥着重要作用，而且随着城市信息化程度的提高，城市场景的三维模型能够得到更加宽广的应用。

工业界对工程和科学计算的规模、复杂度、效率、精度等方面的要求越来越高是促进有限元网格生成技术不断发展的强大动力。由于现有的体网格生成方法，特别是六面体网格的生成和编辑，还难以满足工业界的实际需要，其中一些难题还没有得到真正意义上的解决，因此虽然体网格生成和编辑已经经历了几十年的发展，仍是值得大力开展研究的领域。

网络海量可视媒体智能处理是面向国家重大需求和国际学科前沿的重要科学问题，也是信息网络产业创新的重大瓶颈。本文对国内相关单位在这个方向上的一些代表性工作进行了综述，希望能够为国内从事这个领域研究的学者提供一些有益的参考。

致谢

衷心感谢参与本报告撰写、编辑与审定的人员！

参与本研究进展编写的人员包括（按拼音排序）：陈宝权教授（山东大学，数字化城市建模进展）、高曙光教授（浙江大学，体网格生成与编辑）、刘永进副教授（清华大学，可视媒体）、徐凯讲师（国防科技大学，基于知识的建模/高层语义分析）。

浙江大学鲍虎军教授、冯结青教授编辑并审定了全文。陈雪博士参与了本文的排版与编辑工作。

参考文献

[1] Kaijiya J. Talk by Steven Anson Coons Award Recipient at SIGGRAPH 2011[R] , 2011.

- [2] Müller P, Wonka P, Haegler S, Ulmer A, van Gool L. Procedural Modeling of Buildings [J]. ACM Trans. Graph. , 2006, 25(3) : 614 ~ 623.
- [3] Gal R, Sorkine O, Mitra N J, Cohen-Or D. iWIRES: An Analyze-and-Edit Approach to Shape Manipulation [J]. ACM Trans. on Graph. , 2009, 28(3) : 33.
- [4] Anguelov D, Srinivasan P, Koller D, Thrun S, Rodgers J, Davis J. SCAPE: Shape Completion and Animation of People. ACM Trans. Graph. , 24, 3: 408-416.
- [5] Xu K. Semantics Driven 3D Shape Analysis and Modeling [D]. Ph. D Thesis, National University of Defense Techonlogy, 2011.
- [6] Kalogerakis E, Hertzmann A, Singh K. Learning 3D Mesh Segmentation and Labeling [J]. ACM Trans. on Graph. , 2010, 29(3) : 1 ~ 12.
- [7] Wang Y, Asafi S, van Kaick O, Zhang H, Cohen-Or D, Chen B. Active Co-Analysis of a Set of Shapes [J]. ACM Trans. Graph. , 31(6).
- [8] Xu K, Li H, Zhang H, Cohen-Or D, Xiong Y, Cheng Z. - Q. Style- Content Separation by Anisotropic Part Scales. ACM Trans. Graph. , 29(6) : Article No. 184.
- [9] Sidi O, van Kaick O, Kleiman Y, Zhang H, Cohen-Or D. Unsupervised Co- Segmentation of a Set of Shapes via Descriptor-Space Spectral Clustering. ACM Trans. Graph. , 2011, 30(6) , 126: 1-10.
- [10] Ovsjanikov M, Li W, Guibas L, Mitra N J. 2011. Exploration of Continuous Variability in Collections of 3D Shapes. ACM Trans. Graph. , 30, 4 33: 1-33; 10.
- [11] Kim V G, Li W, Mitra N J, Chaudhuri S, DiVerdi S, Funkhouser T. Learning Part- based Templates from Large Collections of 3D Shapes [J]. ACM Trans. on Graph, 32(4).
- [12] 徐科, 李朝义, 等. 神经生物学纲要 [M]. 北京: 科学出版社, 2005.
- [13] Mitra N J, Wand M, Zhang H, Cohen-Or D, Bokeloh M. Structure- Aware Shape Processing [C]. Eurographics State-of-the-Art Report(STAR) , 2013.
- [14] Talton J O, Yu L, Lesser S, Duke J, Mech R, Koltun V. Metropolis Procedural Modeling [J]. ACM Trans. Graph. , 30, 2, Article 11.
- [15] Zheng Y, Fu H, Cohen-Or D, Au O K-C, Tai C-L. Component-wise Controllers for Structure-preserving Shape Manipulation [J]. Computer Graphics Forum, 2011, 30(2) : 563-572.
- [16] Funkhouser T, Kazhdan M, Shilane P, Min P, Kiefer W, Tal A, Rusinkiewicz S, Dobkin D. Modeling by Example. ACM Trans. Graph. , 2004, 23(3) : 652-663.
- [17] Chaudhuri S, Koltun V. Data-Driven Suggestions for Creativity Support in 3D Modeling [J]. ACM Trans. on Graph. , 2010, 29(6) : 1 ~ 9.
- [18] Chaudhuri S, Kalogerakis E, Guibas L, Koltun V. Probabilistic Reasoning for Assembly- Based 3D Modeling [J]. ACM Trans. Graph. , 2011, 30(4) : 35.
- [19] Kalogerakis E, Chaudhuri S, Koller D, Koltun V. A Probabilistic Model for Component-Based Shape Synthesis. ACM Trans. Graph. , 2012, 31(4) : 55.
- [20] Fisher M, Hanrahan P. Context-Based Search for 3D Models [J]. ACM Trans. on Graph. , 2010, 29(5) : 1 ~ 10.
- [21] Fisher M, Savva M, Hanrahan P. Characterizing Structural Relationships in Scenes Using Graph Kernels [J]. ACM Trans. Graph. , 2011, 30(4) , 34: 1-11.
- [22] Fisher M, Ritchie D, Savva M, Funkhouser T, Hanrahan P. Example- based Synthesis of 3D Object Arrangements [J]. ACM Trans. Graph. , 2012, 31(6) : 135.

- [23] Kim Y M, Mitra N J, Yan D M, Guibas L. Acquiring 3D Indoor Environments with Variability and Repetition[J]. ACM Trans. Graph. , 2012, 31(6).
- [24] Nan L, Xie K, Sharf A. A Search-Classify Approach for Cluttered Indoor Scene Understanding[J]. ACM Trans. Graph. , 2012, 31(6).
- [25] Shao T, Xu W, Zhou K, Wang J, Li D, Guo B. An Interactive Approach to Semantic Modeling of Indoor Scenes with an Rgbd Camera[J]. ACM Trans. Graph. , 2012, 31(6).
- [26] Xu K, Chen K, Fu H, Shun W-L, Hu S-M. Sketch2 Scene: Sketch-based Co-retrieval and Co-placement of 3D Models[J]. ACM Trans. Graph. , 2013, 32(4).
- [27] Stava O, Benes B, Mech R, Aliaga D G, Kristof P. Inverse Procedural Modeling by Automatic Generation of L-systems [J]. Computer Graphics Forum (Special Issue of Eurographics 2010) , 2010, 29 (2) : 665 ~ 674.
- [28] Bokeloh M, Wand M, Seidel H P. A Connection Between Partial Symmetry and Inverse Procedural Modeling[J]. ACM Trans. Graph. , 2010, 29(4) : 1 ~ 10.
- [29] Golovinskiy A, Funkhouser T. Consistent Segmentation of 3D Models[J]. Computers & Graphics, 2009, 33(3) : 262 ~ 269.
- [30] Huang Q, Koltun V, Guibas L. Joint Shape Segmentation with Linear Programming[J]. ACM Trans. on Graph. , 30(6).
- [31] Hu R, Fan L, Liu L. Co-segmentation of 3D Shapes via Subspace Clustering[J]. Computer Graphics Forum, 31(5) : 1703-1713.
- [32] Lv J, Chen X, Huang J, Bao H. Semi-supervised Mesh Segmentation and Labeling[J]. Computer Graphics Forum, 31(7) : 2241-2248.
- [33] Wu Z, Wang Y, Shou R, Chen B, Liu X. Unsupervised Co-Segmentation of 3D Shapes via Affinity Aggregation Spectral Clustering[J]. Computers & Graphics(SMI 2013) , to appear.
- [34] Nguyen A, Ben-Chen M, Welniak K, Ye Y, Guibas L. An Optimization Approach to Improving Collections of Shape Maps. Computer Graphics Forum, 30(5).
- [35] Kim V G, Li W, Mitra N, Diverdi S, Funkhouser T. 2012. Exploring Collections of 3D Models Using Fuzzy Correspondences[J]. ACM Trans. on Graph. , 31(4).
- [36] P Müller, G Zeng, P Wonka, L Gool. Image-based Procedural Modeling of Facades[C], Proceedings of SIGGRAPH, 2007.
- [37] D Hoiem, A Efros, M Hebert. Automatic Photo Pop-up [J]. ACM Transactions on Graphics, 2005: 577-584.
- [38] A Saxena, M Sun, A Ng. Make3D: Learning 3D Scene Structure from a Single Still Image [J]. Transactions on Pattern Analysis and Machine, 2009: 824-840.
- [39] O Barinova, V Konushin, A Yakubenko, K Lee, H Lim, A Konushin. Fast Automatic Single-View 3-d Reconstruction of Urban Scenes[C], Proceedings of European Conference of Computer Vision, 2008: 100-113.
- [40] A Hengel, A Dick, T Thormählen, B Ward, P Torr. VideoTrace: Rapid Interactive Scene Modelling from Video[C], Proceedings of SIGGRAPH, 2007.
- [41] J Xiao, T Fang, P Tan, P Zhao, E Ofek, L Quan. Image-based Facade Modeling[J]. ACM Transaction on Graphics(TOG) , vol. 27, no. 5, Proceedings of ACM SIGGRAPH, 2008, 161: 1-161: 10.
- [42] S Sinha, D Steedly, R Szeliski, M Agrawala, M Pollefeys. Interactive 3D Architectural Modeling from

- Unordered Photo Collections[J]. ACM Transactions on Graphics, 2008, 27(5): 159: 1-159: 10.
- [43] J Xiao, T Fang, P Zhao, M Lhuillier, L Quan. Image-based Street-side City Modeling[J]. ACM Transaction on Graphics(TOG), vol. 28, no. 5, Proceedings of ACM SIGGRAPH Asia 2009, 114: 1-114: 12.
- [44] M Pollefeys, D Nistér, J Frahm, A Akbarzadeh, P Mordohai, B Clipp, C Engels, D Gallup, S J Kim, P Merrell, C Salmi, S N Sinha, B Talton, L Wang, Q Yang, H Stewénius, R Yang, G Welch, H Towles. Detailed Real-Time Urban 3D Reconstruction from Video[J]. International Journal of Computer Vision, 2008, 143-167.
- [45] N Cornelis, B Leibe, K Cornelis, L J V Gool. 3D Urban Scene Modeling Integrating Recognition and Reconstruction[J]. International Journal of Computer Vision, 2008, 121-141.
- [46] D Anguelov, C Dulong, D Filip, C Frueh, S Lafon, R Lyon, A S Ogale, L Vincent, J Weaver. Google Street View: Capturing the World at Street Level[J]. IEEE Computer, 2010, 32-38.
- [47] S Agarwal, Y Furukawa, N Snavely, I Simon, B Curless, S M Seitz, R Szeliski. Building Rome in a Day [J]. Communications of the ACM 54, 10(Oct. 2011), 105-112.
- [48] P Müller, P Wonka, S Haegler, A Ulmer, L Gool. Procedural Modeling of Buildings[J]. ACM Transactions on Graphics(Proceedings of ACM SIGGRAPH 2006), 2006, 614-623.
- [49] J Sun, X Yu, G Baciu, M Green. Template-based Generation of Road Networks for Virtual City Modeling [C]. Proceedings of the ACM Symposium on Virtual Reality Software and Technology, 2002, 33-40.
- [50] G Chen, G Esch, P Wonka, P Müller, E Zhang. Interactive Procedural Street Modeling[C], Proceedings of ACM SIGGRAPH 2008, 2008.
- [51] M Pauly, N Mitra, J Wallner, H Pottmann, L Guibas. Discovering Structural Regularity in 3D Geometry [J]. ACM Transactions on Graphics, 2008, 27(3): 1-11.
- [52] R Gal, A Shamir, T Hassner, M Pauly, D Cohen-Or. Surface Reconstruction using Local Shape Priors [C], Proceedings of the fifth Eurographics Symposium on Geometry Processing, 2007, 253-262.
- [53] Y Livny, F Yan, M Olson, B Chen, H Zhang, J El-Sana. Automatic Reconstruction of Tree Skeletal Structures from Point Clouds[J]. ACM Trans. Graph, 29, 6, Article 151, 2010
- [54] Yotam Livny, Seoren Pirk, Zhanglin Cheng, Feilong Yan, Oliver Deussen, Daniel Cohen-Or, Baoquan Chen. Texture-Lobes for Tree Modelling[J]. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2011), 2011, 30(4).
- [55] Lu J H C, Song I, Quadros W R, Shimada K. (2013). Geometric Reasoning in Sketch-Based Volumetric Decomposition Framework for Hexahedral Meshing[C]. In Proceedings of the 21st International Meshing Roundtable(pp. 297-314). Springer Berlin Heidelberg.
- [56] Lu J H C, Song I, Quadros W R, Shimada K. (2012). Volumetric Decomposition via Medial Object and Pen-Based User Interface for Hexahedral Mesh Generation[C]. In Proceedings of the 20th International Meshing Roundtable(pp. 179-196). Springer Berlin Heidelberg.
- [57] Lu J H C, Song I, Quadros W R, Shimada K. (2010). Pen-based User Interface for Geometric Decomposition for Hexahedral Mesh Generation[C]. In Proceedings of the 19th International Meshing Roundtable (pp. 263-278). Springer Berlin Heidelberg.
- [58] Makem J E, Armstrong C G, Robinson T T. (2012). Automatic Decomposition and Efficient Semi-Structured Meshing of Complex Solids[C]. In Proceedings of the 20th International Meshing Roundtable(pp. 199-215). Springer Berlin Heidelberg.
- [59] Staten M L, Shepherd J F, Ledoux F, Shimada K. (2010). Hexahedral Mesh Matching: Converting non-

- Conforming Hexahedral-to-Hexahedral Interfaces into Conforming Interfaces[J]. International journal for numerical methods in engineering, 82(12), 1475-1509.
- [60] Staten M L. (2010). Sheet-based Generation and Modification of Unstructured Conforming All-Hexahedral Finite Element Meshes. Dissertation Abstracts International, 71(04).
- [61] Gregson J, Sheffer A, Zhang E. (2011, August). All-Hex Mesh Generation via Volumetric PolyCube Deformation[J]. In Computer graphics forum(Vol. 30, No. 5, pp. 1407-1416). Blackwell Publishing Ltd.
- [62] Ruiz-Gironés E, Roca X, Sarrate J. (2012). The Receding front Method Applied to Hexahedral Mesh Generation of Exterior Domains[J]. Engineering with Computers, 28(4), 391-408.
- [63] Wilson T J, Sarrate J, Roca X, Montenegro R, Escobar J M. (2012). Untangling and Smoothing of Quadrilateral and Hexahedral Meshes[C]. In Proceedings of the Eight International Conference on Engineering Computational Technology, Topping BHV (ed.), Civil-Comp Press, Stirlingshire, UK, Paper (Vol. 36).
- [64] Shepherd J F, Dewey M W, Woodbury A C, Benzley S E, Staten M L, Owen S J. (2010). Adaptive Mesh Coarsening for Quadrilateral and Hexahedral Meshes[J]. Finite Elements in Analysis and Design, 46(1), 17-32.
- [65] Yalcin E, Yilmaz A E, Kuzuoglu M. (2013). Performance Comparison of Various Hexahedral Element Quality Metrics via Parametric Distortion of an Ideal Element[J]. International Journal of Computational Methods, 10(04).
- [66] Remacle J F, Toulorge T, Lambrechts J. (2013). Robust Untangling of Curvilinear Meshes[C]. In Proceedings of the 21st International Meshing Roundtable(pp. 71-83). Springer Berlin Heidelberg.
- [67] Maehama H, Date H, Kanai S. Quality Improvement of Finite Element Mesh Models Modified by Mesh Deformation.
- [68] Ben-Chen M, O Weber C Gotsman. Variational Harmonic Maps for Space Deformation[C]. in ACM SIGGRAPH 2009 papers. 2009. New Orleans, Louisiana: ACM.
- [69] Smit M S, Bronsvoort W F. (2009). Efficient Tetrahedral Remeshing of Feature Models for Finite Element Analysis[J]. Engineering with computers, 25(4), 327-344.
- [70] Shontz S M, Vavasis S A. (2010). Analysis of and Workarounds for Element Reversal for a Finite Element-Based Algorithm for Warping Triangular and Tetrahedral Meshes[J]. BIT Numerical Mathematics, 50(4), 863-884.
- [71] Sieger D, Menzel S, Botsch, M. (2013). High Quality Mesh Morphing using Triharmonic Radial Basis Functions[C]. In Proceedings of the 21st International Meshing Roundtable(pp. 1-15). Springer Berlin Heidelberg.
- [72] Xu W, Wang J, Yin K, Zhou K, van de Panne M, Chen F, Guo B. Joint-Aware Manipulation of Deformable Models[J]. ACM Trans. Graph., 2009, 28(3): 35.
- [73] Chen X, Kang S B, Xu Y-Q, Dorsey J, Shum H Y Sketching Reality: Realistic Interpretation of Architectural Designs[J]. ACM Trans. Graph., 2008, 27(1).
- [74] Li G, Liu L, Zheng H, Mitra N. Analysis, Reconstruction and Manipulation using Arterial Snakes. ACM Trans. Graph. 29, 5, Article No. 152.
- [75] Li Y, Wu X, Chrysathou Y, Sharf A, Cohen-Or D, Mitra N J. Globfit: Consistently Fitting Primitives by Discovering Global Relations[J]. ACM Trans. Graph., 30(4): 52.
- [76] Lin H, Gao J, Zhou Y, Lu G, Ye M, Zhang C, Liu L, Yang R. Semantic Decomposition and Recon-

- struction of Residential Scenes from LiDAR Data [J]. ACM Trans. Graph. , 32(4).
- [77] Nan L, Sharf A, Zhang H, Cohen-Or D, Chen B. Smartboxes for Interactive Urban Reconstruction [J]. ACM Trans. Graph. , 29, 4, 1-10.
- [78] Zheng Q, Sharf A, Wan G, Li Y, Mitra N J, Cohen-Or D, Chen B. Non-local Scan Consolidation for 3d Urban Scenes [J]. ACM Trans. Graph. , 29, 4, 1-9.
- [79] Livny Y, Pirk S, Cheng Z, Yan F, Deussen O, Cohen-Or D, Chen B. Texture-Lobes for Tree Modelling [J]. ACM Trans. Graph. , 30(4).
- [80] Shen C-H, Fu H, Chen K, Hu S-M. Structure Recovery by Part Assembly [J]. ACM Trans. Graph. , 31, 6, 184; 1-184; 10.
- [81] Zhang H, Xu K, Jiang W, Lin J, Cohen-Or D, Chen B. Layered Analysis of Irregular Facades via Symmetry Maximization [J]. ACM Trans. Graph. , 2013, 32(4).
- [82] Xu K, Zhang H, Cohen-Or D, Chen B. Fit and Diverse: Set Evolution for Inspiring 3D Shape Galleries. ACM Trans. Graph. , 31(4); 57; 1-57; 10.
- [83] Xu K, Zhang H, Tagliasacchi A, Liu L, Li G, Meng M, Xiong Y. Partial Intrinsic Reflectional Symmetry of 3D Shapes. ACM Trans. Graph. , 28(5).
- [84] Xu K, Zhang H, Jiang W, Dey R, Cheng Z, Liu L, Chen B. Multi-Scale Partial Intrinsic Symmetry Detection. ACM Trans. Graph. , 31(6).
- [85] Wang Y, Xu K, Li J, Zhang H, Shamir A, Liu L, Cheng Z, Xiong Y. Symmetry Hierarchy of Man-Made Objects [J]. Computer Graphics Forum, 2011, 30(2); 287-296.
- [86] Wang Y, Shmulik A, Oliver V K, Zhang H, Cohen-Or D, Chen B. Active Co-Analysis of a Set of Shapes [J]. ACM Trans. Graph. , 2012, 31(6); 165-175.
- [87] van Kaick O, Xu K, Zhang H, Wang Y, Sun S, Shamir A, Cohen-Or D. Co-Hierarchical Analysis of Shape Structures [J]. ACM Trans. Graph. , 2013, 32(4).
- [88] Y Liu, H Hel-Or, C Kaplan, L Gool. Computational Symmetry in Computer Vision and Computer Graphics [J]. Foundations and Trends in Computer Graphics and Vision, 2010, pp. 1-195.
- [89] Q Zheng, A Sharf, G Wan, Y Li, N Mitra, B Chen, D Cohen-Or. Non-local Scan Consolidation for 3D Urban Scene [J]. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2010) , 2010.
- [90] L Nan, A Sharf, H Zhang, D Cohen-Or, B Chen. SmartBoxes for Interactive Urban Reconstruction [J]. ACM Transactions on Graphics (Proceedings of SIGGRAPH 2010) , 2010.
- [91] H Xu, N Gossett, B Chen. Knowledge and Heuristic-Based Modeling of Laser-Scanned trees [J]. ACM Trans. Graph. 26, 4, Article 19.
- [92] Huang J, Tong Y, Wei H, Bao H. Boundary Aligned Smooth 3D Cross-Frame Field [J]. In ACM Transactions on Graphics (TOG) ACM. 2011, 30(6); 143.
- [93] Huang J, Tong Y, Zhou K, Bao H, Desbrun M. Interactive Shape Interpolation Through Controllable Dynamic Deformation [J]. Visualization and Computer Graphics, IEEE Transactions on, 2011, 17 (7), 983-992.
- [94] Li Y, Liu Y, Xu W, Wang W, Guo B. All-hex Meshing Using Singularity-Restricted Field [J]. ACM Transactions on Graphics (TOG) , 2012, 31(6) , 177.
- [95] Li M, Tong R. All-Hexahedral Mesh Generation via Inside-out Advancing Front Based on Harmonic Fields [J]. The Visual Computer, 2012, 28(6-8) , 839-847.
- [96] Chen J, Zhu H, Gao S, Wu H. An Improved Hexahedral Mesh Matching Algorithm [C]. Submitted to 22nd International Meshing Roundtable, 2013.
- [97] Chen X, Peng D, Gao S. SVM-Based Topological Optimization of Tetrahedral Meshes [C]. In Proceedings

- of the 21st International Meshing Roundtable(pp. 211-224). Springer Berlin Heidelberg, 2013.
- [98] Peng D, Gao S, Si H, Zheng Q. GPU-Accelerated Tetrahedral Mesh Improvement[C]. Submitted to 22nd International Meshing Roundtable, 2013.
- [99] Sun L, Zhao G, Ma X. Quality Improvement Methods for Hexahedral Element Meshes Adaptively Generated using Grid-Based Algorithm[J]. International Journal for Numerical Methods in Engineering, 2012, 89 (6), 2012, 726-761.
- [100] Sun L, Zhao G, Ma X. Adaptive Generation and Local Refinement Methods of Three-Dimensional Hexahedral Element Mesh[C]. Finite Elements in Analysis and Design, 2012, 50, 184-200.
- [101] Song W, Liu L. Stretch-based Tetrahedral Mesh Manipulation[C]. In Proceedings of Graphics Interface 2007(pp. 319-325). ACM.
- [102] Liao S H, Tong R F, Dong J X, Zhu F D. Gradient Field based Inhomogeneous Volumetric Mesh Deformation for Maxillofacial Surgery Simulation[J]. Computers & Graphics, 2009, 33(3), 424-432.
- [103] Xian C, Gao S, Lin H, Liu Y, Xiao D. FEA-mesh Editing with Feature Constrained[C]. In Computer-Aided Design and Computer Graphics, 2009. CAD/Graphics '09. 11th IEEE International Conference on (pp. 340-343). IEEE.
- [104] Xian C, Gao S, Zhang T. Tetrahedral Mesh Editing with Local Feature Manipulations[C]. In Computer-Aided Design and Computer Graphics(CAD/Graphics) , 2011 12th International Conference on(pp. 130-137). IEEE.
- [105] Zhu H, Gao S, Li M, Pan W. Adaptive Tetrahedral Remeshing for Modified Solid Models[J]. Graphical Models, 2012, 74(4), 76-86.
- [106] Liu T, Chen W, Liu C H, Fu X. Benefits and Costs of Uniqueness in Multiple Object Tracking: the Role of Object Complexity[J]. Vision Research, 2012, 66, 31-38.
- [107] Zhao K, Wu Q, Shen X, Xuan Y M, Fu X. I Undervalue You but I Need You: The Dissociation of Attitude and Memory toward In-Group Members[J]. PLoS ONE, 2012, 7(3): e32932.
- [108] Zhuang X, Wu C. The Safety Margin and Perceived Safety of Pedestrians at Unmarked Roadway [J]. Transportation Research Part F: Traffic Psychology and Behaviour, 2012, 5(2): 119-131.
- [109] Yong-Jin Liu, Zhan-Qing Chen, Kai Tang. Construction of Iso-contours, Bisectors and Voronoi Diagrams on Triangulated Surfaces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI) , 2011, 33(8): 1502-1517.
- [110] Yong-Jin Liu. Exact Geodesic Metric in 2-manifold Triangle Meshes Using Edge-Based Data Structures [J]. Computer-Aided Design, 2013, 45(3): 695-704.
- [111] Xian-Ying Li, Shi-Min Hu. Poisson Coordinates[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(2): 344-352.
- [112] Xian-Ying Li, Tao Ju, Yan Gu, Shi-Min Hu. A Geometric Study of V-style Pop-ups: Theories and Algorithms[J]. ACM Transactions on Graphics, 2011, 30(4).
- [113] Chuntao Wang, Jiangqun Ni, Jiwu Huang. An Informed Watermarking Scheme Using Hidden Markov Model in the Wavelet Domain [J]. IEEE Trans. on Information Forensics & Security. 2012, 7 (3): 853-867.
- [114] Fangjun Huang, Jiwu Huang, Yun Q Shi. New Channel Selection Rule for JPEG Steganography. IEEE Trans. on Information Forensics & Security. 2012, 7(4): 1181-1191.
- [115] Ming-Ming Cheng, Guo-Xin Zhang, Niloy J. Mitra, Xiaolei Huang, Shi-Min Hu. Global Contrast based Salient Region Detection[C]. IEEE CVPR, 2011, 409-416.

- [116] Tao Chen, Ping Tan, Li-Qian Ma, Ming-Ming Cheng, Ariel Shamir, Shi-Min Hu. PoseShop: A Human Image Database and Personalized Content Synthesis[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(5): 824-837.
- [117] Tao Chen, Jun-Yan Zhu, Ariel Shamir, Shi-Min Hu. Motion-Aware Gradient Domain Video Composition [J]. IEEE Transactions on Image Processing, 2013, 22(7): 2532 - 2544.
- [118] Jin Wei, Chen-Feng Li, Shi-Min Hu, Ralph Martin, Chiew-Lan Tai. Fisheye Video Correction [J]. IEEE Transactions on Visualization and Computer Graphics, 2012, 18(10): 1771-1783.
- [119] Shao-Ping Lu, Song-Hai Zhang, Jin Wei, Shi-Min Hu, Ralph R Martin. Time-Line Editing of Objects in Video[J]. IEEE Transactions on Visualization and Computer Graphics, 2013, 19(7): 1218-1227.
- [120] Youyi Zheng, Xiang Chen, Ming-Ming Cheng, Kun Zhou, Shi-Min Hu, Niloy J Mitra. Interactive Images: Proxy-based Scene Understanding for Smart Manipulation[J]. ACM Transactions on Graphics(ACM SIGGRAPH). 2012, 31(4): 1-10.
- [121] Cui-Xia Ma, Yong-Jin Liu, Hong-An Wang, Dong-Xing Teng, Guo-Zhong Dai. Sketch-based Annotation and Visualization in Video Authoring [J]. IEEE Transactions on Multimedia, 2012, 14 (4): 1153-1165.
- [122] Zhi Wang, Lifeng Sun, Xiangwei Chen, Wenwu Zhu, Jiangchuan Liu, Minghua Chen, Shi-Qiang Yang. Propagation-Based Social-Aware Replication for Social Video Contents[C]. Proceedings of the 20th ACM international conference on Multimedia, 2012, 29-38.
- [123] Xie X, Xu K, Mitra N J, Cohen-Or D, Chen B. Sketch-to-Design: Context-based Part Assembly [J]. Computer Graphics Forum. To appear.
- [124] Ovsjanikov M, Li W, Guibas L, Mitra N J. Exploration of Continuous Variability in Collections of 3D Shapes[J]. ACM Trans. Graph., 2011, 30(4): 33-43.
- [125] Huang S, Shamir A, Shen C, Zhang H, Sheffer A, Hu S. Qualitative Organization of Collections of Shapes via Quartet Analysis[J]. ACM Trans. Graph. , 2013, 32(4): 71-81.
- [126] Tao Chen, Ming-Ming Cheng, Ping Tan, Ariel Shamir, Shi-Min Hu. Sketch2 Photo: Internet Image Montage[J]. ACM Transactions on Graphics, 2009, 28(5): 124.
- [127] Kun Xu, Kang Chen, Hongbo Fu, Wei-Lun Sun, Shi-Min Hu. Sketch2 Scene: Sketch-based Co-retrieval and Co-placement of 3D Models[J]. ACM Transactions on Graphics, Vol. 32, SIGGRAPH 2013.
- [128] Zhou K, Bao H, Shi J. A Unified Framework for Digital Geometry Processing [J]. Chinese Journal of Computers, 25(9).

作者简介

鲍虎军 博士，浙江大学 CAD&CG 国家重点实验室主任、教育部长江学者特聘教授、博士生导师。主要研究方向为计算机图形学、计算机视觉、虚拟现实。CCF 理事，中国计算机学会第七届计算机辅助设计与图形学专业委员会主任。



陈宝权 博士，山东大学计算机学院院长、教授、博士生导师、兼中科院深圳先进技术研究院可视计算研究中心首席科学家。主要研究方向为计算机图形与可视化。中国计算机学会常务理事及青工委主任。



冯结青 博士，浙江大学 CAD&CG 国家重点实验室教授、博士生导师。主要研究方向为计算机图形学。CCF 高级会员，中国计算机学会第七届计算机辅助设计与图形学专业委员会秘书长。



高曙明 博士，浙江大学 CAD&CG 国家重点实验室教授、博士生导师。主要研究方向为计算机辅助设计。CCF 高级会员，中国计算机学会计算机辅助设计与图形学专业委员会委员。



刘永进 博士，清华大学计算机科学系副教授。主要研究方向为计算几何、计算机图形学、计算机辅助设计和模式识别。CCF 高级会员，中国计算机学会计算机辅助设计与图形学专业委员会委员。



徐凯 博士，国防科技大学计算机学院讲师，中科院深圳先进技术研究院博士后研究员。主要研究方向为三维几何处理和几何建模。CCF 会员，中国计算机学会计算机辅助设计与图形学专业委员会委员。



嵌入式系统抗恶劣环境技术的研究进展

CCF 抗恶劣环境计算机专业委员会

摘要

随着信息技术的迅速发展，信息与物理融合系统在工业、医疗和军事领域得到广泛应用。然而这些信息化的嵌入式系统通常需工作在恶劣环境中，主要包括极端的物理环境、复杂的电磁环境和易受攻击的网络环境等。抗恶劣环境技术极大地影响着嵌入式系统的有效应用，成为当前嵌入式领域关注的热点。传统的抗恶劣环境技术主要考虑物理环境的恶劣影响，已不能满足复杂电磁环境和易受攻击网络环境下对嵌入式系统长时间、大范围、安全可靠运行的需求。在简要回顾嵌入式系统抗恶劣环境相关技术发展的基础上，本报告将从嵌入式系统的体系架构、系统管理、容错、网络通信、片上系统、安全软件、环境防护和健康监控等方面总结抗恶劣环境技术的发展现状，比较国内外研究的差距，并对未来的发展趋势提出预测和建议。

关键词：恶劣环境，嵌入式系统，容错技术，安全软件，加固技术

Abstract

With the rapid development of information technology, cyber-physical systems are widely applied in industry, medical, military field and etc., which always run in severe environments, such as extreme physical environment, complicated electromagnetic environment and vulnerable network environment. The technology of resistance to severe environment always affects the efficient of embedded systems, which has become the focused hotspot technology in the embedded field. The traditional method of resistance to harsh environment mainly pay close attention to harsh affection of physical environment, which has not meet the needs of embedded systems running with long time, large-scale, safe and reliable requirements in complicated electromagnetic environment and vulnerable network environment. This paper gives the development status, the gap between home and aboard, and the prospect for the future of the technology of resistance to severe environment from the aspects of architecture, system management, fault-tolerant, network communication, systems on chip, secure software, environment defense and premonition health monitoring.

Keywords: severe environment, embedded systems, fault-tolerant technology, secure software, reinforcement technology

1 引言

随着信息化产业的发展，对计算性能的要求越来越高。当涉及一些特殊的应用场合时，面对高温、低温、潮湿、霉菌、淋雨、沙尘以及强烈振动冲击和电磁干扰等恶劣环

境，嵌入式计算机仍需保持原有性能，进行高效运转。因此，对嵌入式系统抗恶劣环境计算机的研究就显得尤为重要。

嵌入式系统抗恶劣环境技术是指在恶劣环境能够抵抗外部恶劣环境影响情况下嵌入式计算机能够长时间、大范围、安全可靠运行，其中恶劣环境主要包括极端的物理环境、复杂的电磁环境和易受攻击的网络环境等，参见图 1。极端的物理环境包括嵌入式计算机工作于极端高低温、湿度、震动、腐蚀等不利于电子设备正常工作的环境，此时需要为嵌入式计算机提供保温、散热、抗冲击、防潮、防霉、抗腐蚀等技术措施；复杂的电磁环境是指嵌入式计算机所工作的超强、多频谱电磁辐射的复杂环境，此时需要为嵌入式计算机提供抗电磁干扰、抗闪电、抗辐照等技术措施；易受攻击的网络环境是指嵌入式计算机运行于开放式的网络环境，这种环境下存在各种网络的主动和被动攻击行为，例如病毒、后门、入侵、DDoS 攻击、通信截获和窃听等，此时需要为嵌入式计算机提供运行、通信、存储等方面的信息安全防护技术措施。

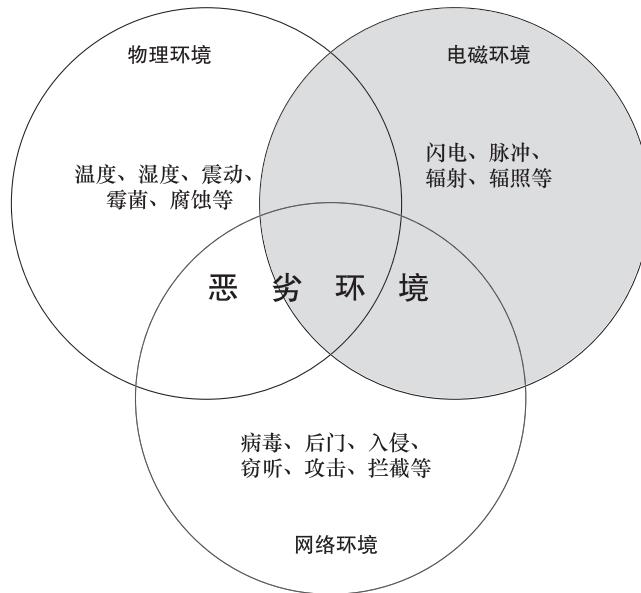


图 1 恶劣环境的定义

嵌入式计算机的抗恶劣环境技术主要是通过增强系统自身的抵抗能力，采用静态被动防御和动态主动预防两方面的技术措施，从系统架构、设计技术、防护技术等方面提高嵌入式计算机抵御恶劣环境的能力，保证在恶劣外部环境下仍然能够正常运行。所涉及的技术措施包括：体系结构、容错技术、可靠网络、片上系统、健康监控、安全软件和环境防护等，参见图 2。

由于嵌入式系统广泛应用于国防、工业、医疗等各个领域，并已快速渗透到物理世界和生活领域，因而其安全性和可靠性更加受到重视。如何有效抵抗物理环境、电磁环境和网络环境带来的影响，是嵌入式系统抗恶劣环境技术研究面临的严重挑战。本文将从国内外嵌入式系统抗恶劣环境技术的各个方面对技术发展现状和趋势进行讨论，力图

为我国嵌入式领域的抗恶劣环境技术发展提出有益的建议和展望。

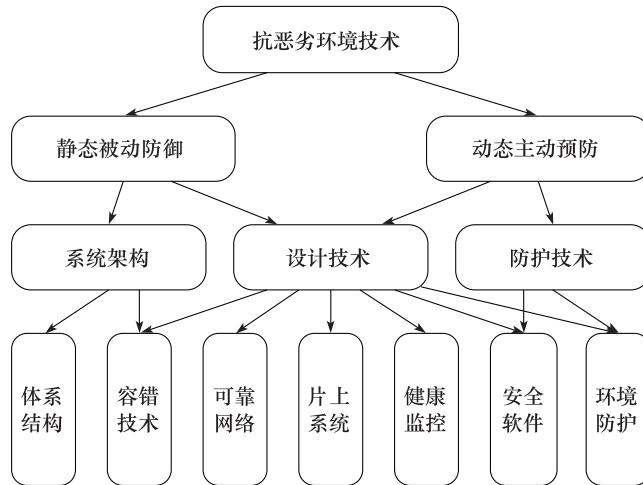


图 2 抗恶劣环境技术的内涵

2 国外研究现状

国外抗恶劣环境嵌入式计算机的设计与生产涉及系统架构、容错技术、可靠网络、片上系统、健康监控、安全软件及环境防护等多项技术。

2.1 体系架构

在抗恶劣环境技术领域，对计算机体系架构的研究重点是安全关键系统的体系架构。欧盟在 2004 年的框架 6 启动 DECOS (Dependable Embedded Components & Systems) 项目进行高可信嵌入式组件和系统研究^[1,2]，提出了高可信的组件化系统的设计架构和实现方法。其核心思想是采用时间触发架构来构建网络层的核心服务，上层应用支持多种安全等级的任务的综合部署，底层网络实现根据应用需求进行调整，但保持核心网络层高可靠、强实时、错误隔离与限制的核心服务，从而提供基于组件化的系统设计架构。2008 年，欧盟框架 7 的研究计划 GENESYS (GENeric Embedded SYStem) 展开了对嵌入式系统的参考架构的研究^[3]，提出了面向组件的架构思想，设计了跨领域的嵌入式系统的通用体系架构，以时间触发网络通信架构为核心，主要考虑了多种系统等级的安全性和健壮性的设计和保障。该体系架构将系统底层平台抽象成核心平台服务，通过连接接口 (Linking InterFace, LIF) 与局部服务组件交互，组合后构成全局应用服务。2008 年，美国海军联合 IBM 开展了开放式体系架构的研究^[4]，给出了一系列技术指导原则和技术实践指南。该研究报告中给出了 7 种非功能性需求对开放式架构的约束和依赖关系，并

对各种非功能需求给出了严格的定义，以满足安全关键领域计算机体系架构抗恶劣环境的需求。2012年，欧盟框架7的SCARLETT（SCAlable & Reconfigurable Electronics platforms and Tools）计划提出了下一代航空领域机载航电系统新架构，将重配置能力引入到计算机体系结构中^[5]。

国外在安全关键领域的计算机体系结构技术发展的主要目标是利用组件化设计思想，通过开放式、高安全的体系结构设计，在提高组件的安全性和可靠性的基础上构建系统的高安全性，满足在恶劣环境下计算机系统安全、可靠运行的需求。

2.2 容错技术

目前，国外对于抗恶劣环境嵌入式计算机容错的研究热点主要集中在分布式容错等技术上。从国外大部分先进技术验证机以及新一代机载系统研制来看，已经普遍采用模块化分布式计算机设计技术，将传感器处理、输入/输出控制、控制决策和控制计算等以外场可更换模块（LRM）的形式独立于飞行控制/飞行器管理计算机外。在大气数据处理单元、作动器控制单元、发动机控制计算机与机载机电设备管理子系统同中央控制计算机之间采用数据总线方式进行信息交换。在容错设计的架构发展方面，国外新提出的IVMS体系结构也呈现出向分布式发展的趋势^[6]。例如，EF-2000战斗机的VMS中部分子系统间采用数据总线进行信息交换；F-35战斗机的IVMS采用全总线化的分布式计算机系统结构。欧洲空客公司（AIR BUS）积极开展采用交换式通信网络的飞行控制系统研究。网络化的容错计算机系统具有结构分散、配置灵活、维修方便、布线简单等优点，它将成为新一代航空器IVMS的基础计算平台。

国外对于抗恶劣环境嵌入式计算机容错技术的研究，更多地集中在分布式容错体系架构，高完整性计算、细粒度故障诊断、隔离技术、故障动态重构与自愈等关键技术上。目标在于通过采用新的容错手段，提高资源的利用率，增强系统的扩展性、适应性，改善维修性，不断提升嵌入式容错计算机的费效比。

2.3 可靠网络技术

目前国外对于抗恶劣环境嵌入式计算机中可靠性网络的研究主要集中在对传统以太网的改进以及新一代基于时间同步的网络等技术研究上。主要成果有AFDX网络和TTE网络。

国际上对于AFDX网络的研究起步较早，空客公司最早提出AFDX网络的概念并已注册为商标，并由ARINC于2005年6月发布为ARINC664P7规范。目前，AFDX网络已经成功应用在空客A380、A400M、波音B787等机型上。在A380上，9套AFDX交换机互联驾驶舱、核心处理、飞行管理、发动机、燃油、液压、客舱等电子设备和分系统的端系统，提供了分布式、高可靠、延时确定的数据通信网络平台。而在波音787上，4套AFDX交换机互联核心处理、引擎&控制、导航、飞行控制、FMS、数据集中器等电子设

备和分系统的端系统，提供了分布式、高可靠、延时确定的数据通信网络平台。除航空领域，AFDX 网络在片上网络^[7]等技术领域也有应用。

目前，国外很多研究机构和电气公司从事着 TTE 的研发工作，主要包括维也纳大学、Malardalen 大学、维罗大学、杜伊斯堡 - 埃森大学等，其中维也纳大学的实时研究小组在 TTE 的研发方面具有一定权威性，正在进行标准的制定工作。从事 TTE 研究的公司主要有 TTTech Computertechnik AG、安森美半导体公司和 GE Fanuc 智能平台的航空电子产品公司等。TTE 技术发展经历了三个阶段，即 TTE 研究概念阶段、TTE 规范总结阶段和 TTE 标准 SAE AS6802 阶段^[8]，该标准已于 2011 年 11 月正式发布。TTE 通过在以太网 802.3 协议 MAC 层之上、LLC 层之下实施时间触发控制（Time-Triggered Control），对全网同步无竞争状态（Contention Free）的时间触发信息传送提供保障^[9]。

在可靠性网络的研究上，国外对于新一代时间同步网络的研究已经基本突破了实验室验证阶段，正在向产品实践验证阶段迈进。

2.4 片上系统技术

SoC 在国外已有几十年的发展历史。以 SoC 技术在航空航天方面的应用来看，目前 SoC 的主要发展方向是努力实现片上系统的高可靠性、低成本及高易用性。

美国 Aeroflex 公司通过应用 SoC 技术，为 NASA 开发了卫星遥测与控制系统^[10,11]。它将 4 个电路板和 1 个背板组成的多模块系统集成为一个芯片，重量由原有的 3200g 降至 34g，功耗由原有的 9 瓦减至 3 瓦，同时将组件之间可能的失效点由 12 处缩减至 1 处。而且，其具有总剂量 300K rad (Si)、LET > 100 MeV·cm²/mg、错误翻转率为 3.7E-8 个错误/天·器件的抗辐射指标，能够满足航天 30 年长寿命高可靠应用的要求。

ESA 公司成功推出了用于空间辐射环境的 LEON 系列微处理器^[12-14]、CAN 总线控制器、HDL 源代码等可供 SoC 设计复用的 IP 以及单片航天用控制器系统等；NASA 的火星探测器也采用了基于 RAD6000 的 SoC 来实现。

Microsemi 公司推出了新一代的 SmartFusion2 SoC FPGA^[15]，用于满足关键性工业、国防、航空、通信、医疗等应用对于先进安全性、高可靠性和低功耗的基本需求。SmartFusion2 在单一芯片上集成了固有可靠性的快闪 FPGA 架构、ARM Cortex-M3 处理器、安全处理加速器、DSP 模块、SRAM、eNVM 和业界所需的高性能通信接口。该产品已应用在飞行数据记录器、武器系统、除颤器、手持无线电设备、通信系统管理应用和工业马达控制等领域。

Altera 公司推出了第十代器件 Stratix 10 FPGA 和 SoC，它以最低功耗实现了业界最好的性能和水平最高的系统集成度。第十代器件采用目前最先进的工艺技术——Intel 14-nmTri-Gate 工艺和 TSMC 20-nm 工艺，工作频率超过 1GHz，是高端 28-nm 内核性能的 2 倍，且功耗降低 70%。

在国际航空航天应用领域，SoC 已经日益成为实现嵌入式计算机微型化、易用化的关键技术，而随着 SoC 的高可靠性的逐渐完善，其必将具有更广泛的应用前景。

2.5 健康监控技术

近年来，有关健康监控（Premonition Health Management，PHM）技术的学术研究和应用研究非常活跃。很多国际知名企也开展了 PHM 理论、技术、软件或者应用解决方案等方面研究，如 Impact Technologies、Honeywell 等。美国的马里兰大学、佐治亚理工学院、田纳西大学、麻省理工学院、加州工学院、斯坦福大学等相关学术机构都开展了各具特色的 PHM 技术研究工作。马里兰大学所属的先进生命周期工程中心成立了故障预测与健康管理联合会，深入开展了电子 PHM 技术方面的研究，并为多家知名企业、研究院所以及各兵种提供培训与技术解决方案；NASA 举办了首届国际宇航“综合系统健康工程和管理”（ISHEM）论坛，将其作为一门新的学科推出，同时开展了多方面 PHM 技术的研究和应用。IEEE Reliability Society 和 IEEE Aerospace and Electronic Systems Society 以及其他 PHM 协会、组织每年也会定期举行国际会议、技术讨论或专项技术交流活动^[20]，积极探索和推动 IVHM 技术发展^[21]。

在军事领域，美国为 F-35 JSF 开发的 IVHM 系统是最早也是目前技术水平最高的应用，F-35 的 IVHM 系统代表了目前 CBM（Condition Based Maintenance）应用的最高水准^[16]。同时以 IVHM 技术为基础的集成应用平台也广泛应用于英、美、加拿大等国研制的各类飞机系统中，该集成应用平台被称作“健康与使用监控系统（HUMS）”。美国各军种及其他机构也开展了与 IVHM/HUMS 类似的技术研究项目，如美国空军研究实验室提出的综合系统健康管理（ISHM）系统方案；海军的综合状态评估系统（ICAS）和预测增强诊断系统（PEDS）项目；陆军的诊断改进计划（ADIP）、嵌入式诊断和预测同步（EDAPS）计划等。

在民用技术领域，IVHM 在民用飞机、汽车、复杂建筑、桥梁、核电站、大型水坝等重要装备和工程设施的监控和健康管理中得到广泛应用。其中 IVHM 技术在民用航空领域的应用尤其突出。比如，波音公司的民机 PHM 解决方案——“飞机状态管理”（AHM）系统已在多家航空公司的多种客运或货运客机上大量应用^[17,18]。由美国 ARINC 公司与 NASA 兰利研究中心共同开发的与 IVHM 类似的“飞机状态分析与管理系统（ACAMS）”也可应用于飞机领域。在航天应用方面，NASA 的第二代可重用运载器已经采用了航天器综合健康管理方案，并选定 QSI 公司开发的综合系统健康管理方案对航天飞机进行健康监控、诊断推理和最优查故，以降低危及航天任务安全的系统故障^[19]；NASA 已拟定了未来 10 年的飞机综合健康管理技术计划^[20]，作为其航空安全项目的一个重要组成部分。

国外在健康监控领域技术方面研究的主要趋势是通过研究健康管理、状态诊断和评估等关键技术，进一步提高航空航天等应用领域计算机的抗恶劣环境能力。

2.6 安全软件技术

目前，抗恶劣环境嵌入式计算机对于安全软件的需求主要集中在具备分区、高安全、

微内核的嵌入式操作系统上。

分区操作系统这一概念源自安全关键应用领域，它的目的是为了实现多种类型的应
用安全的共享同一资源。随后又出现了多重独立安全等级（Multiple Independent Levels of
Security, MILS）^[21]和航空电子应用件标准接口（Avionics Application Software Standard In-
terface, ARINC）^[22]，它们在实时和安全关键等领域被广泛采用。与各种联合式或者分布
式的实时系统实现不同，分区操作系统利用了多种 COTS 的处理器模块和软件组件来构
建综合性系统，允许嵌入式应用程序被合并到整体性的系统当中。目前，已有一些具有
代表性的分区操作系统或相似概念实现。

国外操作系统厂商已推出面向机载领域的高安全嵌入式操作系统。例如，Green Hills
Software 公司首先推出了满足 ARINC653 APEX 接口的 INTEGRITY；Wind River 公司推出
了满足 ARINC653 APEX 接口的 VxWorks653；LynuxWorks 公司推出了 LynxOS-178；BAE
公司推出了 CsLEOS，以及 Honeywell 推出了 DEOS。这些操作系统已经随机载设备通过了
DO-178B 认证，被应用到各类机型中。

INTEGRITY 专门针对机载领域研发，已用于几乎所有主要的下一代商用和军用飞
机，包括 B787、A380、F-35、F/A-22、A400M、X-45C 无人驾驶战斗机（UCAV）、S-92
直升机，也用于 B-52、B-1B、C-17 和 F-16 的改进型。INTEGRITY 以其优异的强实时和
安全特性，在飞行控制系统中得到大量应用。VxWorks653 面向综合化航电系统研制，广
泛应用于 B787、A380 等新型主流飞机的通用核心处理机中。LynxOS-178 主要在显示系
统中得到广泛应用，CsLEOS 主要用于 BAE 的机载设备，DEOS 主要用于 Honeywell 的飞
机发动机控制器中。

为了支持多安全级别的混合处理，对信息安全提供有效支撑，分区操作系统应
满足多重独立安全级别（Multiple Independent Levels of Security, MILS）架构，采用微核
构建分区环境。MILS 架构操作系统的根本设计思想是将操作系统进行层次划分，内核层
仅仅包含提供分区隔离机制的最小功能集合；大多数传统的操作系统功能（如设备驱动、
文件系统、分区安全通信软件等）则以中间件的形式存在，分别驻留于不同的分区，为
应用层软件提供服务。微核方式的分区技术可以保证核心软件足够小，能够提供形式化
验证的工程条件，从而实现安全关键系统所需的高安全隔离特性。INTEGRITY-178B、
VxWorks MILS 2 和 LynxSecure 是采用该方式的典型代表。

基于微核方式分区实现途径，Wind River 公司于 2008 年推出了满足 MILS 架构的安
全（security）操作系统 VxWorks MILS 2。该系统于 2009 年 3 月提交进行 CC EAL 6 认证
评估，已于 2011 年 12 月完成评估。VxWorks MILS 2 分区内核采用 Wind River Hypervisor
虚拟化技术构建，满足 SKPP 1.03 安全需求，支持分区间安全通信。VxWorks MILS Plat-
form 2.0 已被 RTI 公司用于构建 RTI 数据发布服务系统，技术成熟度至少达到 7 级。据报
道，该平台也被 Boeing、Lockheed Martin 和 Raytheon 考虑用于新一代机载设备的开发。

LynuxWorks 公司以 CC EAL 7 为目标设计 LynxSecure，于 2008 年 3 月发布 LynxSecure
2.0，2010 年 5 月发布 LynxSecure 4.0 及配套 IDE 开发环境 Luminosity 4.7，支持分区内
Guest OS 及应用的开发。据报道，LynxSecure 使用虚拟化技术构建，并支持双核处理器，

技术成熟度至少达到 7 级。

可以看出，国外对于抗恶劣环境嵌入式计算机安全软件的研究主要集中在研制分区、高安全、微内核的嵌入式操作系统上，并努力保证对其产品应用实践进行推广。

2.7 工程化技术

国外在抗恶劣环境嵌入式计算机工程化技术上的发展主要集中在液冷、板卡加固、射流冷却、工程仿真等方面。

(1) 液冷技术

作为嵌入式计算机工程化技术的主要技术点，液冷技术一直保持着蓬勃的发展态势，目前国外液冷技术的发展方向有小型化液冷循环、射流冷却技术等，主要技术点也掌握在 DHY、Lytron 等专业公司手中。

液冷冷板和机架的技术主要掌握在 DHY 公司手中，其生产的液冷冷板和机架已经在 F/A-18 和 F-22 飞机上应用。液冷技术的研究主要涉及液冷冷板和机架的关键技术，如真空钎焊工艺、真空气淬工艺、机械加工工艺、除屑工艺以及冷板和机架的检测的技术和方法^[23]。

液冷循环技术涉及各种热组件，包括系统冷却技术和热交换器等，国外主要由 Lytron 公司提供高性能、高效、经济的液冷解决方案。液冷和热管理技术广泛应用于电力电子、军事和航空、核发电、半导体设备以及超级计算机等市场^[24,25]。

(2) 板卡加固技术

嵌入式计算机的板卡散热和加固技术是工程化技术中的一个重要方面，并已形成了相关的国际标准^[26]（ANSI/VITA48.0 ~ 48.5）。国外的 Mercury 公司参与制定了加固级板卡的散热标准，在散热方面拥有多项专利。

(3) 射流冷却技术

从 20 世纪末开始，射流冷却成为电子设备冷却领域的热门话题，特别是美国在这方面投入了大量的人力、物力和财力。以亚利桑那大学和马里兰大学为首的研究机构在射流冷却换热理论上进行了深入的研究。ISR、Parker 等公司可提供部分采用射流冷却技术的产品。为 F22 提供集成机架和 LRM 模块的 DHY 公司拟在未来机载电子设备上采用射流冷却技术。

(4) 工程仿真技术

在抗恶劣环境嵌入式计算机领域，工程化仿真技术主要应用于雷电及 HIRF 防护方面。商用运输飞机主要根据数值仿真的计算结果进行 FAA/JAA 防雷击适航认证。工程仿真技术用于验证全机雷击试验，为飞机局部电磁脉冲试验计算电缆电流和端部电压，在取得大量仿真和局部试验数据后，不再需要进行全机雷击试验。国外 EMA3D 公司利用软件进行工程仿真，相对于全机雷击试验，软件仿真能节省研制经费 \$1.5M^[28]。

在航空领域，工程仿真技术还可以用于对在 HIRF 电磁环境条件下，机舱内部传导干扰和辐射干扰的情况^[29]进行仿真和测试。波音 787 商用喷气式飞机闪电/HIRF 论证：

EMA 公司为其新型商用喷气式飞机波音 787 闪电/HIRF 开展论证项目，首先通过进行闪电间接效应分析来确定航电仪器设备端口上感应的瞬态波形；其次通过仿真了解闪电保护条结构、形状对性能的影响，帮助进行闪电保护条的设计以达到最佳防护效能；然后，通过仿真和测试了解在 HIRF 电磁环境条件下，机舱内部传导干扰和辐射干扰的情况，最终顺利获得试航证。

国外在抗恶劣环境计算机工程化技术的研究上，目标非常明确，就是通过优化的物理工程手段，帮助计算机对抗大气及外太空中的恶劣环境条件。以此目标为基础延伸出的液冷、抗辐照等关键技术，也是研究的热点。

3 国内研究进展

我国在嵌入式计算机的抗恶劣环境技术方面的研究起步较晚，通过近年来积极的探索，虽然已经取得比较明显的进步，但与国外发达国家相比，仍存在不小的差距。

3.1 体系结构

我国从 20 世纪 80 年代开始，逐步在工业控制和航空领域开展嵌入式计算机的研制，通过借鉴国外嵌入式计算机的体系结构，发展自主的嵌入式计算机体系。通过近 30 年的努力，研发和生产了各类嵌入式计算机，在车辆、船舶、飞机、工业医疗等领域得到了广泛应用，促进了信息技术在国民生活中的应用发展。

在“十一五”期间，我国开始关注恶劣环境下嵌入式计算机的系列化发展问题，开展了嵌入式计算机领域的体系结构的系列化研究，形成了单机体系结构、冗余容错体系结构、分布式体系结构和并行体系结构的分类，确定了不同分类的嵌入式系统的标准规范和指标体系，提取出了关键指标，并形成了一些评测方法和评测工具。

近几年，随着物理信息融合系统的快速发展和应用，环境因素对嵌入式计算机的影响愈加严重。国内高校和研究机构针对嵌入式计算机的体系架构设计技术展开了广泛的研究，希望在提高系统安全性和可靠性的基础上，能够形成通用体系架构和设计方法学，从而利用市场规模效应推动嵌入式计算机产业的快速发展。

3.2 容错技术

我国早期研制并获得型号广泛应用的容错计算机多采用集中式结构，这种计算机体积大、成本高、扩展性差，无法满足系统对多个功能的综合和灵活配置的需求。

随后，我国开展了分布式计算机冗余容错技术的研究，在“十一五”期间成功研制了“ 3×2 ”高可靠与高可用计算机系统。该系统的物理形式包括三个结点计算机，分布于三个结构独立的机箱内。结点计算机通过 AFDX 高速网络连接，系统任务在各个结点

计算机上执行。它突破了异构安全等级的功能软件综合配置技术，然而，分布式容错、高确定性通信网络、高精度全局同步、高完整性计算和细粒度故障隔离等关键问题并没有彻底解决。

近年来，我国针对新一代装备控制平台对嵌入式计算机的新要求，进而开展了分布式容错理论、LOCKSTEP 计算技术^[30]、故障动态重构、部件自愈技术、故障预测与健康管理技术^[31]开展了系统研究，取得了一些阶段性成果并开始应用，但总体上，同国际先进水平还存在不小差距。

综合核心处理机采用“健壮分区隔离技术”，为共享平台资源的多个功能和应用提供适当的保护机制，确保在一个分区内的功能或应用的行为不对其他任何分区内的功能或应用的行为产生不可接受的影响^[30]，同时满足所规定的安全性和完整性等级要求。

在抗单粒子翻转方面，我国于 20 世纪 80 年代中期先后利用中子 ^{252}Cf 裂片、重离子、高能质子等粒子开展静态随机存储器的单粒子效应实验研究，参与单位有中国原子能科学院、北京大学重离子物理所以及西北核技术研究所。由于仍处于起步阶段，很多实验理论方法、规范标准都参考了国外专业同行的研究成果。例如，元器件多选用单粒子效应较不敏感的器件，优先考虑反熔丝 FPGA、Flash FPGA、非易失性存储器，降额使用高压功率器件；系统采取冗余设计（空间冗余、时间冗余和信息冗余）；以三模冗余模式进行表决，应用 EDAC 技术（奇偶校验、Hamming 编码、ECC 编码等）；软件措施包括运行程序冗余、参数刷新、过程划分等。从 2012 年开始，在中航工业一飞院的牵头下，由圣涛平研究员主持，我国多家航空电子设备研制单位进行了大气中子单粒子效应实验的试点工作，通过获取相关环境下的单粒子效应数据，在机载电子设备中开展中子单粒子效应防护设计研究，以减小单粒子效应对航空电子设备带来的危害。

综上，我国已掌握了集中式容错机理与关键技术，在分布式容错、确定性网络、器件抗单粒子翻转方面正迎头追赶国际先进水平。

3.3 可靠网络技术

我国的可靠网络技术正从低速向高速、从不确定向确定、从专用向开放发展。

我国的 AFDX 网络研究工作始于 2007 年，在 2008 年国家大飞机立项后取得了较大的进展，技术趋于成熟。以中航六三一所为代表的国内供应商先后突破了 AFDX 网络交换机、端系统、网络配置、网络仿真、网络管理、数据加卸载等技术难点，形成了相关的产品，并通过了协议符合性测试。2010 年 12 月至 2013 年 1 月，中航六三一所按照 DO-254 和 DO-178B 的要求，开展了 AFDX 网络交换机的适航工作，开发过程得到了 FAA DER 的认可，并签署文件。目前，该所已经为 C919 大型客机提供了 AFDX 网络交换机、端系统、RDIU 等原理样件。国内还有一些大学利用 AFDX 网络开展基站通信等方面的应用研究^[32,33]。

TTE 网络的研究相对落后。中航六三一所利用多模光纤，在网络交换机和端系统上分别实现了高精度的分布式时间同步算法，并以此为基础突破了监护、时间触发通信调

度、高速交换等技术，研制出原理样机。然而，由于网络容错理论的欠缺，更新一代的样机研制缺乏深层次的理论指导。

3.4 片上系统技术

我国的 SoC 技术处于高速发展之中，许多高校、研究所和企业都投入了大量的人力和物力来进行关键技术公关。虽有一些芯片级产品问世，但仍需要进行进一步的试验测试。

复旦大学设计了一款基于 65nm 工艺的 RISC 处理器^[34,35]，从架构设计、电路设计、芯片后端设计多个层次保证其高性能、低功耗的特点，该处理器可以将 RS 前向纠错解码算法的吞吐率提高 4~70 倍。

某航天研究所研发了一款适用于下一代卫星使用的小型化 SoC^[36]。该 SoC 以 FPGA 上的 SoC2008 芯片为核心，集成了 SPARC V8 兼容型 CPU、UART、Timer、Watching-Dog、中断控制器、存储器控制器、1553B 等空间计算机常用外设控制单元，同时在 SoC2008 芯片内部实现了星载各种敏感器、执行机构和遥控遥测接口。为了进一步提高 SoC2008 的抗辐射性能，针对 SEU、SEL、TID 等进行了抗辐射加固，采用 EDAC、自动纠错回写、流水线重启、CACHE 校验、冗余触发器、三模时钟树等组合策略实现了对软错误的容错处理。

中航六三一所也相继推出了 1553B、AFDX、659、6713 等航空专用 SoC。同时，在最新研究中提出了基于部分可重构技术的片上系统架构^[37]，可以满足高性能、高灵活、高可靠和可扩展的应用需求，有效降低了机载嵌入式计算机的体积、重量和功耗 (SWaP)。

我国的芯片级可靠性试验主要包括温度、恒定加速度、老练测试、电测试、密封检漏、质量一致性检验等。更高层次的系统级应用验证尚不够完善，目前还没有统一的国家级试验标准。以航空应用为例，虽然已经制定了《GJB 360B—2009 电子及电气元件试验方法》、《GJB 128A—1997 半导体分立器件试验方法》、《GJB 548B—2005 微电子器件试验方法和程序》、《GJB 4027A—2006 军用电子元器件破坏性物理分析方法》等标准或规范，但是它们一般采用跟随整机模块进行试验的方法，多属于常规检验。由于缺少大型、专用试验设备，无法在地面模拟所有可能的真实系统工作环境，且受试飞时间和成本等诸多因素的影响，模拟的（恶劣环境）场景组合非常有限。

SoC 虽具有集成度高、密封性好、抗干扰能力强等先天优势，但由于其设计复杂、前期制造和试验验证工作繁琐，加之我国还没有形成完整、系统的体系和标准，SoC 的抗恶劣环境能力有待进一步提高。

3.5 健康监控技术

我国在故障诊断、预测和健康管理方面的研究需求和研究对象主要集中在航空、航

天、船舶、兵器等复杂高技术装备领域，研究主体以高校和科研院所为主，主要研究内容涉及体系结构及关键技术研究、智能诊断和预测算法研究（基于模型的方法、基于数据的方法和基于统计的方法）以及测试性和诊断性研究等方面。

然而，从工业制造部门和装备使用者的角度看，我国对该技术的理论研究和应用研究尚没有有效衔接，应用需求尚未获得足够的系统而明确的理论分析和引导。以航空应用为例，飞机综合健康管理技术（Integrated Vehicle Health Management，IVHM）^[38]方面近年来虽然出现了一些基础性研究成果，但由于缺乏良好的研究管理机制、研究体系分散、统一高效的协调机制欠缺等因素，造成理论和应用脱节，基础研究缺乏背景支撑和实验验证等致命缺陷。在飞机状态监控方面，虽然已成功采用非航空电子监控、BIT等技术实现在线监视，且具有较高的测试覆盖率，使得飞机的测试性、维修性得到极大的改善，但这些技术的作用仅仅局限于有限的设备状态监控，对设备进行状态趋势分析和预测的能力尚不成熟，此外，也不能完成空、地一体的实时飞机状态数据传输、综合维护和管理。

3.6 安全软件技术

在高可靠嵌入式操作系统方面，我国在陆、海、天、空领域基本上均有一家较有实力的科研院所推出了面向本领域的嵌入式操作系统，能够满足现阶段现役武器装备的需要，但在技术水平和满足未来武器装备需求方面，仍与国外专业嵌入式操作系统厂商存在差距。

在微核分区操作系统方面，我国工业界尚未见到成熟的嵌入式产品出现，学术界也形成研究热点。对于构建 MILS 架构操作系统的重要途径之一——虚拟化技术，学术界也主要关注服务器虚拟化技术，对嵌入式领域的虚拟化分区技术鲜有考虑。目前，只有北京航空航天大学在中航六三一所机、弹载计算机航空科技重点实验室航空科学基金的资助下，开展了面向 IMA 系统的虚拟化平台关键技术研究，对机载领域综合化系统中虚拟化分区技术的应用进行了有益的探索。

在信息安全方面，中科院相继研制和开发出了红旗 Linux 安全操作系统 RFSOS、SECIMOS 和 SecLinux^[40]。RFSOS 在 Linux 核心层为上层软件提供了基本的安全功能支持，主要包括增强的身份标识与鉴别、细化的自主访问控制、特权用户职责划分、强制访问控制、审计跟踪以及安全管理等内容。与 RFSOS 相比，SECIMOS 主要增加了形式化规范与验证、隐蔽信道分析和可信路径等安全功能和保证，而且系统的整体结构和具体的实现方法也不同。SecLinux 以 TCSEC 标准的 B1 安全等级和中国等级准则第三级的要求为设计目标，提供身份标识与鉴别、自主访问控制、强制访问控制、最小特权管理、安全审计、可信通路、密码服务和网络安全服务等方面的支持。此外，南京大学开发了基于 Linux 的 Softos 安全操作系统，该系统提供了强制访问控制、审计、禁止客体重用、入侵检测和扩展的系统资源访问控制等功能模块。

综上，我国对嵌入式操作系统的研究较为活跃，已经出现了一批应用于各自专业领

域的嵌入式操作系统，并且通过采用时空隔离等技术具备了一定的可靠性支持。但是，对构建高安全 MILS 架构所需要的微内核和虚拟化技术以及嵌入式信息安全技术的研究尚未成熟。从架构入手，合理布局，并采取一定的安全可靠性验证手段和设计方法，是推动高安全软件发展的有效途径。

3.7 工程化技术

在散热技术方面^[41,42]，我国主要采用热管散热。各类计算机芯片普遍采用铝质、铜制散热器外加风扇，依靠强迫对流空气来冷却发热器件。随着嵌入式计算机向小型化、高性能、高可靠的方向发展，微型热管和回路热管成为研究的重点。目前，我国已经研制出系列风冷平板型散热器、IGBT 桥式整流器模块用热管散热器、螺栓型半导体用热管散热器以及各种组合式热管散热器，并且发布了相应的行业标准。但是，热管散热在电子电器设备中的应用主要集中于大功率电器设备，在微电子设备中仍应用不多。此外，我国正在开展液体冷却和射流冷却方面的工程化技术研究，但与国外水平还有较大差距。在浸入式冷却方面，我国仅限于理论研究。

在抗振防冲技术方面，我国对基于振动被动控制的外加固理论进行了深入研究，已获得一些实施性强的研究成果，例如钢丝绳减振器和无谐振峰隔振缓冲器已在计算机外部设备加固中得到应用。随着对较低频振动、强冲击等抗振防冲要求的提高，研究振动主动控制理论及研制相应的振动主动控制装置成为大势所趋。然而，我国目前在计算机外部设备加固中尚无对此项技术的实用研究成果和实际应用。

在电磁兼容技术方面，对于密闭式机械结构的低功耗嵌入式计算机来说，国内外技术水平差距不大。但是，对于开放式、高电压、大功耗的嵌入式计算机，我国尚有很长的路要走。

在防 HIRF 技术方面，中航六三一所较早开展了专门的研究工作，取得了一定的成果和经验，并逐渐形成了一套以雷电、HIRF 一体化防护理念为基础的符合现有国情、紧跟国际发展趋势的工程设计理论。然而，由于我国领空 HIRF 环境测评工作的缺失，中国民航和其他适航当局还不能准确地定义或控制航空器使用过程中将会曝露于 HIRF 的能量水平，我国尚无条件制定相关标准，目前不具备完整、独立的开展 HIRF 试验验证的条件。

在嵌入式系统小型化、集成化要求越来越高的大背景下，我国的工程化技术在新材料、新工艺的应用方面还存在欠缺，在灵活性、高效性上更是十分薄弱。

4 国内外研究进展比较

嵌入式系统抗恶劣环境技术在国外已经有了广泛和深入的研究，尤其美国处于技术领先地位，国内与国外在嵌入式系统应用的关键抗恶劣技术的比较如下表所示：

项目	国外情况	国内情况	技术差距和问题
体系结构	DECOS 框架、GENESYS 框架、SCARLETT 架构、开放式架构	有分类的体系架构，如单机体系结构、冗余容错体系结构、并行体系结构等	技术相差 1 代，大约 5~10 年，国内没有完善的标准规范，未建立统一的架构
容错技术	分布容错、强 EDAC 存储校验、LOCKSTEP 计算 ^[43,44] 、故障静默、消息同步。故障检测与诊断，动态重构、自愈	集中容错、EDAC 存储校验、故障逻辑、故障检测	技术相差 1 代，先进分布式容错理论和方法未建立
可靠通信总线网络	千兆/10Gb 以太网、AFDX 网络、FC 网络、IEEE1394、SPIR-WIRE 总线、实时网 ^[45]	10 ~ 100 兆以太网、AFDX 网络、1553B 总线、CAN 总线	应用技术相差 5~10 年，未掌握核心技术，高性能通信芯片不能自主保障
片上系统	建立有专门的应用验证组织机构和验证方法，制定了较完备的标准	行业试验标准和方法，没有国家级试验标准	技术相差 1 代，大约 5~10 年；没有国家级试验标准体系与专门机构
健康监控	PHM 理论、技术、软件或者应用解决方案研究较深入，开始广泛应用于实际应用协同	针对专门领域开展基础理论与方法研究	技术相差大约 5~8 年；基础理论与方法不完善、理论研究和应用研究尚没有有效衔接
安全软件	EAL6+ 认证的 Integrity 操作系统、EAL6 认证的 VxWORKS 操作系统	成功研制嵌入式操作系统 ACoreOS653、ReWORKS 等	技术差距较大，没有自主的高安全操作系统内核
环境防护技术	热导精确仿真、射流液冷。基于仿生机制和模型的全新防护，研发出使用智能电磁防护材料	热导仿真、风冷。屏蔽、滤波、接地	技术相差 5~10 年。无新型防护材料、没有可行模型

总之，统一体系架构、高可靠器件的设计验证、PCI-E 体系、高速数据传输、高效液冷、LOCKSTEP 等高完整性技术、抗粒子辐照防护技术、智能电磁防护材料、基于仿生机制和模型的全新防护技术、高可靠操作系统是国外嵌入式系统抗恶劣环境技术的关键和亮点，国内技术水平同国际先进水平有 3~10 年的差距。

5 发展趋势与展望

嵌入式系统抗恶劣环境技术是一项多学科综合的技术领域。在相关支持技术的发展推动下，嵌入式系统正在朝着微型化、网络化、智能化、综合化的方向发展，可以预测抗恶劣环境技术的发展趋势如下：

(1) 体系结构

应用于安全关键系统的嵌入式计算机，由于技术成本高、专业应用领域窄，从而技术发展成本居高不下。为了能够利用微电子的工业化产能和市场的规模效应，需要研究高安全嵌入式计算机的通用体系架构和设计方法学，充分利用 COTS 产品技术，减少开发成本并缩短开发周期，从而推动产业的快速发展。

(2) 容错技术

安全关键应用将采用全网络化分布式开放系统架构，通过片上容错、板级容错、总线容错、动态重构和功能自愈的方法，利用安全分区操作系统，在鲁棒的时间、空间分区保护机制下，实现模块级不同安全级别功能综合以及系统级功能多余度配置和故障自动重构，从而提高系统冗余资源的利用率，大幅提高系统可靠性与可用性。此外，系统在满足高可靠容错需求前提下，未来将结合数字验证、原理验证以及系统综合验证等技术途径，构建基于综合化模块、高可靠的嵌入式容错系统。未来嵌入式系统的容错技术将向分布式、基于消息传输的容错管理方向发展，容错粒度将由通道提高到模块以及芯片。

(3) 网络技术

嵌入式网络系统将在提高网络通信速率的同时，加强网络的强实时、高安全等保障能力。目前 AFDX 网络虽然已经成功应用在 A380、A400M、B787、A350 等机型中，其实时性仍无法满足安全关键系统的应用要求。TTE 采用了时间触发的 TDMA 通信方式，基于千兆以太网能够同时兼容三种业务流的传输，包括用于安全关键系统的时间触发 TT 消息、用于非安全关键系统的速率限制 RC 消息（也即 AFDX），以及用于维护等无实时性要求的尽力而为 BE 消息（普通以太网）。TTE 的这种高速、强实时、高可靠、多业务流等特性，使其能够满足未来统一网络的应用要求，因此时间触发以太网 TTE 技术将是未来发展热点。

(4) SoC 技术

SoC 系统实现了单位面积上多电路的集成封装，基于 SoC 的嵌入式系统具有体积小、重量轻等优势，因此 SoC 技术的主要应用领域将是嵌入式系统。目前 SoC 芯片设计技术已经得到突破，但是 SoC 要满足嵌入式系统的环境使用要求，未来的重点将是加强对各种恶劣环境对 SoC 的作用机理研究，建立一套全方位、全体系、多层次以及极端条件下 SoC 设计、验证及试用体系，稳步提高 SoC 抗恶劣环境的水平，为嵌入式系统提供高安全、高可靠、自主可控的集成电路产品。

(5) PHM 技术

PHM 是实现自主后勤保障的核心，能够有效降低维护成本，提高系统安全性。目前，嵌入式系统的健康监控仍然处于理论研究阶段，主要技术难点是构成嵌入式系统的硬件部件失效无法准确度量，在故障发生前没有明显的特征参数；大多数嵌入式系统的 PHM 缺乏失效机理的准确分析；智能的嵌入式系统普遍采用 VLSI 电路实现，规模大、功能相关性弱，PHM 诊断代价较高，由于成本因素，PHM 通常无法实现。未来 PHM 技术的研究方向将包括特定应用领域的嵌入式系统的失效机理及故障特征提取方法，寿命预测模型及方法，VLSI 的片内测试机理和片上测试方法，“征兆”数据的融合算法研究。

(6) 安全软件

嵌入式计算机主要应用在工业、交通、医疗、金融和通信等领域，一旦出现安全问题，不但会给个人以及社会造成巨大的经济损失，甚至会危害国家和地区的安全。在信息安全的需求牵引下，将开展嵌入式系统软件的微内核实现的基本安全（Security）服务

机制、细粒度的安全访问控制技术和分区间的安全信息通信等技术的研究。在 MILS 架构的约束下，合理布局和实现系统的安全支持功能，并进一步通过形式化验证，使高安全软件从根本上具备信息安全支持能力。

(7) 工程化技术

随着新材料、新技术的应用，嵌入式系统的小型化、功能集成化将为嵌入式系统应用提供更强的灵活性和更多功能。未来工程化技术将以复合材料为突破口，综合研究高效散热、抗振防冲以及电磁防护等工程化技术。高效散热技术主要研究 APG 高导热材料与智能热管理技术，抗振防冲技术主要研究主动减振技术，电磁防护技术主要研究强电磁脉冲防护等技术。

6 结束语

随着信息技术的迅速发展，抗恶劣环境技术的目标已从传统的抵抗恶劣的物理环境，逐步发展到抵抗复杂的电磁环境和易受攻击的网络环境。在这个背景下，国外开展了体系架构、容错技术、可靠网络、安全软件等方面新技术的研究，在理论和工程方面取得了一批有影响力成果，有效提升了抗恶劣环境计算机的适应能力。这些新技术给我们带来了巨大的挑战，也给我们带来了前所未有的发展机遇。

根据上述分析，我们认为面向安全关键系统的体系架构、分布式容错技术、强实时高可靠网络、抗恶劣环境 SoC、PHM 故障机理、高安全嵌入式软件、面向小型化和极端恶劣环境的工程化技术等是国内抗恶劣环境计算机技术研究的重点。国内的研究者将紧随国外技术发展趋势，结合抗恶劣环境嵌入式计算机具体需求，深入开展研究，努力缩短与国际水平的差距。

参考文献

- [1] Hermann Kopetz. An Integrated Architecture for Dependable Embedded Systems. Proceedings of the 23rd IEEE International Symposium on Reliable Distributed Systems , 2004.
- [2] Neeraj Suri. DECOS: Dependable Embedded Components and Systems. Darmstadt University of Technology , 2007.
- [3] Roman Obermaisser, Hermann Kopetz. A Candidate for an ARTEMIS Cross-Domain Reference Architecture for Embedded Systems. Vienna University of Technology , 2009.
- [4] Eric M. Nelson. Open Architecture Technical Principles and Guidelines 1.5.8. IBM Federal CTO Office , 2008.
- [5] P. Bieber, F Boniol, M Boyer, E Noulard, C Pagetti. New Challenges for Future Avionic Architectures [J]. Aerospace Lab Journal Issue 4, May , 2012.
- [6] 张喜民, 徐慕. 先进战斗机的综合飞行器管理系统综述[J]. 电光与控制, 2011, 18(11): 1-8.

- [7] Ben Achballah, Ahmed, Ben Saoud. Investigating The Use of The AFDX Protocol As A Network- On-Chip. International Conference on Design & Technology of Integrated Systems, Nanoscale Era, 2012.
- [8] SAE AS6802. AEROSPACE STANDAND, 2011.
- [9] TTEthernet Specification[S]. TTTech Computertechnik AG, 2008.
- [10] Sandi Habine, Barry Cook, Paul Walker, et al. Using FPGAs and a LEON3 FT Processor to Build a Flying Laptop. RESPACEMAPLD, 2011.
- [11] Greenbelt, Maryland. SpaceCube 2.0: An Advanced Hybrid Onboard Data Processor [J]. Electronics/ Computers. NASA Tech Briefs, 2011-02.
- [12] Jan Andersson, J. G. , Roland Weigand. Next Generation Multipurpose Microprocessor. Data Systems in Aerospace, 2010.
- [13] Technical Note on NGMP Verification Next Generation Multipurpose Microprocessor. Aeroflex Gaisler AB, 2010-11.
- [14] Luo Pei, Zhang Jian. A High Reliable SOC On-board Computer Based On Leon3[C]. Proceedings of IEEE International Conference on Computer Science and Automation Engineering(CSAE) , 2012.
- [15] 世界电子元器件, 2012-11.
- [16] Dr W Scheuren. Joint Strike Fighter Prognostics and Health Management, AIAA98-3710.
- [17] Abhinav Saxena, Indranil Roychoudhury, Jose R Celaya. Requirements Flowdown for Prognostics and Health Management, AIAA 2012-2554.
- [18] Abhinav Saxena, Indranil Roychoudhury, Jose R Celaya. PHM Framework Design Based on Data Warehouse[C]. 2012 Prognostics & System Health Management Conference(PHM-2012 Beijing).
- [19] Leonardo Ramos, Takashi Yoneyama, Cairo L Nascimento Jr. How Aircraft Operators Can Benefit from PHM Techniques, 2012.
- [20] Deepak kumar singh, Mrs. A. Prabha. Industrial Machine Monitoring Using MEMS Technology [C]. International Conference on Computing, Electronics and Electrical Technologies, 2012.
- [21] ALVES-FOSS J, HARRISON W S, OMAN P, et al. The MILS Architecture for High Assurance Embedded Systems[J]. International Journal of Embedded Systems, 2006, 2(3/4) : 239-247.
- [22] Aeronautical Radio, Inc. ARINC Specification 653-1 Avionics Application Software Standard Interface [S]. Annapolis: Aeronautical Radio, Inc, 2003.
- [23] Holyoke Lakeville, Heat exchangers for Avionics & Electronics, Di Hed Yokes. Inc, 2008: 2-15.
- [24] Jungho Kim, Ken Kiger, Lytron Thermal Productions Introduction. Lytron Market, 2010: 8-21.
- [25] Robert Bergholz. Thermal Management and Advanced Cooling [R]. General Electric Aircraft Engines, 2008: 1-9.
- [26] Randy Banton. A VITA Based Framework for Ruggedized COTS Electronics[J]. Mercury Computer Systems, 2009. 06 : 2 ~ 28.
- [27] Scottsdale, Arizona. Advanced Cooling Technology Workshop[J]. ISR SprayCool, 2008. 07 : 7-21.
- [28] Kanchan M Kelkar. Modeling Tools for Improving Productivity of HIRF Design of Avionics. University of Maryland, Inc. 2009: 12-14.
- [29] Kanchan M. Kelkar. Modeling Tools for Improving Productivity of HIRF Design of Avionics. University of Maryland, Inc. 2009: 15-21.

- [30] 张喜民, 魏婷. ARINC 659 容错数据总线测试验证系统研[J]. 西安电子科技大学学报, 2010, 41(6): 164-169.
- [31] 姜斌, 赵静, 齐瑞云, 等. 近空间飞行器故障诊断与容错控制的研究进展[J]. 南京航空航天大学学报, 2012, 44(5): 603-609.
- [32] Liu, Xiao-sheng; Zhang, Peng-yu; Ren, Hui-fen; Zhao, Zhen-feng. Applying AFDX to Improve the Real-Time Performance Of Substation Communication Network[C]. IEEE 7th International Power Electronics and Motion Control Conference - ECCE Asia, Harbin, 2012.
- [33] Lei Rao; Wang, Qixin; Liu, Xue; Yufei Wang. Analysis of TDMA Crossbar Real-time Switch Design For AFDX Networks[C]. IEEE Annual Joint Conference, 2012.
- [34] 权衡, 肖瑞瑾, 欧鹏, 等. 基于 65 nm 工艺的高性能低功耗处理器设计[J]. 计算机工程, 2012, 38(19): 250-253.
- [35] Quan Heng, Xiao Ruijin, You Kaidi. A Simple High- efficient Intercore Communication Mechanism for Multi-core Systems[C]. Proc. of Asia-pacific Signal & Information Processing Association Annual Summit & Conference, 2011.
- [36] 张少林, 杨孟飞, 刘鸿瑾. 空间应用 SoC 研究现状简介[J]. 航天标准化, 2012, 3: 14-20.
- [37] 赵谦, 赵小冬, 刘作龙. 机载嵌入式动态可重构计算机设计技术研究[J]. 航空计算技术, 2013, 43(2): 113-116.
- [38] 常琦, 袁慎芳. 飞行器综合健康管理(IVHM)系统技术现状及发展[J]. 系统工程与电子技术, 2009, 31(11): 2652-2657.
- [39] 徐晓光, 叶宏. 分区间通信在航空电子系统中的设计与实现[J]. 航空计算技术, 2005, 35(1): 45-48.
- [40] 赵志科, 卿斯汉, 李丽萍. 支持动态多策略的安全体系结构应用研究[J]. 计算机工程, 2004, 30(3): 63-66.
- [41] 张娅妮, 陈菲尔, 田沣. 机载电子设备冷却散热技术的发展[J]. 航空计算技术, 2012, 42(4): 113-116.
- [42] 齐广峰. 军用计算机热管理综述, 航空计算技术, 2009. 9: 132-134.
- [43] Next Generation Integrated Avionics System[R]. GE Aviation, 2009.
- [44] Application Note: PowerPC 750 GX Lockstep Facility[T]. IBM. 2008.
- [45] D A Gwaltney, J M Briscoe. Comparison of Communication Architectures for Spacecraft Modular Avionics Systems. NASA, 2006.

作者简介

牛文生 博士, 中航工业西安航空计算技术研究所总工程师, 研究员, CCF 高级会员, CCF 抗恶劣环境专业委员会委员, 主要研究方向是嵌入式计算机体系结构。



李亚晖 博士，中航工业西安航空计算技术研究所高级工程师，CCF会员，主要研究嵌入式计算机体系结构。



赵小冬 硕士，中航工业西安航空计算技术研究所研究员，主要研究嵌入式计算机体系结构。



谢建春 博士，中航工业西安航空计算技术研究所工程师，主要研究嵌入式计算机图形图像处理技术。



文鹏程 博士，中航工业西安航空计算技术研究所工程师，主要研究嵌入式计算机图形图像处理技术。



张喜民 博士，中航工业西安航空计算技术研究所工程师，主要研究嵌入式计算机容错技术。



王乐 硕士，中航工业西安航空计算技术研究所工程师，主要研究嵌入式计算机体系结构。



孙允明 硕士，中航工业西安航空计算技术研究所研究员，主要研究嵌入式计算机体系结构。



普适计算的研究进展与趋势

CCF 普适计算专业委员会

摘要

近年来，普适计算相关技术领域已发展到全新阶段：一方面，物联网技术、社交网络、移动互联网、智能终端、感知及交互技术的迅速发展，推动了信息空间、物理空间和社会空间的融合；另一方面，基础设施的广泛布设和智能终端的泛在嵌入，计算能力以前所未有的程度深入到人们的生活，由此衍生出大量“围绕人的服务”。在此基础上，“人—机—物”三元世界逐步深度融合，正催生未来普适计算的一种愿景——iCPS (I and Cyber-Physical-Social Systems)。其中，人是核心，智能是目标，“人—机—物”三元融合是途径。围绕 iCPS，本文介绍了国内外研究机构在用户建模、基础设施和服务平台以及智能终端感知交互技术等方面开展的一系列工作及成果。然后，结合学术界和工业界提出的不同概念原型，探讨了“人—机—物”模型、普适智能服务、智慧终端和人机交互等方面的未来发展趋势。最后分析了普适计算领域存在的挑战。

关键词：普适计算，普适服务，用户模型，智能终端，人机交互

Abstract

Ubiquitous computing technologies have developed to a bran-new stage. On one hand, the rapid development of the technologies about things of internet, social network, mobile internet, smart terminals, sensing and interacting promotes the tight integration of the cyber space, physical space and social space; on the other hand, with the ubiquitousness of hardware facilities and smart terminals, computing is deeply embedded into people's living environment, many human-centric services being derived as a result. On the basis of the rapid development of the technologies, the condition of nowadays, that humans and smart things being tightly integrated, are making the future vision of ubiquitous computing, iCPS (I and Cyber-Physical-Social Systems) . iCPS is a human-centric system, with the goal of providing intelligent services by integrating humans and smart things. For the vision of iCPS, many research effort have been concentrated on user models, infrastructure, smart terminals, and HCI. On the basis of the achievements, many demos have been proposed. By analysing the key concepts of demos, the paper discussed the future trend and the potential challenges of ubiquitous computing.

Keywords: ubiquitous computing, ubiquitous service, user model, smart terminal, human-machine interaction

1 引言

1.1 背景：新技术革命推动普适计算发展

3D 打印、能源互联网、新型纳米材料……一波新的技术浪潮正在改变世界，其特征

是新能源、新材料与信息技术的高度融合与发展。2012年4月，著名期刊《经济学人》对当前这次新技术浪潮做了专题论述^[1]，认为它可与前两次技术革命——17世纪的“蒸汽技术革命”和20世纪初（1912年前后）的“电力技术革命”相媲美，可称为“第三次工业革命”。

不论这次技术浪潮是否达到了“工业革命”的程度，但它的确已对世界产生了重大影响。2010年8月起，伊朗核设施遭受“震网”计算机病毒攻击，部分离心机被迫关闭，其核计划可能被延迟两年；自2010年底起，突尼斯、埃及、阿尔及利亚、叙利亚等国相继发生所谓“茉莉花革命”，Twitter、YouTube、Facebook等新在线社交媒体成了这次政治剧变的时代标志；2010年美军成立了“赛博司令部”，2011年白宫公布了美国“赛博空间国际战略”，从国家战略的高度研究、规划、建设自身在信息空间中的作战能力^[2]，对世界各国产生很大影响。

基于对上述一系列事件的判断分析可知，目前电子信息技术和产业作为基础性、战略性、支柱性、先导性产业的特征日趋显著，向其他产业的渗透性更加突出；信息技术与生物、新材料、新能源等技术的交叉创新，可能引发新一轮产业革命；与此同时，人类社会、信息空间和物理世界的不断融合，将改变传统的计算模式，要求建立与之适应的新型信息基础设施，并将对诸多关键技术领域提出新的要求和挑战。

自Mark Weiser于1991年提出“Ubiquitous Computing”（普适计算）理念以来^[3]，普适计算已走过20多年的发展历程。普适计算最初的设想就是将信息技术融入到日常生活用品中，使它们从人们的视线中消失，以最自然的方式提供服务，彻底突破桌面计算模式。

可以看出，信息空间与物理空间融合的思路从一开始就渗透在普适计算的理念中。但相比Weiser的时代，目前的相关技术领域已发展到一个全新阶段。

首先，信息技术整体发展已进入到所谓的“四后”时代，即“后PC”、“后Moore”、“后IP”和“后图灵”（涌现新型计算模式），信息技术不断向其他产业渗透。

其次，各类新型支撑技术为普适计算提供新的发展平台。云计算技术的迅猛发展，近乎无限地扩展了普适终端的计算、存储能力；物联网技术重塑了物物关系，特别是加快了物体间的信息活化、通信以及共享；社交网络不仅彻底改变了用户间获取和传播信息的方式，而且给出了全新的人与人、人与信息、人与物的关系。

再次，智能终端及移动互联网的爆炸式增长，使用户和终端紧密耦合，围绕用户需求的、上下文敏感的各类服务（例如LBS服务）迅猛发展；同时，感知和计算能力也随着终端的普及，泛在地嵌入到物理环境中，智能手机、智能电视、智能汽车和各种传感设备都是最典型的终端实例。

综合上述情况，我们认为：信息产业正向其他产业快速扩张，在带动起一系列新兴增长极的同时，也将信息能力融入其他产品形态，客观上推动了信息空间、物理空间和社会空间的融合；同时，得益于基础设施的广泛布设和智能终端的泛在嵌入，计算能力以前所未有的程度深入到人们的生活，“计算能力的普适化”正在出现；移动终端与用户紧密绑定，由此衍生出大量“围绕人的服务”，相应的新型产业生态系统正被构建起来。

1.2 未来普适计算愿景——iCPS

互联网的出现为人类信息的及时共享提供了基础，移动互联网使得人们与信息空间（Cyberspace）连接的可持续性得到了保证，物联网将人们活动所依赖的物理环境中的各种组成单元接入到信息空间，社交网络和脑机接口正成为人接入信息空间的两种典型方式（前者为松耦合方式、后者为紧耦合方式）。随着这种“人—机—物”三元世界的逐步深度融合，正催生未来普适计算的一种可能的愿景——iCPS。在此系统中，人是核心，智能是目标，“人—机—物”三元融合是途径。下面将通过一组生活场景来说明 iCPS 将给人类生活带来的影响。

（1）场景 1：从智能家庭到智慧城市

根据 IBM 的预测，到 2050 年，70% 的人口将会生活在城市。环境的智能化程度和范围正不断扩大，从一个个智能家庭发展为智慧城市，我们将迎来更加智能、便捷、绿色环保的城市生活。

未来的人居环境将由一个个具有智能的家庭单位连接成网络，从而实现资源的共享和优化配置。一方面，通过各种感知设备的大量应用，智能家庭单位中的功能单元将能理解用户的行为习惯和生活规律，并通过与用户社交网络的融合，主动与其他进行人性化互动。例如，空调系统可以在用户下班之前提前开启，将房间调节到舒适的温度；吸尘器等清洁装置可以自动调整打扫房间的时间和频率；电视、平板电脑等设备可以自动为用户下载感兴趣的内容。在交互方面，用户将能直接通过语音、表情、手势和姿态动作等与家居设备进行交互。例如，通过语音打开或关闭灯光，或者通过手势或肢体动作进行游戏操控等。另一方面，由各种智能家庭单元连接成的网络将能更好地协调和配置资源。例如，电网将能够根据各个家庭的用电情况进行负载均衡；网络系统能够根据各个家庭的使用情况和带宽变化自动缓存内容和分配带宽；以家庭为单位的社交网络可以促进居民在节能减排、环境保护方面的行动，从而为绿色人居环境的建立提供基础。

由各个智能家庭单元组成的未来城市将成为低能耗、低污染、资源优化配置和高度可回收的典范。在节能方面，智能的电气设备能够根据用户的习惯自动调整工作状态，并通过时序优化降低电网峰值，从而达到节能的目标；在减排方面，自动化的交通系统将对车辆进行统一调度和控制，从而避免交通拥堵，降低污染；在资源利用方面，通过 RFID 等追踪技术，可以对设备的使用进行管理，从而提高资源的利用率。

（2）场景 2：从车联网到智能出行

车联网是物联网的一部分，车联网时代的智能汽车与城市交通信息网络、智能电网以及社区信息网络完全地连接起来，为人类提供诸多信息服务。未来的交通出行将建立在智慧路网和智能车联网的基础上，提供快速、安全、环保的出行体验。一方面，未来的汽车将具备高度智能的车载信息系统，可以与城市交通信息网络、智能电网以及社区信息网络连接，随时随地获得即时资讯，并且作出与交通出行有关的明智决定；另一方面，未来的路网将设计成立体式交通网络，以允许不同速度和模式的车辆快速通行。汽车自动驾驶技术将成为现实，并且在“车联网”的保护下实现零交通事故率。智能的

“车联网”甚至可以帮助司机订票、寻找停车场，以及自己找到充电站完成充电。

在车联网和自动驾驶的时代，老人、孩子甚至残疾人都可以拥有自己的汽车，人们的出行将变得更加轻松。一个典型场景如下^[4]：Anna 自小双目失明，行走都需要导盲犬的帮助。一天 Anna 的父亲举办个人音乐会，他需要提早去场地彩排，Anna 只能一个人前往音乐会。她一个人坐上车，按动导航系统的开关，不用为看不见路担忧。导航系统显示出父亲事先给她设定好的行车路线，汽车在自动驾驶的状态下高速前进。途中汽车提示电量不足，Anna 只需告诉汽车去充电，汽车便会临时改变路线，带她到最近的充电场地，充电完毕后，汽车继续按照原来路线高速前进，Anna 准时到达父亲的演出现场。

(3) 场景 3：移动 + 社交 + 网络

未来的移动社交网络不再是完全虚拟的社会网络，而将逐步把在线交互与位置、行为活动等线下感知信息相结合。物理世界中原本不相识的个体越来越有可能通过移动社交网络平台成为朋友，并可以通过实时分享彼此的物理状态（如位置、活动等）增强线上线下之间的交流。下面以智慧校园（如图 1 所示）为例做说明。



图 1 移动社交网络在智慧校园中的应用

学生 A 和 B 是某大学一年级的新生，他们两人虽然不认识，但由于两人在移动社交网络的签到都包括图书馆、A 教学楼、羽毛球馆等场所，而且两人专业相近，故移动社交平台通过匹配两人的历史活动记录和个人信息发现二者的相似性，对两人进行朋友推荐。在得到双方确认后成为移动社交网络上的好友，并且能够相互查看对方的状态。这样，A 在自习时遇到不会的问题，看到 B 也在同一教学楼就可以寻求 B 帮助；而当 B 想去打球时也可以查看 A 的活动状态，进而决定是否 A 有时间陪他去打球。移动社交平台还可以根据二人的相似性把 A 常去的兴趣地点推荐给 B。

(4) 场景 4：社会感知

社会感知是普适计算的高级阶段，群智感知是实现社会感知的有效途径。在未来群智感知中，每个人都会充当多个角色。例如作为记者，去发现和报道所经历的活动和事件；又或如环境保护者，进行周边环境情况共享等。下面以突发性社会事件的感知为例进行说明。

某城市正在举行大型比赛，很多群众围观比赛并拿出手机进行拍照或者拍摄视频。这时，突然在赛场某个角落发生了爆炸。在这种危急情况下，所有人会迅速散开并逃跑。在警局收到报警并赶到现场后，就很难查出行凶者的线索。这时，当时围观比赛的群众会在到达安全地点后将手机拍摄的照片和视频上传到警局的网络上。大量群众拍摄的突发事件发生前事发现场的照片和视频将使警局的工作人员在没有目睹这场突发事件的情况下就能够基本还原突发事件发生前现场的情况和围观群众的状态，从而帮助警局人员能够快速锁定可疑人员并进行侦察，加速案件的侦破过程。采用这种公众提供集体智慧的方式，警局工作人员就不必再大费周折地寻找目击证人并进行采访，大大减少了他们的工作量，并且群众的智慧更能反映出突发事件发生时的真实情况。波士顿爆炸案发生后，FBI 采用现场用户贡献的数据进行嫌犯确认就是一个很好的实例（如图 2 所示）。



图 2 社会感知与波士顿爆炸案

2 国际研究现状

从上述计算机相关领域的研究现状与进展来看，未来普适计算正在朝着类似 iCPS 的愿景发展。这主要体现在：“人—机—物”进一步充分互联，各类异构终端普遍分布，信息能力泛在融合到日常物体中；各类异构交互设备进一步隐藏于周边环境，交互方式更加丰富和自然；海量知识的积累使得计算环境对用户需求和个性化理解更为准确；“围绕人的计算”仍将是未来普适计算最鲜明的特征。

在现阶段研究成果的基础上，学术界和工业界打造了层出不穷的概念原型，从不同角度阐述了对未来信息产品和计算模式的理解。窥一斑而知全豹，我们选取了若干较具代表性的原型，尝试对其进行分析并探讨目前国际相关领域的现状及趋势。

2.1 “人—机—物”模型

(1) 社交互联网 (Social Web of Things)：人和物充分互联

随着物联网应用的发展，物联网技术正在与社交网络结合，形成一种以私有物体信息上网为媒介，以社交娱乐为目的的新网络——社交物联网。社交物联网利用物联网的感知检测技术，原本在我们生活中的普通物体可以实现实时的信息化，通过网络技术、云计算技术、云存储技术等，物体就可以“活生生”的“生长”在社交网络上。比如，用户将自己的微信账号和家电的微信账号绑定后，每天出门的时候，或许只需对着微信

说一句“我要去上班了”，家里的电视、空调、电灯就会自动关闭。2011年2月，爱立信推出了称为“社交互联网”(social web of things)的概念视频^[5]，设想人和各种联网设备以平等的方式在社交网络上互动，如图3所示。该概念突破了传统的“人控制设备”的模式，让人与设备以更平等、更智能的方式沟通和协作。2011年底，浙江大学也尝试将物理设备（饮水机）连接到人的社交网络（新浪微博）上^[6]。当水烧开或者没有热水时，饮水机会自动发布微博通知用户。

可以预见，那些原本仅属于我们生活空间中的“私有物品”一旦接入社交网络，就可以通过这些既有的平台，以全新方式与全球用户互动。

(2) 智能玻璃：信息和物体的深度融合

美国康宁(Corning)玻璃公司制作了一段关于想象未来生活的视频^[7]。在影片当中，女主角走进洗手间，一边刷牙、洗脸，一边利用镜子的触控功能安排今日的工作行程，甚至是回复朋友信息，如图4所示；男主角在床边，利用墙壁上的触控式荧幕，就可得知今日的天气以及交通状况。智能的玻璃具备感知周围物体的能力，与移动终端、家居用品、交通工具等友好地融合在一个智能环境中。智能的玻璃尝试将信息以最便捷的方式直接叠加到日常环境中，体现的是一种将虚拟空间和真实空间深度整合的思路，预言了未来日常生活中物体和信息深度融合的趋势。

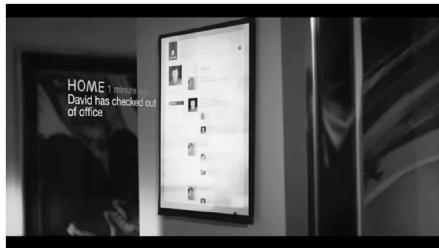


图3 社交互联网



图4 智能玻璃

从上述两个概念原型可以看出，一方面，未来的社交网络不仅是人与人的网络，也将是人与物、物与物的网络，信息化设备通过“社交物联网”与人进行平等的互动，这种互动方式充分利用了现有社会网络平台、接口和协议，社交网和物联网将以类似的方式走向融合。另一方面，信息技术不断整合到日常生活物体中去，这种深度整合的思路不仅将催生出多种新产品，还使得信息基础设施真正无处不在，可能孕育出全新的计算模式，从而极大地改变人们的生活和工作。

2.2 智慧终端

(1) 智能手机：随身交互与计算中心

目前智能手机的计算和存储能力已经超过早期的PC，并具备PC没有的传感、通信

等功能，同时其能力正被云计算无限扩展。截至 2012 年底，全球智能手机用户数已突破 10 亿^[8]。为满足用户需求，越来越多的传感器件和传感技术被运用到智能手机上，催生了一系列基于智能手机感知的应用，并掀起了移动感知的研究热潮^[9]。我们可以很清晰地看到，未来的智能手机将成为重要的感知和计算设备，越来越多的新型应用将围绕手机展开，并最终成为“人—机—物”三元世界相互沟通和融合的门户。

(2) 可穿戴设备：计算机的未来形态

随着传感技术的进步，越来越多的微型可穿戴传感装置不断涌现。

一方面，一些可穿戴设备可以通过与智能手机相连接，提供友好的人机交互界面、数据存储和云服务访问。例如，MYO 臂带通过检测手势变化时的肌肉放电特征来识别用户手势。借助这种新型的生物电感应技术，MYO 臂带可以检测手部运动时肌肉的电信号传输，甚至在手指细微动作执行前的瞬间就能感知到运动信号。利用蓝牙与智能手机连接后，用户可以通过手势控制设备开关、音乐播放，甚至小型四翼飞行器；另一些可穿戴设备则隐藏于衣服、鞋子等日常用品中，对人的活动进行无时无刻的感知。例如，Nike Lunar 篮球鞋通过内置的芯片可以记录运动时每一次跳起的高度、跑步的步速等数据，并将数据转化为衡量运动量的 NikeFuel 数值。通过智能手机应用和云端账号便可即时查看数据，还可提供每周排名、个性定制训练课程、查看比赛中细节数据等功能。MYO 臂带和 Nike + 篮球鞋如图 5 所示。



图 5 MYO 臂带和 Nike + 篮球鞋

另一方面，一些功能更强大的可穿戴设备也在日益兴起，并有望成为类似智能手机的随身服务设备。例如，2012 年 4 月，谷歌推出了谷歌眼镜，其形态是一副眼镜，但功能类似一台整合了谷歌云服务的智能手机，其独特形态便于用户随时获取各类信息，并通过投影显示器以增强现实的方式提供给用户（如图 6 所示）。

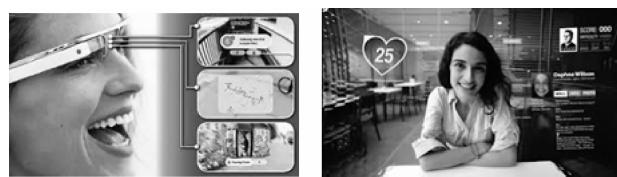


图 6 谷歌眼镜

可以预见，在云计算、增强现实、三维显示等技术的支持下，可穿戴计算设备的计算、存储、显示等能力得以大幅扩展，形态日新月异，异构化、小型化、随身化的各类设备将持续涌现。计算能力正在摆脱传统的“计算机”形态，通过各种穿戴设备，以

“呼之即来，挥之即去”的方式将信息提供给用户。

(3) 智能汽车：一种新的信息服务终端

智能汽车关注汽车本身的智能化，能够辅助司机完成更加安全和可靠的驾驶任务。车联网和自动驾驶是目前智能汽车领域的两个热点方向，为未来更加安全、高效的智能交通系统描绘了雏形。

在车联网方面，通用汽车于2010年3月在上海发布了基于车联网的概念车EN-V^[10]，如图7所示。EN-V是面向未来城市个人交通的解决方案，致力于城市交通的零油耗、零排放、零堵塞和零事故。EN-V通过城市智能交通网络和车辆自身驾驶系统实现智能导航和行驶，并可通过智能手机定位车的位置。多辆EN-V智能车之间能够组成网络，共享道路信息，并能前后排列成队列在道路上行驶；EN-V具有障碍检测功能，可以自动避开行驶过程中的障碍物；作为两轮电动车，采用智能平衡技术保证车身平衡。



图7 通用EN-V智能概念车

与此同时，汽车自动驾驶技术已经日趋成熟。截至2012年，谷歌公司的无人驾驶车累计驾驶里程已达30万英里（48万公里），仅发生1起事故（等红灯时被追尾）。随后，谷歌无人驾驶车在加州和内华达州拿到上路许可证。国家自然科学基金委亦于2008年设立了由郑南宁院士领衔的“视听觉信息的认知计算”重大研究计划，其任务之一就是集成人的听视觉认知机理等成果，研制无人驾驶车辆验证平台。2012年11月，军事交通学院自主研发的无人驾驶汽车成功进行了从北京至天津的行驶测试，预计于2015年将测试从北京行驶至深圳。谷歌无人驾驶汽车和国内无人驾驶汽车如图8所示。



图8 国内外无人驾驶汽车

综上所述，车联网和自动驾驶技术将改变汽车产业的形态。用户将从驾驶行为中解放出来，汽车则从人操控的交通工具逐步向自主机器人演变，这将对人车交互方式产生

深远影响。汽车将不仅是交通工具，还是集聚信息服务的重要平台。“围绕人的计算”正取代“围绕机械的计算”成为汽车价值的核心。

2.3 人机交互

普适人机交互充分融合了传感器技术、智能计算技术和嵌入式系统等领域的最新技术。交互设备将逐渐隐藏于环境之中，交互的深度和广度不断扩展，朝着以人为本、增强体感、方便用户的方向不断向前发展。

(1) 交互设备逐渐隐藏到幕后

随着环境设备的不断增强，用于满足普适交互的基础设施逐渐完善。一方面，体现在环境中多尺度显示屏的普及。以智能手机为代表的小型交互屏幕解决了交互屏的可携性问题，提供了无处不在的交互界面；以智能电视为代表的大型屏则具有较好的视觉效果和沉浸感，为基于肢体动作的交互提供了可能；而大型投影屏则提供了大范围多用户信息共享和协作的可能，并可借助智能手机等终端设备进行远程交互。另一方面，可穿戴设备的普及带来了交互体验的提升。通过隐藏在眼镜、手表、鞋子，甚至植入于体内的芯片，这些设备将提供无缝的感知和交互体验，用户无需借助专用设备即可通过自然的方式使得设备理解其行为意图。

一个典型的例子是 MIT 的第六感项目^[11]。该项目使用的交互技术在一定程度上将现实空间与虚拟空间融为一体。整个系统由 4 个套在手指上的彩色标记环、一个小型摄像头、一个便携式投影仪等设备组成（如图 9 所示）。通过摄像头追踪标记环的运动，可以识别用户手势信息，进而结合当前的用户情境，推断出用户动作的意图，例如拍照、文本输入等。执行结果和操作界面可以通过投影仪投影到任何合适的显示位置，达到增强现实的效果。

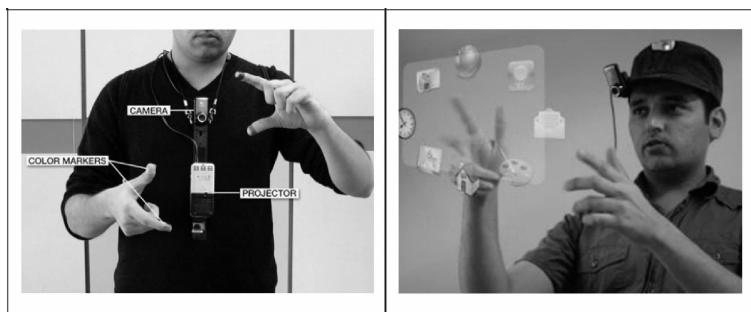


图 9 第六感装置示意图

(2) 交互的广度和深度不断扩展

依托各类新设备和交互模态，人机交互的广度和深度也将不断得到扩展。

一方面，感知技术的进步使得多种交互手段融合成为可能，交互方式不断推陈出新。表 1 列举了一些典型异构终端及代表性应用。

表 1 可连接到智能手机的异构终端

传感类型	原理	代表性产品
脑电感知装置	通过头皮表面放置的电极，记录脑部生物电活动的波形图（EEG），从而检测人脑兴奋程度	NeuroSky MindWav 头箍 NeurowaveMico 耳机
手势感知装置	通过红外感应器跟踪和识别手指运动；通过检测手势变化时的肌肉放电特征来识别手势和动作	LeapMotion 手势控制器 MYO 臂带
姿态感知装置	结合彩色摄像头和红外摄像头建立人体三维模型，从而捕捉肢体动作；使用加速度传感器感知肢体运动	Kinect 体感控制器 Amiigo 腕带
行为感知装置	监测运动行为，估计运动量，从而给出健康建议	FitBit 腕带 Nike FuelBand 腕带
生理感知装置	通过生理传感器感知用户的心率、血压、体温等数据，进而监测睡眠质量等健康信号	Basis Band 手表 Larklife 腕带

在这些异构设备中，肢体和手势识别技术进展迅速。包括以 Kinect 为代表的深度摄像头感知、LeapMotion 为代表的红外感知、MYO 臂带为代表的生物电信号感知等。微软的 Kinect 体感控制器^[12]结合了摄像头和红外传感器，可以识别约 30 米范围内的深度信息。利用 Kinect，可以对用户的姿态和手势进行实时识别，并用于 3D 建模和控制虚拟场景中的物体摆放等操作；LeapMotion 则通过红外传感器实现了高精度的深度信息感知，可以感知其上方 8 立方英尺内的手指动作，识别精度可达 0.01 毫米，还可以同时识别不同手指间以及铅笔的差别，应用前景十分广阔。Kinect 和 LeapMotion 手势控制器如图 10 所示。



图 10 Kinect 和 LeapMotion 手势控制器

另一方面，复杂信号感知技术（特别是脑机接口等技术）使人机交互的层面不断深入。如何理解人的意图是神经科学、计算机、心理等学科的重要交叉研究课题，该领域近年的进展带来了一系列新型交互模式，脑机交互是这类更“深入的”人机交互模式的代表。

一个典型的例子是 Campbell 等人提出的 NeuroPhone^[13]。NeuroPhone 可以让用户通过意念控制手机进行拨号（如图 11 所示）。当用户注视的联系人头像闪烁时，系统能检测到特殊的脑电信号，从而触发拨号操作。NeuroPhone 展示了脑机交互的初步应用，未来可以运用于用户意图识别等方面。



图 11 NeuroPhone 和 FlyingBuddy2：利用脑电信号进行交互

3 国内研究进展

围绕 iCPS 等可能的普适计算愿景，研究者建立了大量“围绕人”的普适计算模型，试图基于普适感知等手段，将人在生活与工作过程中的各种外在表现和内在变化完整地刻画和记录在信息空间中；同时，在普适计算基础设施和服务平台方面，研究者们也做了若干探索性工作，尝试为理论模型提供支撑平台，促进信息空间和物理空间的融合；而随着材料科学、电子科学、云计算以及相关硬件和制造业的发展，智能手机和大量异构可穿戴设备等智能终端也不断涌现出来。这些终端设备融合了感知技术和交互技术领域的最新进展，直接与人交互，形成围绕人的普适感知源和服务源。下面将从用户模型、基础设施和服务平台、新型普适智能服务、智能终端感知和交互技术等方面介绍相关国内研究进展。

3.1 用户模型

普适计算区别于传统计算技术的本质特性在于强调计算模式的“以人为本”。从这种意义上说，人类活动是在普适计算空间中实现信息空间与物理空间融合的纽带。用户实体与计算实体之间、计算实体与计算实体之间的相互交互需随用户移动或环境变化而在不同时空动态演变^[14]。因此，需要以人及其周围环境为核心建立模型，记录、存储、描述“I”在信息空间、物理空间和社会空间的活动轨迹，这也是我们提出 iCPS 愿景的原因之一。

在普适计算中，用户建模的过程是获取、表达和维护用户兴趣、社会关系和行为规律等相关知识的过程。用户模型是推荐系统、行为预测等相关技术的基础。根据应用领域不同，用户建模的主要研究方向包括用户兴趣建模、朋友关系建模和移动性建模等：

- 用户兴趣模型不是对用户个体的一般性描述，而是具有一种面向算法的、具有特定数据结构的、形式化的用户描述。它由表示用户兴趣的对象组成，每个对象都有一个权值信息，权值越高，表明用户对方面的信息兴趣越浓厚。
- 朋友关系建模随着社交网络的兴起受到了前所未有的重视，通过用户在社交网络上留下的数字足迹，融合用户在线上和线下交互的特征，可以对朋友关系进行评估和建模。
- 用户的移动性模型是从研究用户个体的运动机制出发，详细描述用户的具体运动参数（如位置、轨迹、方向、速度等）并分析用户的移动规律，主要可以应用于用户的位置跟踪、预测、兴趣点发现等。带 GPS 的智能手机的普及，使得大规模采集用户的移动轨迹成为可能，为用户移动性建模提供了数据支持。

此外，还有一些用户模型更宏观地刻画了用户模型在物理空间和信息空间中的不同特性，以及用户和环境、服务之间的关系。典型的有浙江大学的“影子模型”^[15]等，较准确地体现了普适计算服务“无处不在”的时空特性、普适计算用户界面的“透明”交

互模式及普适计算“以人为本”的基本理念。

浙江大学从物理空间和信息空间融合的背景出发，根据以人为中心的不同实体之间相互作用的基本原理，提出了基于物理场理论的智能影子模型（如图 12 所示）。用户在信息空间中存在一个自己的数据影子，随时随地与物理空间中真实的“我”相互感知与交互，记录人生信号；计算基础设施中存在一个服务影子，形成个性化、移动式的虚拟化智能空间，随时随地为用户提供智能化普适服务。智能影子的计算范式是从对物理世界的“感知”到反作用至物理世界的“交互”这两个操作交替螺旋式反馈上升的过程，从而达到从人生记录到人生服务。

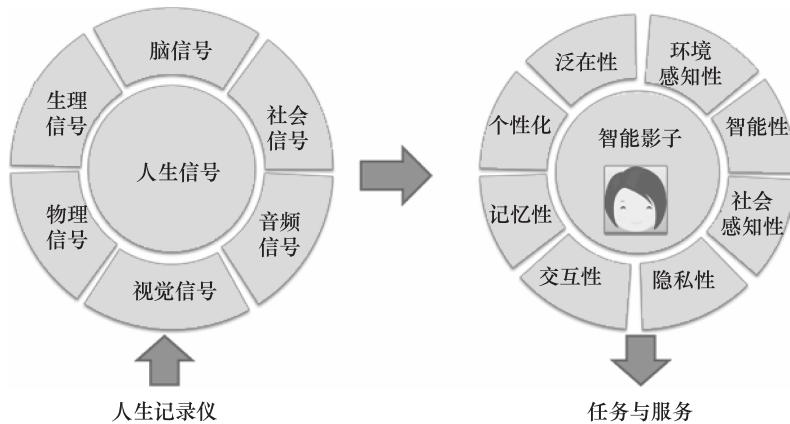


图 12 智能影子的用户模型：从人生记录到人生服务

3.2 基础设施和服务平台

普适计算环境中的基础设施和服务平台对普适计算中大量的联网的设备、物体、计算实体进行管理，为它们之间的数据交换、消息交互、服务发现、任务协调、任务迁移等提供系统级的支持^[16]。正是这些基础设施和服务平台为上述用户模型提供了支撑平台，使得设备和服务能无缝地嵌入生活和工作空间，促进了信息空间和物理空间的融合，在改变提供信息和服务方式的同时，也改变了人们未来的生活和工作。

(1) 移动社交网络系统支撑平台

人类同时生存于物理和虚拟两个社会空间中，但在过去很长一段时间里，这两个社会空间是分离的。移动社交网络的发展为物理和虚拟两个社会空间的连接提供了条件，融合人们线上、线下体验，形成了混合社交空间。在移动社交网络中，感知随时随地发生，交流无处不在，服务触手可及，为实现面向社会发展的感知计算（社会感知计算^[17]）提供了一条切实可行的技术途径。因此，感知计算在移动社交网络中显得尤为重要^[18]。然而，感知数据的实时连续性、服务的及时性等特点对移动社交网络系统支撑平台提出了新的要求。连续实时的用户状态感知功能使移动终端，尤其是智能手机成为一个理想的数据感知平台，借助于智能移动终端不断增强的感知能力，捕获用户周围的状

态信息。同时，服务的及时性以及移动终端的弱计算能力需要以中心服务器为依托，通过分析异构多源感知数据，及时地为目标用户提供个性化的服务。因此，移动社交网络系统支撑平台由手机端的智能感知和服务器端的分布式服务中间件两个部分构成。

移动感知中间件通过智能移动终端中的各种感知设备，包括 GPS 定位、加速计、蓝牙等，实时地采集用户的状态信息，实现用户数字轨迹的实时检测和记录。西北工业大学提出了一个参考智能感知中间件架构，可以划分为感知设备层、汇聚分析层和应用传输层三个层次，如图 13a 所示。

分布式服务中间件是构建移动社交网络系统平台的灵魂，是整个应用系统的中枢。其结构如图 13b 所示，包括网络通信层、基础服务层和应用层。网络通信层主要完成跨协议的通信对接，实现移动终端到服务器、终端到终端之间的互通通信，从而达到屏蔽底层硬件设备差异的目的。基础服务层实现公用服务的封装，为上层应用提供统一的服务机制和调度接口。基础服务层以开放服务网关（Open Service Gateway initiative，OSGi）框架为基础，实现对于公用服务的封装。应用层则面向具体的需求，以公共服务为基础开发各种服务应用，从而实现为移动社交网络用户提供适时的个性化服务。

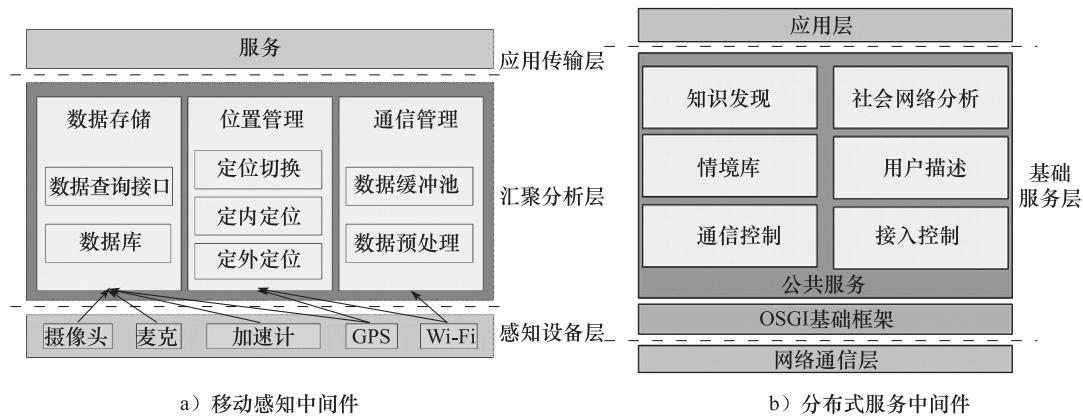


图 13 移动社交网络系统支撑平台

在移动社交网络平台服务方面，国内还有很多其他研究工作。例如，微软亚洲研究院提出了结合手机端移动轨迹感知数据和后台数据集成分析的移动社交网络位置服务平台，能根据用户间行为相似度进行旅行推荐^[19]。诺基亚中国研究院刘震等认为移动社交网络的关键是实现线上线下的真正融合，提出了瞬时社交网络（Ephemeral Social Network）概念来为用户提供更好的移动社交服务体验^[20]。清华大学唐杰则对移动社交网络平台中的用户行为预测方法进行了研究^[21]。

（2）群智感知计算

社会及城市感知任务具有范围广、规模大、任务重等特点；传统感知技术依赖于专业人员和设备，覆盖范围受限，维护成本高。在此背景下，群智感知计算作为一种新的感知模式被提出来。它利用大量普通用户使用的移动设备作为基本感知单元，通过物联网/移动互联网协作，实现感知任务分发与数据收集利用，最终完成大规模、复杂的社会

与城市感知任务。

清华大学刘云浩团队对群智感知计算的起源、概念与特质进行了归纳和总结，提出群智感知计算的主要挑战，包括弱联网环境下数据采集、数据质量与数据选择、大数据处理等^[22]。其团队的研究人员还成功把群智感知的概念应用到智能交通、室内定位等领域^[23]。微软亚洲研究院研究人员利用群体贡献的手机数据进行用户相似度匹配和好友推荐^[24]。浙江大学利用从杭州及上海收集的大规模出租车移动轨迹数据来分析人们的移动规律、城市动态及地点语义^[25]。北京邮电大学利用用户贡献的数据来改进机会网络中的数据路由性能^[26]。

西北工业大学提出了群体智能的基本计算架构^[27]，共分为5层（如图14所示）。从大量的数字脚印中挖掘和理解个人与群体活动模式、大规模人类活动和城市动态规律，把这些信息用于各种创造性服务，例如，社会关系管理、城市感知、环境监控等。



图14 群智感知计算体系架构

(3) 异构空间情境管理

近年来，普适计算研究者提出不同的情境管理方法来满足不同智能空间的需求，如智能家庭、智能办公室、智能医院等。这些情境管理被设计来存储本地空间产生的情境信息并为本地应用提供调用接口。随着普适计算技术的发展和普及，我们可以预测在不久的将来，现实生活中将布满各种智能空间，人们在这些不同的智能空间之间移动，在不同的智能空间留下不同的情境，这些情境综合起来才能构建出用户真实而完整的上下文信息。这样就产生了一个新的挑战，当智能空间由单个转为多个时如何管理在这些异构空间中产生的情境。由于用户的移动性，用户情境可能分布于不同智能空间，当一个

空间查询某个用户或某移动实体的情境时，究竟应该从哪个空间获取。

单一空间的情境管理已有相关研究，如 Jose^[28] 依靠电信运营商建立了“Vader system”，根据用户当前所在位置确定用户所在空间并获取情境信息。近年来，异构空间的情境管理引起了研究者的兴趣，如西北工业大学提出的“iCROSS”^[29]解决方案。图 15 为“iCROSS”系统架构。该方法中移动用户通过向所处空间的情境服务器申请访问，情境服务器进行授权并在一个统一的注册中心更新用户进入点，从而解决移动用户注册问题。同时提供了一个跨空间情境查询算法来解决移动状态下用户情境查询问题。

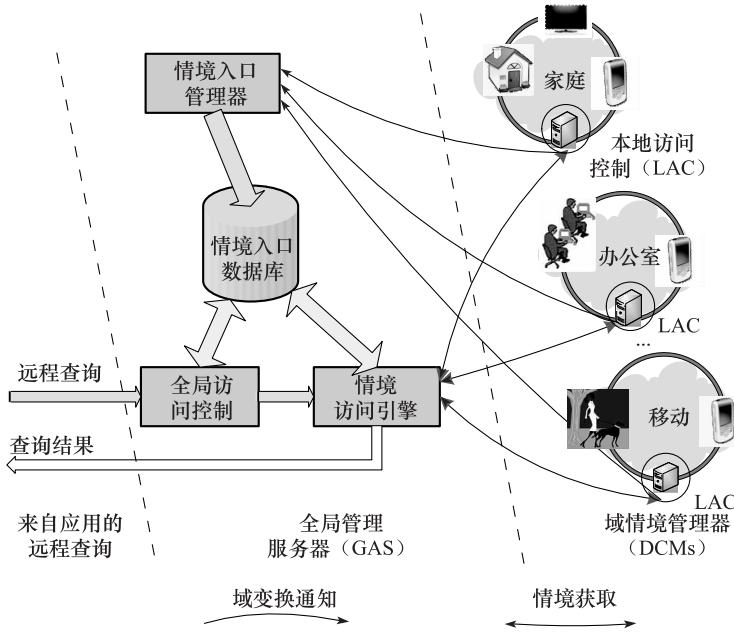


图 15 iCross 异构空间情境管理平台

在虚拟空间和真实空间融合的背景下，情境在“虚实”之间的融合也相当重要。针对移动终端上的微博浏览问题，微软亚洲研究院谢幸等人综合利用时间、地点、浏览历史、社会关系和活动构成用户浏览时的“情境”。首先学习微博内容和用户情境之间的关系，然后通过检测当地的微博话题，根据用户的利益偏好推荐用户相关话题，直至找到用户首选的内容^[30]。

3.3 新型普适智能服务

(1) 基于移动感知的精准广告推送

精准广告是随着互联网发展而出现的新事物，属于个性化推荐的范畴。例如，百度和谷歌等利用用户关键词、所在地域和查询时间等对用户兴趣进行分析，进而提供相关联的广告推送。基于移动感知的精准广告则从传统 PC 迁移到更为广阔的物理空间。主要依托移动终端，如智能手机、平板以及智能可穿戴设备等，通过传感器采集用户的历史

数据在服务器端建立用户的偏好模型和用户行为模型，在进行广告推送时根据用户的偏好、用户当前或未来一段时刻所处的实时情境给特定用户推送广告，以达到精准推送的目的。图 16 给出了华为与西北工业大学合作的精准广告平台的一组场景。如果从用户的行为模型可以知道用户经常在早上 8 点左右乘坐地铁，从用户的偏好模型发现用户平时喜欢浏览时事新闻，那么当通过用户移动终端上的传感器发现用户在乘坐地铁，且当前处于空闲状态时，推送引擎就可以向用户推送新闻套餐服务。如果用户为学生，根据用户行为模型发现该用户在节假日有外出游玩的爱好，而在离用户家不远的一个游乐场准备在即将来临的节假日搞优惠活动，那么推送引擎会将该优惠活动推送给该用户。

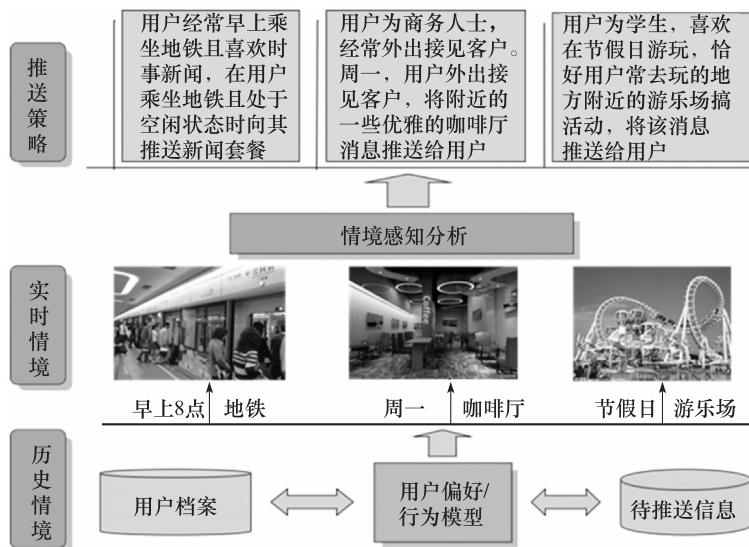


图 16 移动精准广告推送

随着带丰富感知功能的智能手机和平板的普及，移动精准广告推送将获得前所未有的关注。通过感知数据分析获得用户兴趣偏好和实时状态（位置、忙闲状态等），它一方面可以让商家能快速地定位目标客户，另一方面让用户及时有效地获得自己需要的信息。移动精准广告的火爆，充分展示了在大数据的支持下，目前的计算系统正以前所未有的广度和深度认知用户的行为，这为普适信息服务提供了机遇和支持，同时将不断推动研发人员和用户在隐私保护、安全等相关问题上的思考与实践。

（2）社会感知服务

社会感知计算（Socially-Aware Computing）是通过人类生活空间部署的大规模多种类传感设备，实时感知、识别社会个体的行为，分析、挖掘群体社会交互特征和规律，辅助个体社会行为，支持社群的互动、沟通和协作^[17]。社会感知计算的核心在于“感知”二字，有两层涵义，首先感知现实世界（Sensing），然后觉察并做出响应（Aware）。具有社会感知能力的计算机系统可以帮助人们了解当前状况，提升社交技巧，促进良性交互，满足个人的社会性需求。前面曾提出智慧校园的例子对社会感知进行说明。为了促进校园中的社会交互，提高公共资源的使用率，西北工业大学实现了一个信息资源共享

的智能校园系统，其中以自习室查询应用 Where2Study、群智感知应用 I-Sensing、社会活动推荐服务 GroupMe 等为典型代表，如图 17 所示。

- Where2Study^[18]是一个自习室状态共享的应用查询系统。通过检测 Wi-Fi 信号的 RSSI 信号强度，实现用户的室内定位功能。用户通过移动终端可以查询每个教室的使用情况，同时能够清楚地了解每个教室座位的剩余数量。
- I-Sensing^[18]是一个基于群智感知的应答系统。系统中的每一个用户充当一个可移动的感知结点，为系统提供相应的输入信息。使用者可以就校园内公共设施的使用情况，例如运动场地是否空闲等问题向系统发起一个任务请求。系统引入了社会竞争机制以达到激励群体用户参与和接收任务的目的。
- GroupMe^[31]是一个社会活动采集、挖掘及智能化推荐系统。系统能自动记录用户在校园社区内与其好友活动交流的历史（如一起自习、聚餐、开项目会议等），根据社会图（Social Graph）分析及群组计算（Group Computing）等方法进行活动群组发现。当用户组织活动时，可根据移动感知得到的用户情境（所在位置、时间、和谁在一起等）及群组亲密度信息等进行智能化群组推荐。



图 17 智慧校园感知服务：Where2Study (a, b)、ISensing (c)、GroupMe (d, e)

此外，南京大学和香港中文大学的研究人员对移动社交网络中基于用户协作的高效内容分发服务进行了研究^[32]。微软亚洲研究院利用群体贡献的手机感知数据（如声音、照片等）进行分析，对用户所处的位置语义进行识别^[33]。

(3) 协作式智能健康辅助

普适计算领域顶级期刊《IEEE Pervasive Computing》的编委文斯·斯坦福（Vince Stanford）指出，基于普适计算的老年人生活辅助技术将成为未来解决老年人日常生活问题最有效、最具活力的技术之一^[34]。近年来，随着社交网络的出现，基于社会交互与普适健康感知融合的协作式健康辅助成为新的研究热点。

为了督促老年人按时服药，西北工业大学设计了基于移动社交网络的协作式老年人服药系统^[18]。该系统在智能家庭的基础上引入了移动社交网络的思想，通过社会化的提醒，一方面使老年人能够按时服药，另一方面增强老年人之间的社会交互。首先，借助智能家居中的感知能力，系统自动检测老人是否按时服用药物，当检测到用户没有按时服用药物时，系统自动形成一个服药提醒任务，并通过交互式终端向目标主体发起服药提醒，目标主体在接收并完成服药任务后，通过社会竞争机制给提醒服务的发起者以反馈。图 18 给出了协作式服药系统的网络结构图。另外，北京大学在情境感知服务系统 CBS³（Context-aware Based Smart Service System）的基础上提出了 iCARE 服务^[35]，避免患有老年痴呆或由于其他疾病造成自主意识丧失的老人丢失。iCARE 的主要业务是通过在老人居住社区内布置的蓝牙结点采集数据，判断老人是否走出安全区域。当老人走出正常区域时会被非正常区域的结点检测到，iCARE 会向老人的子女或社区保安报警，以避免老人丢失。中科院计算所提出了基于智能手机的 PPCare 系统，能够监控老人的日常行为并提供健康服务支持。该系统包括 4 个功能：运动规划、卡路里消耗预测、跌倒报警、生理监视。该系统已经在两个老年人社区进行了用户体验并取得了较好的反馈^[36]。

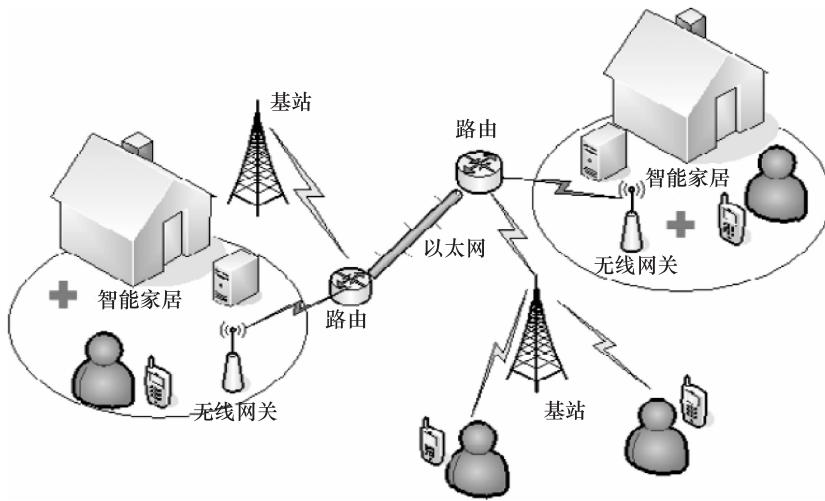


图 18 基于移动社交网络的协作式服药系统架构

3.4 智能终端感知和交互技术

随着硬件技术的发展，各种智能终端方兴未艾。智能手机正在迅速成为个人计算和

通信的核心设备，以可穿戴设备为代表的新型异构终端不断涌现。这些终端设备融合了感知技术和交互技术领域的最新进展，将形成以人为中心的普适感知源和服务源。下面将从智能手机和异构终端这两个方面介绍智能终端方面和交互技术方面的进展。

(1) 智能手机

智能手机具有丰富的感知能力。通过各种内置的传感器，可以实现对手机自身状态和周围环境的感知。例如，通过内置的加速度传感器和陀螺仪可以感知手机的运动状态，进而识别用户手势^[37,38]；通过手机摄像头和麦克风可以识别人脸、声纹（Voice Print）等关键生物信息，从而确定用户身份；通过近场通信（NFC）技术，可以从周围的 NFC 标签或 NFC 手机读取信息，从而进行媒体共享或信息传输。基于这些感知能力，已有研究者提出多种新型应用。例如，浙江大学的 MagicPhone^[37] 通过感知智能手机的空间朝向，提出了指向操控的概念，使得智能手机成为环境设备的统一控制器，体现了自然交互的思想；浙江大学的 TouchPhone^[39] 则通过近场通信技术（NFC）感知周围设备的信息，从而实现了智能手机与环境设备之间的任务迁移与协作。例如，当用户播放音乐时将手机与音箱接触，则音乐会自动迁移到音箱中播放。MagicPhone 和 TouchPhone 工作场景示意图如图 19 所示。



图 19 MagicPhone 和 TouchPhone 工作场景示意图

不仅如此，在手机基础信号的基础上，可以提炼出以人为中心的若干智能手机可知的内容，包括位置感知、姿态感知、行为感知、身份感知、情绪感知和社会关系感知等。位置感知通过获得智能手机所处的空间坐标及其所对应的语义信息，使得智能手机和移动互联网应用能够提供基于位置的服务；姿态感知则通过捕捉人的身体动作、手势和表情变化等来识别动作意图。与传统的人机交互方式相比，姿态控制和交互更为直接、自然；行为感知则通过对用户长时间的活动信号进行感知，抽取出用户的活动规律，从

而在适当的时间地点为用户提供个性化的交互方式；而情感感知则试图通过人的语音、表情甚至大脑活动信号来推测人的情感状态和特征，从而更好地理解用户的喜好。

社会感知相对更复杂一些，主要是试图借助前述各种感知手段，从社会活动中发现用户之间的社会交互特征和规律，进而理解用户的社会关系。社会关系感知能帮助用户协调社会行为，促进社群的互动、沟通和协作。中国科学院陈益强等人对终端收集到的蓝牙信息进行分析，可有效地捕获用户周围的环境特征，例如正式会议、与朋友聊天、和朋友逛街，或者独自一人等，特定的情境识别率以及情节推测的准确率达 87%^[40]。

综上所述，基于智能手机的感知交互手段正在经历从单一到多维、从数据到语义、从个体到群体的发展趋势，并将朝着以人为本、无处不在、智能感知与交互的方向不断向前发展。

(2) 异构终端

便携式的头戴式脑电信号采集设备是一类有潜力的终端设备，它可以采集头部特定位置的脑电信号，并经蓝牙传送到智能手机，从而根据信号特征检测大脑的放松程度（attention）和专注程度（meditation）等活动状态^[41]。通过使用多个电极同时采集大脑不同位置的脑电信号，可以获得更丰富的电位特征。其中，事件相关电位中的 P300 信号已获得广泛研究^[13, 42, 43]，浙江大学的 FlyingBuddy2^[44]就是通过检测 P300 电位来识别用户意图。FlyingBuddy2^[44]实现了通过脑电信号控制四翼螺旋飞机的原型。当用户思考特定动作时，系统能检测出脑电信号变化，从而转换成控制四翼螺旋飞机的指令。

此外，还出现了一些基于常规交互设备的新颖交互手段。清华大学通过获取用户移动鼠标形成的光标轨迹，从而呈现出用户控制鼠标时个性化模式。通过研究某个用户使用鼠标时的轨迹，就能识别出该用户是谁，识别率高达 87.5%^[45]。中国科学院的樊银亭等人基于认知心理学相关理论，提出了基于经验感知的自适应用户界面模型^[46]，从界面静态组成元素、动态交互行为和自适应策略三个方面建立了该模型的统一描述。该模型能够根据用户交互历史和上下文环境，准确预测用户意图，实时调整界面布局和交互行为以及主动地适应用户。这些新型的交互方式体现了对用户复杂意图的感知能力，将为未来人机交互提供更广阔的空间。

4 国内外研究进展比较

我国的移动互联网市场规模已成为全球第一，智能终端广泛普及，通信网络大量铺设，在终端研发、网络运营、服务提供等方面形成了一批领军企业，积累了大量关键技术和产品，正在逐步形成非常有利于普适计算发展和应用的生态环境。在这种情况下，通过与国外研究现况的对比，我们欣喜地发现，国内普适计算领域已不断涌现达到国际先进水平的成果。与此同时，我们也发现一些可改进的地方。总的来看，国内外普适计算研究体现出如下几方面特点：

1) 国外大量研究工作由工业界团队参与甚至主导，产业导向鲜明，学科交叉充分，从概念到产品的研发能力更完备。

国外研究群体除高校、研究所等学术机构之外，一些国际工业巨头的实验室和小型科技公司也积极参与普适计算领域的研究工作。这些工业界研究者的参与，使得研究工作的目标和导向更为清晰，研发工作打破专业限制，进行按需整合。其研究内容不仅包含计算机科学，还包含了材料科学、人机工程、通信等领域，甚至医学、认知科学、心理学等领域成果也被充分吸纳。同时，优秀成果也能及时进行市场转化，出现了如Kinect等优秀产品。与此相比，国内研究团队主要来自专业学术机构，学科交叉不够充分，成果转化路径不成体系，成果产业化程度不高。

2) 与国外相比，国内在装备类、器件类方面的研究偏弱，软件、算法和服务方面研究整体水平相对较高。

从具体研究领域来看，国内在普适计算用户模型、基础设施、新型服务和交互理论等方面研究均颇有亮点，部分成果受到国内外研究者的好评和引用。从创新能力角度来看，一些重要、关键的创新概念仍由国外研发机构主导和推动，国内原始创新能力仍有待加强；同时，在可穿戴设备、感知设备方面，基础硬件（如臂带、手环、各种传感器、脑机接口设备等）主要源自国外，部分体现了我国在“研发→制造”产业链条上的弱点，也进一步体现学术机构和生产企业对接的必要性。

5 发展趋势与展望

目前工业界和学术界大量概念展示和相关技术的进展，较为清晰地刻画出了未来普适计算的一系列趋势。同时，新技术浪潮对普适计算提出了新的挑战和思考，包括：

1) 物联网、云计算、CPS、社交网络等技术改变了人与设备、人与环境的关系，设备智能化程度不断提高，人与设备互动性加强。在这种计算环境下，传统普适计算模型需如何演变？

2) 依托脑机接口、大数据分析、社交网络等技术，有可能从信号、数据、社会关系等多个层面认知用户行为，据此可构建怎样的新用户模型？

3) 设备互联技术和云端服务近乎无限地放大了终端的能力，终端成为用户接入计算环境的入口。在这一背景下，未来终端形态会如何变化？

4) 信息空间、物理空间和社会空间进一步交融，信息被深度嵌入到物理环境中，信息服务将具有怎样的新形态？应建设怎样的信息基础设施为构建普适计算环境服务？

5) 新材料、新能源、生物技术和信息技术不断融合，各种新型人机界面和交互范式层出不穷^[20]。在这些新技术和交互模式基础上，如何进一步实现“以用户为中心”的交互体验？

6) 随着大量异构传感设备的使用，大规模、异构、多模态数据的挖掘和处理成为一类挑战性问题，其中包括低质量数据的过滤，大数据的理解和处理技术等。

7) 在大量异构终端和基础设施构成的物理环境中，不断深化的“感知人”计算将最终推动“围绕人”的计算。具体体现在：随着云计算、社交网络和物联网及其他基础设施的发展，智能手机将逐渐成为打通“人—机—物”三元世界的门户，为人机交互、健康护理、社群智能等领域带来革命性的变化；各种异构终端的兴起，将形成以人为中心

心的普适感知源和服务源。以可穿戴设备为典型代表的一批微型终端，通过与智能手机相连接，可以极大地扩展现有设备的感知和交互能力，并能记录关于人的各个维度的信息，从而提供个性化的服务。

8) 普适人机交互领域正处于茁壮成长时期，充分融合了智能计算技术、传感器技术以及嵌入式系统等领域的最新技术，在深度和广度上均达到前所未有的高度。每一种新的生理信号均可能延伸为一种新型人机交互通道，利用这些新型生理信号的基于手势、脑电等新型感知方式的人机交互模式日渐兴起，将提供更为直观、自然的交互体验。

9) 普适计算与社交网络深度融合带来线上线下行为的相互映照，通过线上数据可以对物理事件进行预测，而通过物理世界交互记录可以对线上交流进一步增强。未来的挑战包括如何基于线上线下特征对用户进行建模，如何通过线上线下特征的融合进行行为预测等。

6 结束语

正在涌动的新技术浪潮正不断影响和推动普适计算的发展。现有的基础设施、软硬件环境、各种支撑技术、计算能力、新材料打造了优越的“硬”条件；以4V为特征的大数据、丰富的云-端服务、弹性互动的社交网络构成的“软环境”则将用户、物理环境与计算系统紧密联系在一起。以人—机—物三元深度融合为途径进一步打造“围绕人的计算”正当其时，相关研究领域、工业应用和商业模式正在迎来全新的发展契机。

参考文献

- [1] Chun-Han Lin, King Chung-Ta, Hsiao Hung-Chang. Region Abstraction for Event Tracking in Wireless Sensor Networks. 2005.
- [2] House White. International Strategy for Cyberspace: Prosperity, Security, and Openness in a Networked World. 2011.
- [3] Weiser Mark. The Computer for the 21st Century [J]. Scientific American, 1991, 265(3): 94-104.
- [4] 高弋坤. 车联网开启城市交通新时代 [J]. 通信世界, 2010, 25(24): 22.
- [5] Alendal Mats. The Social Web of Things- a Social Network for Your Devices, Ericsson, http://www.ericsson.com/thinkingahead/idea/110217_social_network_for_you_1968920151_c.
- [6] Chen Longbiao, Yaochun Li, Zeming Zheng, Li Zhang, Dan He, Xiaolong Li, Sha Zhao, Shijian Li, Gang Pan. WaterLady: A Case Study for Connecting Physical Devices into Social Networks [C]. in Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing (UIC/ATC), 2012 9th International Conference on. 2012: IEEE.
- [7] Doherty L, Pister K S J, El Ghaoui L. Convex Position Estimation in Wireless Sensor Networks. in INFO-COM 2001. 2001.
- [8] Analytics Strategy. Global Business Smartphone Sales 2010—2016. 2011, Research report, Strategy Analytics.
- [9] Lane Nicholas D, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, Andrew T Camp-

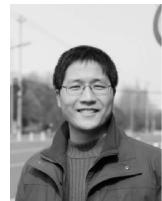
- bell. A Survey of Mobile Phone Sensing [J]. Communications Magazine, IEEE, 2010, 48(9): 140-150.
- [10] 高美嘉. 通用 EN-V 电动联网概念车[J]. 轻型汽车技术, 2010(9): 36-36.
- [11] Mistry Pranav, Pattie Maes. SixthSense: a Wearable Gestural Interface [C]. in ACM SIGGRAPH ASIA 2009 Sketches. 2009: ACM.
- [12] Norrie Lauren, Roderick Murray-Smith. Virtual Sensors: Rapid Prototyping of Ubiquitous Interaction with a Mobile Phone and a Kinect [C]. in Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services. 2011: ACM.
- [13] Campbell Andrew, Tanzeem Choudhury, Shaohan Hu, Hong Lu, Matthew K Mukerjee, Mashfiqui Rabbi, Rajeev DS Raizada. NeuroPhone: Brain-Mobile Phone Interface Using a Wireless EEG Headset [C]. in Proceedings of the second ACM SIGCOMM workshop on Networking, systems, and applications on mobile handhelds. 2010: ACM.
- [14] 吴朝晖, 潘纲. 中国计算机学会文集: 普适计算 [M]. 北京: 清华大学出版社.
- [15] Zhang Li, Gang Pan, Zhaohui Wu, Shijian Li, Cho-li Wang. SmartShadow: Modeling a User-Centric Mobile Virtual Space [C]. in Pervasive Computing and Communications, 2009. PerCom 2009. IEEE International Conference on. 2009: IEEE.
- [16] 徐光佑, 史元春, 谢伟凯. 普适计算 [J]. 计算机学报, 2003, 26(9): 1042-1050.
- [17] 於志文, 於志勇, 周兴社. 社会感知计算: 概念, 问题及其研究进展 [J]. 计算机学报, 2012, 35(1): 16-26.
- [18] 於志文, 周兴社, 郭斌. 移动社交网络中的感知计算模型、平台与实践 [J]. 中国计算机学会通信, 2012, 8(5): 15-21.
- [19] Lu-An Tang Yu Zheng, Jing Yuan, Jiawei Han, Alice Leung, Chih-Chieh Hung, Wen-Chih Peng. On Discovery of Traveling Companions from Streaming Trajectories [C]. in IEEE International conference on Data Engineering (ICDE). 2012. Washington, DC: IEEE.
- [20] 刘震, 秦永健, 王霞, 徐彬. 瞬时社交网络——从在线社交网络到移动社交网络 [J]. 中国计算机学会通信, 2012, 5.
- [21] 唐杰, 杨洋. 移动社交网络中的用户行为预测模型 [J]. 中国计算机学会通信, 2012, 5.
- [22] 刘云浩. 群智感知计算 [J]. 中国计算机学会通信, 2012, 8(10).
- [23] Yang Zheng, Chenshu Wu, Yunhao Liu. Locating in Fingerprint Space: Wireless Indoor Localization with Little Human Intervention [C]. in Proceedings of the 18th annual international conference on Mobile computing and networking. 2012: ACM.
- [24] Dragos Niculescu, Badri Nath. Ad Hoc Positioning System(APS) [C]. in Global Telecommunications Conference, GLOBECOM'01. IEEE. 2001. San Antonio: IEEE Press.
- [25] Gang Pan Guande Qi, Wangsheng Zhang, Laurence Yang, Shijian Li, Zhaohui Wu. Trace Analysis and Mining for Smart Cities: Issues, Methods, and Applications [J]. IEEE Communications Magazine, 2013, 51(6): 120-126.
- [26] Peiyuan Yuan, Huadong Ma. Hug: Human Gathering Point Based Routing for Opportunistic Networks [J]. in Wireless Communications and Networking Conference (WCNC). IEEE: Shanghai. 2012: 3024-3029.
- [27] 郭斌 张大庆, 於志文, 周兴社. 数字脚印与“社群智能” [J]. 中国计算机学会通信, 2011, 7(3): 53-60.
- [28] José Rui, Filipe Meneses, Adriano Moreira. Integrated Context Management for Multi-Domain Pervasive Environments. 2005.
- [29] Guo Bin, Daqing Zhang, Lin Sun, Zhiwen Yu, Xingshe Zhou. iCROSS: Toward a Scalable Infrastructure

- for Cross-Domain Context Management[J]. Personal and Ubiquitous Computing, 2013, 17(3): 591-602.
- [30] Han Jonghyun, Xing Xie, Wootack Woo. Context-Based Microblog Browsing for Mobile Users[J]. Journal of Ambient Intelligence and Smart Environments, 2013, 5(1): 89-104.
- [31] Guo Bin, Huilei He, Zhiwen Yu, Daqing Zhang, Xingshe Zhou. GroupMe: Supporting Group Formation with Mobile Sensing and Social Graph Mining[C]. in Proceedings of the 9th international conference on mobile and ubiquitous systems: computing, networking and services(MobiQuitous'12). 2012.
- [32] 胡海洋、李忠金,胡华,赵格华.面向移动社交网络的协作式内容分发机制[J].计算机学报,2013,36(3): 613-625.
- [33] Yohan Chon Nicholas D Lane, Fan Li, Hojung Cha, Feng Zhao. Automatically Characterizing Places with Opportunistic Crowdsensing Using Smartphones[C]. in Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012, ACM: New York. p. 481-490.
- [34] Stanford Vince. Using Pervasive Computing to Deliver Elder Care[J]. Pervasive Computing, IEEE, 2002, 1(1): 10-13.
- [35] 莫同,李伟平,吴中海,褚伟杰.一种情境感知服务系统框架[J].计算机学报,2010,33(11): 2084-2092.
- [36] Yan Tang, Shuangquan Wang, Yiqiang Chen, Zhenyu Chen; PPCare: A Personal and Pervasive Health Care System for the Elderly[C]. in Ubiquitous Intelligence & Computing and 9th International Conference on Autonomic & Trusted Computing(UIC/ATC). 2012: Fukuoka. 935-939.
- [37] Wu Jiahui, Gang Pan, Daqing Zhang, Shijian Li, Zhaojun Wu. MagicPhone: Pointing & Interacting [C]. in Proceedings of the 12th ACM international conference adjunct papers on Ubiquitous computing-Adjunct. 2010: ACM.
- [38] Du Yuan, Haoyi Ren, Gang Pan, Shijian Li. Tilt & touch: Mobile Phone for 3D Interaction[C]. in Proceedings of the 13th international conference on Ubiquitous computing. 2011: ACM.
- [39] Chen Longbiao, Gang Pan, Shijian Li. Touch-driven Interaction via an NFC-enabled Smartphone. in Pervasive Computing and Communications Workshops(PERCOM Workshops), 2012 IEEE International Conference on. 2012: IEEE.
- [40] Chen Yiqiang, Zhenyu Chen, Junfa Liu, Derek Hao Hu, Qiang Yang. Surrounding Context and Episode Awareness Using Dynamic Bluetooth Data. in Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012: ACM.
- [41] Crowley Katie, Aidan Sliney, Ian Pitt, Dave Murphy. Evaluating a Brain-Computer Interface to Categorise Human Emotional Response[C]. in Advanced Learning Technologies(ICALT), 2010 IEEE 10th International Conference on. 2010: IEEE.
- [42] Linden David E J, The P300: where in the brain is it produced and what does it tell us? The Neuroscientist, 2005, 11(6): 563-576.
- [43] 吴边,苏煜,张剑慧,李昕,张吉财,陈卫东,郑筱祥.基于P300电位的新型BCI中文输入虚拟键盘系统[J].电子学报,2009,37(8): 1434-1438.
- [44] Yu Yipeng, Dan He, Weidong Hua, Shijian Li, Yu Qi, Yueming Wang, Gang Pan. FlyingBuddy2: a Brain-Controlled Assistant for the Handicapped. in Proceedings of the 2012 ACM Conference on Ubiquitous Computing. 2012: ACM.
- [45] Yu Chun, Yue Shi, Xinliang Wang, Yuanchun Shi. Watching You Moving the Mouse, I know Who You Are. in Proceedings of the 2012 ACM Annual Conference Extended Abstracts on Human Factors in Computing Systems Extended Abstracts. 2012: ACM.

- [46] 樊银亭, 滕东兴, 杨海燕, 马翠霞, 戴国忠, 王宏安. 基于经验感知的自适应用户界面模型[J]. 计算机学报, 2011, 34(11): 2211-2223.

作者简介

李石坚 浙江大学计算机学院副教授, CCF 普适计算专委会委员。主要研究方向为普适计算、移动互联网、社会计算等。E-mail: shijianli@zju.edu.cn。



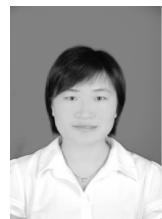
郭斌 博士, 西北工业大学计算机学院副教授。主要从事普适计算、群智感知、移动社交网络等领域研究工作。E-mail: guobin.keio@gmail.com。



於志文 教授, 博士生导师, 洪堡学者, 西北工业大学计算机学院副院长, 普适与智能计算研究所所长。主要从事普适计算、移动互联网、人机交互、个性化服务等领域研究工作。E-mail: zhiwenyu@nwpu.edu.cn。



赵莎 博士生, 浙江大学计算机学院, 主要从事普适计算方面研究工作。E-mail: szhao@zju.edu.cn。



吴朝晖 浙江大学计算机学院教授, 系统所所长。CCF 会士、常务理事、普适计算专委会主任。主要研究方向为人工智能、普适计算与服务计算等。E-mail: wzh@zju.edu.cn。



微处理器芯片可靠性设计研究进展

CCF 容错计算专业委员会

摘要

随着云计算时代的到来，人们对计算资源的渴望有增无减。在一些任务关键领域，例如银行交易数据库信息管理、航空航天、某些恶劣环境下的军用计算机系统等，都对芯片的可靠性有着很高的要求。即便在普通的 IT 领域，高可靠微处理器也是保证信息服务质量的基石。然而，随着微处理器芯片的复杂度和规模的不断增加、制造工艺的不断细化，保证芯片的高可靠变得越来越困难。其中的挑战包括高能中子、阿尔法粒子等放射性粒子导致的软差错（soft error），芯片运行时功耗变化带来的动态电压降导致的时序故障，以及芯片非正常老化导致的系统失效等问题。ITRS 预测表明，这些问题导致的故障率将逐渐上升。如何克服上述问题以提高处理器可靠性成为近年来容错计算领域研究的热点问题。国内在高可靠微处理器上的发展相对滞后，科研投入明显不足，至今没有类似 IBMZ 系列等高可靠芯片产品投入市场。本报告将综述国际上微处理器芯片可靠性设计的研究进展，重点分析现有先进的容错设计方法和微处理器可靠性评估体系，介绍我国在这一领域的现状和差距，并分析技术主要发展趋势。

关键词：高可靠微处理器，可靠性设计，瞬态故障，永久故障，可靠性评估

Abstract

The relentless quest for computing capacity keeps growing with the advent of Cloud Computing era. The reliability is an essential requirement in mission-critical fields, such as database management for bank transactions, aerospace and military computer systems in harsh environments. Even for civil IT industry, high reliable microprocessors are also fundamental components to ensuring quality of service. Unfortunately, ensuring the reliability of processors is getting increasingly challenging given the growing chip complexity and integrated density, and shrinking feature sizes. The culprits include soft error caused by high- energy radiations such as neutrons, alpha particles, timing faults resulted from dynamic voltage fluctuations, and system failures due to abnormal aging effects. According to the forecast of ITRS, the error rate will keep increasing; hence, it's unsurprising that how to tackle these challenges has become a research hot topic in fault tolerant computing community. In terms of R&D of high-reliable microprocessor, China has lagged behind, and also the investment to corresponding researches is far from sufficient. So far, China has not domestic IBM Z-series-like microprocessor product in the market. This report is devoted to survey the international research progress on design for reliability, with emphasizing on introducing state-of-the-art fault-tolerant design approaches and the systematic evaluation framework, presenting the current situation of China and the gap to the cutting-edge countries, and lastly analyzing the technology trend of design for reliability.

Keywords: High- reliable microprocessor, design for reliability, transient fault, permanent fault, reliability evaluation

1 引言

高可靠微处理器芯片是一类有特殊的使用场合的处理器芯片，这些场合包括航空航天电子、国防军工、环境恶劣的工业控制、银行系统等。以一般的工业控制机为例，其平均无故障时间（MTTF）一般要求在10万小时以上，平均修复时间（MTTR）在5分钟以内，可用性要求达到99.99999%的水平。在航天领域，例如同步卫星载微处理器系统，在轨时间约为八年，就意味着需要保证星载微处理器八年不失效，这对可靠性的要求极其苛刻。在民用领域，例如支撑云计算的数据中心，虽然对可靠性的要求不如工控和航空航天等领域那么高，但是提高微处理器的可靠性无疑是提升不间断服务的质量，同时降低硬件维护成本的有效途径。

然而，随着集成电路制造工艺的不断细化，晶体管和片上互连线的抗干扰能力随之下降。根据国际半导体技术发展报告（ITRS）的预测，在16纳米工艺下，由于软差错导致的失效率将会达到每片处理器芯片每天一次^[1]，这意味着MTTF的期望值仅为24小时。此外，芯片上由于辐射剂量累积效应、老化效应等，将会导致间歇性、甚至永久性故障的概率的上升。这对于可靠性设计提出了新的挑战。据最新的消息显示，美国空军研究实验室（AFRL）和国家航空航天局（NASA）于2013年3月发布了一个研究征询公告，目的是研究2020~2030年空间抗辐照多核处理器架构，用于空间载人和无人航天器，总投入预计超过2000万美元。在目前我国载人航天科学发展的重要阶段，研究高可靠微处理器芯片将有更加重要的意义。

本文将先总结现有的可靠性设计发展现状和可靠性评估体系，分析比较国内外研究现状，最后对未来可靠性的一些关键点进行技术展望。

2 国际可靠性设计研究现状

2.1 容芯片制造缺陷

正如人身体的病变大都从某些细胞开始，芯片的故障也大都可以追溯到某些晶体管的缺陷。这些缺陷包括生长多晶硅本身的晶格缺陷，制造过程中由于环境杂质导致的缺陷，氧化物绝缘层缺陷，掩模刻蚀中引入的形状缺陷，粒子掺杂中浓度漂移缺陷等。这些缺陷是限制芯片成品率的重要因素。

容忍缺陷最简单有效的办法就是冗余设计。由于缺陷出现的概率较低，若在设计之初就部署冗余资源，那么当目标单元由于缺陷而失效时，便可以利用冗余（无缺陷）的资源来替换缺陷单元，达到容缺陷的目的。这一思路已经广泛运用在大容量的存储器设

计中，例如计算机的主存芯片。而且，芯片内部还内建了检测失效单元的逻辑功能，根据检测结果，重新进行地址映射，排除缺陷单元即可。冗余所导致的芯片面积开销通常也可以保持在较低的水平。例如，研究表明针对现有 SRAM 型存储器的通常故障率，这些内建的冗余资源所导致的开销小于 5%，但可以显著提高芯片的成品率。

冗余设计在多核处理器设计中也有所应用。随着多核、众核处理器的兴起，单片处理芯片中某个核出现故障的概率增加了。IBM、Sony、Toshiba 公司于 2001 年启动，历经五年共同研制成功 Cell 多核处理器^[2]，其 8 个协处理单元和一个 PowerPC 处理器核由高速环形总线互连，由于芯片规模大、核数目较多，据报道成品率还不到 20%（芯片的成品率通常都是公司商业机密，根据成品率就可以大致推算出芯片的成本，所以极少公开发布）。Sun 公司的 UltraSPARC T1 处理器含有八个同构的 SPARC 处理器核，通过交叉开关互连。由于存在核的缺陷，该处理器被分级为两类：无故障核的 8 核芯片，用于 SUN 高端服务器 SUN Fire T2000；而存在一个或两个故障核的芯片降级为 6 核芯片，用于低端的 SUN Fire T1000 服务器。

随着制造工艺的进步，晶体管的特征尺寸不断细化，晶体管对这些缺陷也变得越来越敏感，其功能性已经不能简单地用“正常”、“失效”这样的二元判定标准来划分了。某些存在缺陷的晶体管在功能上是正常的，但可能开关速度有所下降，也可能更容易发生老化现象。这就带来了工艺偏差相关的可靠性问题^[4]。解决工艺偏差最“治本”的方法是改进工艺。例如，摒弃传统的二氧化硅、多晶硅等材料，采用高 K 材料、金属栅材料等新材料。采用更先进的过程控制技术可以实时发现制造过程中的扰动、参数漂移，迅速反馈，确保良率。

但仅凭改进工艺以期获得高可靠性是远远不够的，正如再强健的人也难免生病一样，通过了产品测试的芯片在运行过程中也会发生故障。例如，由于芯片环境温度的剧烈变化，或者供电系统电压波动带来的芯片时序变化，都会导致芯片运行时的故障。于是一系列偏差容忍的设计方法应运而生，形成了可靠性设计的一个细分研究领域，即 PVT 偏差优化^[6]。其中，“P”即指工艺偏差（process variation），“V”指电压波动，“T”指温度的变化。其中，P 是静态偏差，在芯片制造过程中引入，而 V 和 T 是动态偏差，影响芯片运行过程，极易导致芯片运行过程中的瞬态故障。可以通过协同优化的方法来克服 PVT 偏差造成的瞬态故障^[6]。

2.2 容芯片中的瞬态故障

现在的芯片虽然是个极其复杂的系统，但其基本的模型还是时序电路，遵守一定的时序规则。简单来说，就是内部的状态寄存器必须在“正确的时间”锁存到“正确的值”，并且将其保存一个（或多个）完整的时钟周期。这里“正确的时间”即是芯片中每个寄存器的时钟有效沿，“正确的值”就是在有效沿到来之前必须计算完成的信号值。其中任何一个环节出现问题，都会导致所谓的瞬态故障。

软差错是典型的瞬态故障之一^[7]。软差错会破坏保存在寄存器中的“正确的值”。

早在 1979 年就发现在动态存储器中存在软差错。软差错最初被定义为发生在存储单元中的随机、不重复出现、单比特反转的故障。但现在软差错的概念已经不局限于存储单元和单比特反转故障，而是扩展到了任何由于高能粒子导致的芯片内部随机性故障。诱发软差错的原因是地球低强度背景辐射下的中子和高能粒子引起的晶体管内部电荷重分布。这种改变虽然不会对芯片产生永久损坏，但有可能使受影响的电路结点产生差错数据并造成芯片的瞬态故障。随着晶体管的尺寸不断变小，每个晶体管本身对中子和高能粒子的影响更加敏感；同时芯片规模的指数级增长也意味着芯片上某一部分遭受一个软差错的影响的概率大幅提高。图 1 是研究人员根据统计得出的软差错率随工艺的变化趋势，并预测了未来的趋势^[1]。根据结果推测在 16 纳米工艺下，由于软错误率的上升，每天每片芯片中出现一个用户可见的失效将成为可能。

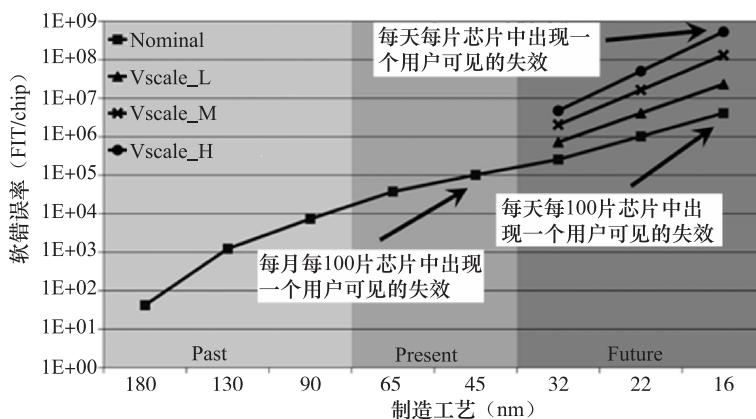


图 1 软差错率随特征工艺尺度的变化趋势^[1]。其中，Nominal 表示标称的工作电压，Vscale_L、Vscale_M、Vscale_H 分别代表可能采用从保守“L”到激进“H”的电压缩放级别的软差错率

软差错在组合电路与时序电路里的表现不同，普遍采用的故障模型是：在组合逻辑中，软差错的影响是产生一个窄脉冲，即单事件瞬态（Single Event Transient, SET）；在时序逻辑中，软差错的效应是导致存储元件位翻转（bit flip），即单事件翻转（Single Event Upset, SEU）。对于 SRAM 等结构，单个粒子还有可能导致多个位发生差错，即导致多事件翻转（Multiple Event Upset, MEU）。MEU 与 SEU 本质相同，但发生的概率目前来看要小得多。

电压的波动是导致瞬态故障的另一个主要原因。例如，瞬态电压下降过大将会减慢晶体管的开关速度，宏观来看就是电路变慢了，这会导致正确的值不能在规定时间内计算完成，从而导致故障。电压的波动主要来源于两种效应：IR 降（IR-drop）和感应噪声（inductive noise）^[8]。由于电压对（动态）功耗的“平方”贡献关系，降低电压被公认为是最有效的降低芯片功耗的手段之一，然而，随之而来的负面效应就是噪声容限的降低。当前我们已经面临“电压墙”（voltage wall）的困境，芯片的供电电压已经降到了 1 伏左右，在现有的 CMOS 技术下继续降低的空间有限；相反，芯片的总功耗却在不断上升。这两个趋势带来的直接效应就是流经芯片的电流密度不断增大，即使较小的寄生电阻都

有可能导致较大的 IR 降损失。同时，感应噪声的影响正比于电流的变化率，芯片功耗持续上升，以及芯片内部功耗门控、时钟门控（power gating, clock gating）等低功耗技术的广泛采用，都对电流的剧烈变化起到推波助澜的作用，从而造成输送到芯片上电压的波动。

抑制电压波动（V 分量）的方案主要有两类：1) 减小供电网络的阻抗和感抗；2) 抑制电流的变化率。例如，现有的大规模芯片广泛利用倒装芯片（flip chip）的封装方式来减小电感和平衡阻抗，片内利用银做导线材料来减小电阻。在体系结构级，利用步进控制的方式来逐步唤醒某些功耗较大的微结构来抑制剧烈的电流变化，达到缓和电压噪声的目的。

瞬态故障通常都具有较强的随机性和不可预测性，最简单有效的方法就是“重做”。就如我们在个人电脑出现“死机”等状况时，首先想到的就是“重启”。然而，在很多场合是没有“重启”的条件的，例如不容许重启带来的时间开销或数据丢失。如何容忍瞬态故障也是可靠性设计的主要挑战之一。

传统针对软差错的优化大多建立在 SEU 和 SET 故障模型的基础上。例如，Mitra 等人在电路级提出了一种具备自检测能力的触发器设计^[9]，其基本思想是利用现有的扫描触发器本身的冗余资源来检测 SEU，即主触发器和扫描触发器构成双模冗余，扫描部分的触发器在功能状态下也处于工作状态。这一方案可以很容易地扩展成具备一定 SET 故障检测能力的触发器。还有一类基于校验码（ECC）的电路及方案。该类方案已经成功运用到 IBM Power4、Power5 系列处理器中用于缓存系统的保护^[10,11]。但是，基于校验码的方案对于电路中的非规则结构并不适用。

在高可靠处理器容错设计领域，尤其是对高性能通用处理器的容错设计，一些研究人员提出线程级冗余来进行容错^[12]。为了降低设计的硬件开销，这类容错技术多以同时多线程处理器（SMT）和片上多核处理器（CMP）为基础，从而可以利用已有的片上资源来进行容错。这类方案的核心思想可以概括为：输入复制，冗余执行，输出比较。SMT 技术使得多个线程可以同时共享处理器硬件资源，而只有很少的线程切换开销。这为利用线程级冗余提供可靠性保证的方案提供了硬件条件。然而，SMT 技术是以复杂的处理器前端设计为代价，带来了功耗效率下降，热量管理困难等问题。同时，由于原线程与冗余线程的“相关”性很高，容易导致资源竞争的问题。这类方案的最大的缺点就是吞吐量的显著下降和冗余线程带来的功耗损耗。从计算有效性来看，计算的带宽几乎损失一半，因为每一个线程都需要一个执行核、一个校验核。

Wang 等人于 2005 年提出基于猜测（speculation）的可靠性设计思想^[13]。Reddy 等人于次年结合指令部分冗余（partial redundancy）思想改进了该设计^[14]，其原型系统运用于具有较强的猜测执行能力的高性能超标量处理器。这类可靠性设计能有效工作的前提是：处理器自身具备相当高的猜测准确率，比如分支预测。这使得我们可以以较高的置信度做出如下推测：导致某次猜测失败的最可能的原因是瞬态故障，而不是硬件猜测本身的局限性。虽然这类可靠性设计的硬件开销几乎为零，但故障覆盖率很难超过 90%，而且会出现显著的误检测率。其次，还有一大批处理器（特别是现有的嵌入式处理器）并不具备强健的硬件所支持的猜测执行（speculative execution）能力，例如 LEON、ARM 系列处理器。这些局限性极大限制了这类基于猜测的方案的应用场合。

最后一类策略是基于软件多模冗余故障检测及容忍。这类方案的基本思想是在编译阶段就考虑实现指令集的双模或三模冗余^[15]。优点是不需要硬件的修改，缺点是均有超过60%的性能开销、加倍的二进制代码容量，以及应用时需要对原有的代码重新编译。

2.3 容芯片中的永久故障

不同于瞬态故障，芯片中的永久故障效应是可重现的。通常认为当芯片出现永久故障意味着芯片已经“寿终正寝”，但随着多核处理器的发展，单芯片核数目不断上升。例如，Tilera公司的TILE-Gx系列的众核芯片包含72个处理器核；IBM最新的高端服务器芯片Power7包含8个核；正在开发的国产的龙芯三号一款处理器芯片也包含高达16核^[16]。因此，某一个核出现永久故障的概率也随之增加。因为某一个核出现故障而牺牲整片处理器芯片是一种极大的浪费，所以将故障核屏蔽，允许芯片降级使用是容忍永久故障的一个有效途径。

自检测的概念在存储器中使用比较普遍，称为存储器“内建自测试（BIST）”，即在设计存储器芯片时就将测试模块加入（“内建”），在出厂前启动内建自测试。根据预先设定的测试流程来检测所有存储单元功能是否正常。如果不正常，还可以根据故障位图来进行诊断，通过将故障单元地址重映射到预留的冗余（无故障）单元来屏蔽故障，达到修复的目的。

然而，BIST对于复杂的处理器芯片就显得力不从心了。首先，与存储器的规则结构相比，处理器内部的结构极不规则，相应的测试数据就会大得多，一般情况下必须借助外部的测试仪才能将大量的数据施加到芯片上。其次，要进行充分的测试，测试数据本身必须通过复杂的EDA软件来生成，片上伪随机生成的测试数据很难达到高故障覆盖率。最后，即便检测并定位到故障点，也很难通过简单的冗余替换策略来修复。所以，针对处理器芯片永久故障的容错设计通常是在更粗的粒度上开展的，例如以处理器核为基本单位。

2.4 国际高可靠微处理器实例：HP NonStop系统

容错计算是一项系统工程，并不是通过一片可靠的微处理器芯片或者一套高可靠软件系统就能实现。现有的较为著名的容错计算机包括惠普公司的NonStop系列计算机^[19]，IBM公司的Z系列高可靠计算机等。这些高可靠计算机系统都是硬件、软件协同可靠性设计的典型范例。

以NonStop系统结构为例，系统最大的特征就是硬件提供冗余（双模或三模）、软件管理冗余。图2所示即为配置为三模冗余的4路处理器NonStop系统结构，每个处理器都具备自检测功能（Self-Checked），并独享存储，处理器之间通过系统区域网络（System Area Networks，SAN）互连。为了保证可靠性，SAN都有一套备份。所有的存储接口单元、IO单元都有双模冗余。所以，在NonStop体系结构下，单故障不会“停止（Stop）”系统的服务，这也许就是NonStop名字的由来。然而，NonStop系统并不仅要做到硬件冗余，如何高效管理这些冗余资源来实现高可用的服务才是关键。

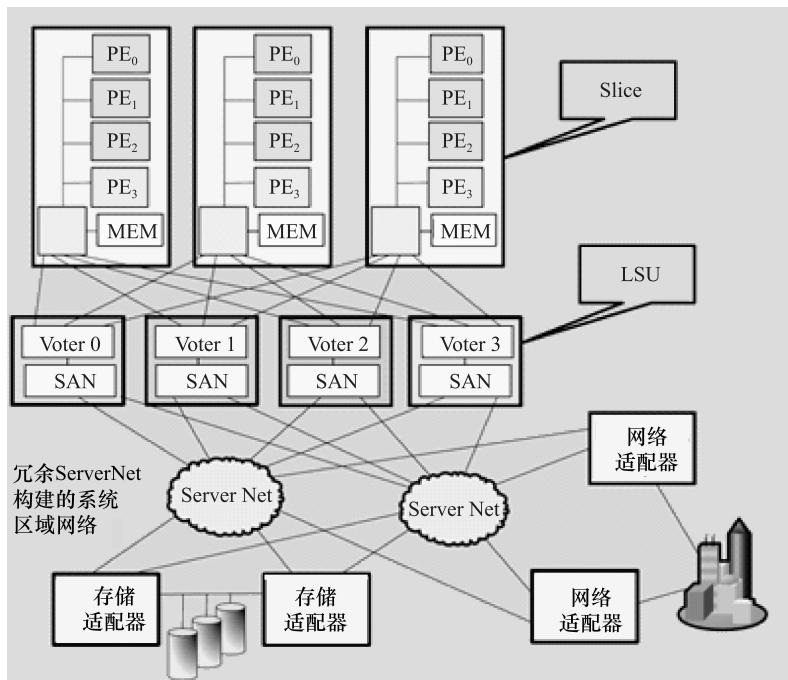


图 2 配置为三模冗余的 4 路处理器 NONSTOP 系统结构^[19]

HP 为 NonStop 计算机开发了 NonStop Kernel 作为操作系统。该系统功能上与传统的操作系统类似，包括内存管理、进程管理等，但实现上有一套复杂的消息传递系统和失效恢复机制。所有的关键系统软件实现为由一个主进程和一个备份进程组成的进程对。这些进程分布在不同的处理器上，防止某个处理器失效导致整个进程崩溃。主进程和备份进程通过处理器间的消息进行通信，如果主进程崩溃会触发一条消息，通知备份进程接管剩下的工作，并且会重新启动一个备份进程实现主 - 备份进程对。

处理器间通信通过双模冗余的 SAN，HP 称之为 ServerNet，是一套基于包交换的高速、低延迟的互连网络。每个处理器至少通过两套 SAN 互连，所以，任何处理器间网络的故障都不会导致互连的失效。此外，NonStop Kernel 之上还建立了一套分布式事务处理的关系数据库，与 Kernel 紧密耦合来保护 NonStop 中的数据完整性。但这些复杂的系统都是对用户透明的。对于应用程序而言，NonStop 与普通的服务器没有区别。这也是 NonStop 受到欢迎的主要原因之一。

3 国内可靠性设计研究现状

国内在芯片可靠性设计方面也有一些创新性的方法和创新点。在核级冗余设计方面，中科院计算所在 973 项目的支持下提出一套设置冗余核从而有效提升芯片级别的方法，我们称之为 $N + M$ 方法，即标称 N 个核的处理器，额外增加了 M 个冗余核作为备份。通过改变互连就可以屏蔽故障核，将芯片降级使用；或者通过互连拓扑的重构利用冗余的

核来替换故障核、达到修复的目的^[17,18]。与存储器类似，当出现故障核时，利用冗余（无故障）核替换，达到尽可能提升芯片的性能级别的目的。与存储器不同的是，核间的互连要比存储器阵列间的互连复杂得多。做核间替换后，处理器互连的拓扑结构极有可能发生较大的变化，例如原先相邻的核变得不再临近，造成核间通信的性能降低。所以，如何放置冗余核的物理位置，选择哪一个冗余核替换故障核都是需要权衡的。这就是 $N+M$ 方法中的拓扑重构问题^[3]。如图 3 所示，对于一个目标是 3×3 的多核处理器（如图 3a 所示），可以预留一列处理器，即实际制造 3×4 的多核处理器（如图 3b 所示）。当发现有故障核时，如图 3c 所示，我们可以利用冗余核来替换故障核。

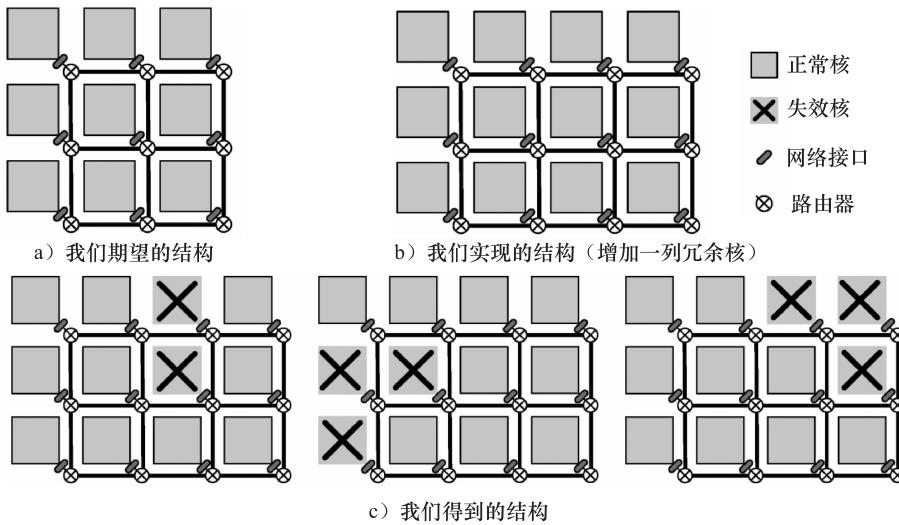


图 3 基于核级冗余的多核处理器故障修复^[18]

这一方法既可以应用在芯片制造的环节提升芯片的良率，也可以动态地重构来容忍芯片在使用过程中由于老化等原因^[5]出现的永久故障。但是，与缺陷容忍不同，容忍服役期出现的永久故障要求芯片本身有自检测、自诊断、自修复的能力。中科院计算所提出了一套集自检测、自诊断、自修复“三位一体”的可靠性保障体系，称之为“3S”技术，该技术处于国际领先水平。

在容瞬态错误方面，中科院计算所国家重点实验室与北京控制工程研究所和浙江大学合作完成的成果“星载微处理器系统验证 - 测试 - 恢复技术及应用”荣获 2012 年度国家技术发明奖二等奖。此外，国防科技大学、中科院微电子研究所、同济大学等单位在处理器的抗辐射设计技术、航空可靠处理器容错设计等方面也有贡献。

在容错计算机系统方面，2011 年，在我国科技部“十一五”863 计划“高端容错计算机”重大专项支持下，浪潮集团成功研制的天梭高端容错服务器，可用性达到 99.999%。

此外，2007 年，在国家自然科学基金“半导体集成化芯片系统基础研究”重大研究计划重点项目“微处理器自修复设计基础技术研究”支持下，中科院计算所成功研制的自修复处理器 SRP，能够对片内 SRAM 上不多于 10 个固定型故障达到 100% 修复率。

SRP 处理器的最大特征是通过在片上存储器中增加内建自测试（Built-In Self-Test，

BIST)、内建自诊断 (Built-In Self-Diagnosis, BISD) 和内建自修复 (Built-In Self-Repair, BISR) 逻辑 (简称 3S 结构)，使之具备故障检测、定位和修复能力，从而仅以较低的硬件开销即可显著提高处理器的成品率和可靠性^[20]。自测试模块一旦检测到故障则启动自诊断功能，由自诊断模块定位故障地址和判定故障类型，并将故障信息输入自修复模块。由于芯片内部设计了专门的冗余位和冗余字，一旦自诊断完成，自修复将根据故障信息采取优化的修复策略，将访存故障存储单元的操作重定向到冗余资源，从而实现自修复。利用内容可寻址技术结合冗余位的自修复方法具有修复率高、面积开销小以及实现简单等优点。图 4 显示了 SRP 处理器的版图和封装后的芯片。

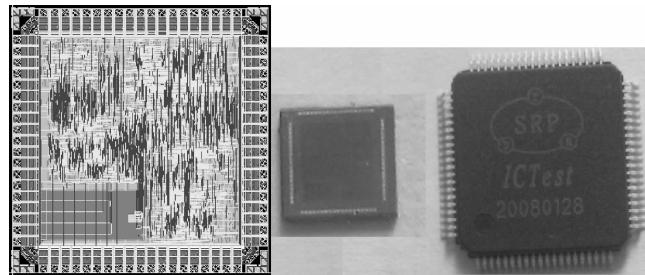


图 4 基于内容可寻址技术的自修复处理器 SRP：版图与封装后的芯片

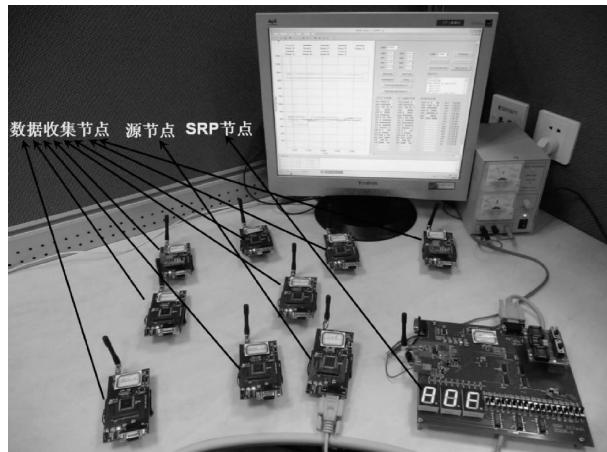


图 5 采用自修复处理器 SRP 搭建的无线传感网络结点

4 可靠性评估体系

4.1 可靠性评估概述

传统上对集成电路可靠性问题的研究主要是面向芯片制造过程而展开的。所关注的模型有属于失效物理模型的缺陷分布模型、有反映软（硬）差错影响的可靠性模型，也

有与可靠性相关的成品率计算模型。在纳米工艺下，随着集成度的提高，除缺陷之外，参数扰动也严重地影响到电路功能及性能的实现，导致建立可靠性解析模型变得越来越困难。而且，相关技术是针对电路成品的，难以直接帮助设计人员在电路设计（特别是高层设计）阶段改善电路的可靠性。

从 20 世纪 90 年代开始，人们认识到，在电路高层（含门级及以上抽象层次，如系统级（包括体系结构级和微体体系结构级））分析软差错的影响有助于体系结构设计师提出高可靠的结构设计方案。因为单粒子翻转对微处理器设计带来极大的挑战，为采取节省开销的技术应对瞬时故障，设计者需要准确估计处理器差错率，以评估容错技术的效果，或找出处理器结构中的薄弱环节，从而制定兼顾开销和可靠性的措施。

4.2 电气级可靠性估计

4.2.1 电气级模拟故障注入方法

电气级的软差错分析多采用模拟故障注入方法。其工作过程如下：首先分析故障机制，建立电气级故障模型；再使用 SPICE 等工具模拟电路，并依据故障模型在电路内部结点上注入故障，模拟由粒子撞击而形成的瞬时脉冲，即 SET；然后通过瞬时脉冲在敏化通路上的传播，分析 SET 的削弱程度，并结合触发器上升沿与下降沿，分析锁存窗屏蔽效果；最后，推导 SEU 的发生概率。

电气级故障模型主要包括电流模型和电压模型。由于存在结间势差，粒子撞击漏极区产生的电子空穴对聚集在漏极和衬底之间，这些电子空穴对产生电流脉冲。电流脉冲宽度取决于最大电流、结电荷收集时间常数和离子踪迹建立时间常数等几个技术参数，常用双指数电流脉冲模型来描述电流脉冲，如式（1）所示：

$$I(t) = I_0(e^{\alpha t} - e^{\beta t}) \quad (1)$$

其中， I_0 表示最大电流， $1/\alpha$ 表示结电荷收集时间常数， $1/\beta$ 表示离子踪迹建立时间常数。双指数电流脉冲模型主要针对 alpha 粒子对 PMOS 器件的撞击进行建模，在电路级模拟粒子撞击产生的电气影响时通常采用该模型。

电压模型有矩形脉冲模型、梯形脉冲模型和基于威布尔概率函数的电压模型。为了简化模拟过程，把电压脉冲转化为一个或多个矩形脉冲，矩形脉冲的宽度取决于瞬时脉冲电压超过逻辑阈值的部分。瞬时脉冲传播后，波形用梯形表征更准确，因此，用梯形脉冲模型来描述电压脉冲。但在瞬时脉冲刚产生时，并不是标准梯形，于是用基于威布尔概率函数的经验模型拟合电压变化。电压脉冲可用 SPICE 等工具进行模拟，因此电气级故障模型常用电压模型。

要使注入的故障能在模拟的电路中准确地传播，必须提供准确和完整的时序信息。模拟故障传播过程中使用标准延迟格式文件提供时序信息，并分别在电路的组合逻辑与时序单元中进行模拟实验。

电气级模拟软故障注入的自动化方法^[21]使用 LL(k) 语法分析技术自动解析 SPICE 网表以自动提取故障注入目标的信息，由程序将网表自动展平，克服了多次例化的子电

路无法注入单故障的问题，通过对展平后网表的遍历自动插入表征软差错的电流源。

另外一类基于故障注入的方法是通过分析电荷收集量与关键电荷值的关系，以变化的电荷量为主要参数代入解析模型，直接计算软差错率（Soft Error Rate，SER）。在解析式中，软差错率与粒子流强度、撞击结点横截面积、电荷收集效率呈正比，与关键电荷值呈反比。解析方法一般不考虑逻辑屏蔽作用，所以有一定的局限性。

在电路级，通过故障注入来分析 SER 已经有了许多方法和工具。例如，模拟故障传播的 SERA^[22]、SEAT-LA^[23]，针对存储单元的 ROBAN^[24]、DASIE 方法^[25]、SMC-DASIE^[26]等。

以上通过模拟电路内部属性来分析故障影响的方法属于静态因素分析，在此基础上，还有学者考虑了环境等外界条件影响，即结合了动态因素进行分析。乔治亚理工学院集成 Cadence Spectre 模拟器开发了可靠性分析工具，用以分析电迁移和热载流子效应影响下芯片的可靠性^[27]。Ramakrishnan 基于 SEAT-LA 工具分析了不同电源电压、工作温度、热载流子注入、负偏置温度不稳定性对组合电路 SER 的影响^[28]。

4.2.2 电气级模型解析法

噪声影响分析考虑了 SET 的产生、传播、被锁存，以及时序电路的三种屏蔽作用。如图 6 所示，噪声影响分析分为四个部分：①“瞬态噪声建模”用矩阵表示噪声概率密度函数（Noise Probability Density Function，NPDF），以概率形式描述 SET 的发生。②在“电路噪声免疫度分析”中量化了逻辑屏蔽、时序屏蔽和电气屏蔽，这些屏蔽作用有助于防止瞬态噪声成为可观测的错误。③通过“NPDF 转换”过程反映噪声免疫对 SET 分布和传播的影响。在转化过程中，电路被表示为矩阵网表，屏蔽效应表示为 NPDF 矩阵的操作。④最后，当转换到达 D 触发器时，根据输入的 NPDF 计算“噪声捕获率”，通过噪声捕获率可以衡量 SET 错误率。

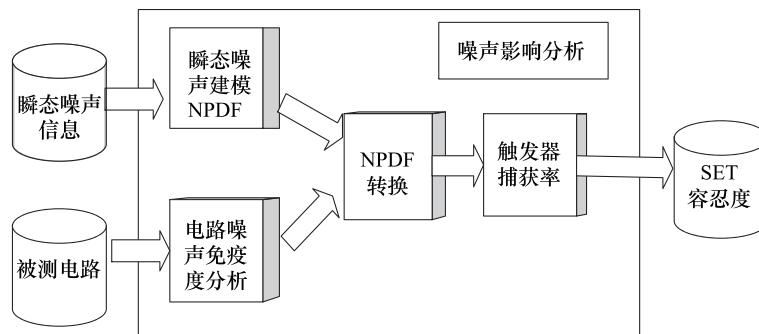


图 6 噪声影响分析

4.3 门级可靠性估计

4.3.1 门级模拟故障注入方法

门级模拟故障注入方法首先使用 SPICE 模拟电路，依据双指数电流脉冲模型或者其

他电流模型注入故障；在电路级收集瞬时脉冲的产生以及传播中衰减的信息，用以分析电气屏蔽；然后通过建立逻辑级模拟器描述逻辑屏蔽；最后依次考虑电气屏蔽、逻辑屏蔽和锁存窗屏蔽，计算锁存概率。

瞬时故障仿真器^[29]（FAult Simulator for Transients，FAST）提供门级瞬时故障模拟环境，它源自于真实的故障模型。FAST 使用时序故障模拟器对 SET 建模，同时用零延迟并行故障模拟器跟踪位错误，从而大幅度降低了时间开销。根据输出错误数占被锁存的故障的百分比计算锁存率。实验发现，多位锁存错误出现的几率很小，这说明单位翻转是锁存差错的主要表现形式。

门级工具还有 ASERTA^[30,31]、软差错自动分析仪（ASSA）^[32]、组合逻辑 SER 分析工具^[24]，这些工具的复杂度不高，但由于依据的是一定假设条件下的简化模型，损失了精确性。因此，后面的研究中提出了一些兼顾准确性与速度的方法。SonzaReorda 采取了将故障传播过程中时间信息折合成结构信息的方法^[33]，即扩充电路的结构，把不同时间间隔后的同一个门看做不同门，作为扩展部分重新模拟电路结构，从而实施零延迟的故障模拟而不损失准确性。在 ISCAS85 基准电路上的实验表明，与 VHDL 的时序模拟器实现方法相比，注入同样多故障时，其速度可以提高 5000 倍，而输出的差错相同。FASER 工具^[34]在门级运用 BDD（Binary Decision Diagrams）结构对故障传播建模，并叠加电路级的脉冲信息，通过电路分割减小复杂度。实验表明，FASER 方法的实施速度是基于 SPICE 模拟方法的 9000 倍，而误差为 12%。

SERA 方法结合物理层因素推导了高层抽象模型^[35]，运用图论和故障模拟方法分析逻辑屏蔽机制，用 HSPICE 模拟分析电气屏蔽和锁存窗屏蔽作用，与蒙特卡罗模拟方法相比，其加速比达到 105 的数量级，而误差损失不到 5%。

Mohanram 在分析 SER 时给出了三个因素的解析模型，计算组合电路中的软差错从发生结点 n 传播到锁存器 l 的概率^[36]，即 $Rseu(n) \times Psensitized(n, l) \times Platched(n, l)$ ，其中， $Rseu(n)$ 表示在结点 n 发生 SET 改变逻辑值的概率， $Psensitized(n, l)$ 表示从结点 n 到 l 存在敏化通路的概率， $Platched(n, l)$ 表示 n 点差错被 l 锁存的概率。

Nieuwland 面向组合电路，分析了三个指标：门的 SER(Gser)、毛刺可观察度 Glitobserv、电气屏蔽因子 EMF^[37]。其中，Gser 通过输入向量出现的概率乘以这种输入组合下易感性系数得到；Glitobserv 表示毛刺在没被其他门逻辑屏蔽的情况下传播到原始输出的概率，可以使用软件工具后向遍历算得；EMF 表示一个毛刺从产生位置到原始输出的传播通路上不被其他门削弱的概率，它主要依赖于通路长度及历经门的类型。这样，门失效率为 $FRgate = Gser * Glitobserv * EMF$ ，整个电路的软差错率 SERcircuit 是所有门 FRgate 之和。其中，FRgate 的取值依赖于不同输入向量出现的概率、门的类型、电路的结构等。FRgate 越高，表示这个门对电路 SER 的影响越大。SERcircuit 可以用来比较容忍软差错的不同设计。

时序电路的蒙特卡洛可靠度分析方法^[38,39]首先随机产生 nSamples 个输入矢量，加载到时序电路的原始输入端，然后根据电路差错概率 p 随机地在电路内部结点注入故障；注入故障后，SEU 在电路的敏化通路中传播；最后，在电路的原始输出端统计输出结果

出错的个数 n_{Errors} , 根据 $R = 1 - n_{Errors}/n_{Samples}$ 计算电路的可靠度。 n_{Sample} 越大, 可靠度评估结果越精确。如果电路原始输入端数为 n , 内部结点个数为 m , 那么 $n_{Samples} \geq 2^n \times m$, 则该类方法时间复杂度为 $O(2^n \times m)$ 。该方法比较精确, 但时间开销大。

4.3.2 门级模型解析法

门级模型解析法是通过计算软差错从发生结点传播到电路原始输出的概率来衡量失效率, 考虑电路的拓扑结构和传播路径信息, 与电路中各个门的类型和连接方式来分析电路的可靠度。门级概率评估源于信号可靠度方法, 信号可靠度是 20 世纪 60 至 70 年代提出的一种度量电路输出正确性的指标, 电路中一个线端的信号可靠度即为这根线上逻辑信号正确的概率^[40]。信号可靠度的推导中借助了信号概率概念及计算方法。比较典型的方法有计算单个输出结点 SER 的 TP 算法^[41]、通过计算差错传播率表征电路 SER 的 EPP (Error Propagation Probability) 方法^[42], 以及通过 PTM 模型计算整个电路可靠度的 PTM 方法^[43]。

PTM 模型能精确评估门级电路的可靠性。它假设所有基本门电路的故障概率是相同的一个常数, 并把电路作为一个不可分割的整体来对待。然而, 对于规模较大的电路, PTM 方法的时空开销随电路引线数的增加呈指数级增长。

之前所实现的 PTM 方法只适用于较小规模的电路。文献[44]引入了电路划分的思想, 先把电路分割成一组适宜用原始 PTM 方法直接计算其可靠度的模块, 然后计算出这些模块的可靠度, 再依据串行可靠度模型, 将所有模块可靠度合成为整个电路的可靠度。而且, 用实验的方法得到了合适的电路分割参数, 即分割宽度, 改进的方法可以用于更大规模的无冗余组合电路。

经过上述改进的 PTM 模型在估计电路可靠性时存在精度不足的问题, 文献[45]通过对电路拓扑结构的分析, 提出了宏门的概念。其思路是把一个电路划分成一定数量的宏门, 再以宏门为单位, 采用迭代的 PTM 方法计算从原始输入到电路任意引线位置的部分电路的可靠度。该方法的复杂度与宏门的数目呈线性关系, 具有较高的精度。

PGM (Probabilistic Gate Model) 模型以门为单位, 通过迭代方法求得输出为 1 的概率作为电路的可靠度^[46]。若电路中不带有扇出分支, 该模型可用于较大规模电路的可靠度评估, 否则, 由于采用多路径算法^[47]来计算扇出分支结构的可靠度, 使得其时空复杂度也随着扇出分支数的增加而呈现指数级增长的趋势。该方法未完全考虑电路结构对故障所具有的屏蔽特性以及忽略了输出为 0 的概率对可靠性的影响, 使计算精度受到了一定程度的影响。基于离散时间马尔可夫链的概率模型检验 (Probabilistic Model Checking, PMC) 方法^[48]与 PGM 方法高度相似, 因此它们存在着相同的问题。

文献[49]中给出的方法结合 WAA (Weighted Averaging Algorithm) 方法^[50]和关联系数方法^[42], 考虑了电路的扇出分支, 使可靠度的计算精度得到提高。但是, 对于每一个基本门的输入概率分布, 该方法无论是采用模拟方法, 还是 BDD 方法都有较大的计算代价, 且互连线关联系数的估计复杂度较大, 而且对于扇出分支的关联系数的估计是近似的, 因此该模型计算结果的精度还有进一步改善的空间。

只定位在门级估计 SER 的方法一般面向组合电路或者时序电路的组合逻辑部分。假定门的差错感染率是已知的，计算电路原始输出或者触发器输入出错的概率，用于分析逻辑屏蔽的影响，属于局部分析方法。加上电路级的模拟或者统计信息，便可以分析电气屏蔽、锁存窗屏蔽，从而更加完整地估计 SER，如文献[51]所述将电气屏蔽因素结合到 EPP 方法当中。

将 EPP 方法扩展，就形成了时序差错传播概率分析方法^[52]，可用于评估时序单元（触发器、锁存器和存储单元）受粒子撞击时的差错锁存概率。假定轰击只能发生在时序单元中，触发器被轰击后，在其输出端统计差错概率，没有考虑电路内部结点受到轰击的情况，也没有考虑逻辑屏蔽的影响。

组合电路可靠度估计的 SP 方法的基本思想是：根据电路拓扑结构，在一个阶段内计算出每个门的差错分量和传播差错分量的累积效应，最后计算出电路的可靠度。如果电路不存在重汇聚扇出，该方法可以精确地计算出电路的可靠度。SP 算法采用相关系数处理存在重汇聚扇出的电路。

时序电路可靠度估计的 MP 方法^[53]是通过电路的迭代方式评估时序电路的可靠度，是 SP 方法的扩展。该方法首先对时序电路近似层次划分，确定电路反馈的位置；然后，把电路转换为组合电路和反馈电路，组合电路部分用 SP 方法计算每个结点的可靠度，反馈电路部分用式（2）计算反馈信号对时序电路的影响；最后用迭代方式计算电路可靠度。其中梯度因子（gradient factor, gr）为反馈信号对电路可靠度的影响程度，gr 的取值范围为 0~1，gr=0 表示反馈信号对电路的可靠度没有影响，gr=1 表示反馈信号对可靠度影响最大，实验中一般使 gr=0.5。

$$P_{pr}(b_PI_i) = P_{new}(b_PI_i) + (P_{new}(b_node_i) - P_{new}(b_PI_i)) \times gr \quad (2)$$

MP 方法的不足之处在于：①假设反馈信号对时序电路的影响为一个梯度因子，而且 gr 的取值均为 0.5；②对于多输入门，把它拆分为多个两输入和单输入的门进行可靠度计算，这样改变了电路的原始拓扑结构；③时间复杂度为 $O(4^m)$ ，m 为重汇聚扇出源的数目。

基于贝叶斯网络（BN）的电路可靠性评估模型^[54]的理论依据是条件概率，其本质与 PTM 模型类似，故能准确评估电路的可靠性。对于规模较大的电路，该方法借鉴蒙特卡洛方法的思想，利用两个优化的随机抽样算法来统计电路发生故障的频率，得到的结果作为电路的失效率，也是一种近似的电路可靠性评估方法。

时序电路存在反馈，因而不能表示成标准的 BN，只能用动态 BN 来描述电路的差错模型。基于 BN 的时序电路可靠度分析方法把时序电路建模为动态 BN 模型^[55,56]，用 BN 分析工具推理计算电路的平均差错概率，再计算电路的可靠度。

BN 方法用 BN 分析工具（如 Hugin Expert^[57]、GeNIE^[58]等）进行 BN 推理，计算电路的平均差错概率，最后计算电路的可靠度。BN 推理算法分为精确推理算法和近似推理算法。对于特定拓扑结构的网络，其复杂性取决于结点数。因此，精确推理算法一般用于规模较小的时序电路。对于中大规模电路，一般采用近似算法进行推理，它可以简化推理和计算过程，但不能够提供精确概率值。

基于时序条件概率矩阵（Sequential Conditional Probability Matrix, SCPM）的时序电路可靠度分析方法把时序电路转化为组合电路，并按迭代方式分析电路的可靠度。SCPM 方法用时序条件概率矩阵表示每个逻辑门的信号概率，并用条件概率表处理重汇聚扇出^[59]。使用条件概率可以在保持可靠度估计准确性的同时解决可靠度评估的可扩展性问题。

多周期软差错模型（Multi-Cycle Soft error Modeling, MSEM）是用于分析时序电路 SER 的方法^[60]，并能够识别时序电路中最脆弱的门和触发器。MSEM 方法首先计算从差错发生点到原始输出端的差错传播概率。然后运行多周期差错传播分析程序估计电路的 SER。该方法无需进行耗时的故障注入或逻辑模拟，通过计算在多个时钟周期内组合电路和时序电路的差错传播概率来估计时序电路的 SER。

触发器是时序电路的重要组成部分，文献[61]提出了一种基于 PTM 的触发器可靠度估计（reliability estimation of Flip-flops based on Probabilistic Transfer Matrix, F-PTM）方法。F-PTM 方法首先从触发器电路的特征方程出发，再用电路 PTM 判定定理推导计算触发器电路的 PTM，最后，根据输入信号的概率分布函数计算出触发器的可靠度。F-PTM 方法既能计算组合电路的 PTM，也能计算带有反馈的电路的 PTM，因此可扩展到时序电路的可靠度评估。

文献[62]提出了基于 PTM 的时序电路可靠度估计方法（reliability estimation of Sequential circuits based on PTM, S-PTM），先把待评估的时序电路划分为输出逻辑模块和次态逻辑模块，然后用时序电路 PTM 计算模型得到电路的 PTM，最后根据输入信号的概率分布计算出时序电路的可靠度。

门电路的故障发生概率通常是以经验值或人为给定的方式出现，考虑到门级电路的结构、工艺技术、应用环境及所用材料等都是影响其故障水平的重要因素，文献[63 ~ 65]首先分析了缺陷影响下的互连导线的故障情况，并构建了相关的故障概率模型，把其计算结果作为基本门的失效率。由于栅氧化层的故障也是影响电路可靠性水平的重要因素，于是提出了广义门的概念，它包括输入导线和（或）基本门；考虑到广义门电路的故障概率与其类型、输入负载和时间等因素的关系，分别选取了基于弱链接模型的双峰对数正态分布和 Weibull 分布作为输入导线和栅氧化层失效的时间分布模型，并构建了考虑输入负载的各广义门电路故障概率的时间分布模型。

4.4 系统级可靠性估计

4.4.1 系统级模拟故障注入方法

系统级模拟故障注入首先选取合适的故障模型，同时建立集成电路的系统级模型，并用硬件描述语言实现，然后在模拟系统中嵌入故障注入机制。向系统级模型注入故障可以方便地控制故障的注入时间和注入位置，能够直接观测故障在模型中的行为，但要仔细平衡抽象层次、模拟速度以及准确度三者之间的关系。

在系统级分析软差错影响多采用位翻转故障模型，使用 VHDL/Verilog 描述电路，注入故障并运行一组典型工作负载，观察系统运行情况，收集数据，分析其可靠度。

IBM 公司用软差错蒙特卡洛建模（Soft-Error Monte Carlo Modeling, SEMM）方法分析器件的 SER^[66]。分析 SER 时，需要提供芯片详细的版图、工艺和临界电荷等信息。SEMM 方法将芯片划分成 3D 长方体网格，每个小长方体代表芯片中的某一个区域，具有一定的物理性质。然后采用蒙特卡罗随机模拟的方法对每个小长方体进行研究，利用辐射源能谱等特性随机生成一系列粒子，分析这些随机生成的粒子如何在网格中生成电子-空穴对，这些电子-空穴对如何传播，是否超过了本区域的临界电荷。最后，验证芯片整体的 SER 是否达到设计要求。

SINJECT 系统既可以在门级，也可以在 RTL 和系统级模型上注入故障。

文献[67]对龙芯 1 号进行了软差错敏感性分析，同时运行龙芯 CPU 和出错 CPU，监视两个处理器的运行过程，收集故障信息并比较。（软差错敏感性的定义为产生处理器运行失效的软差错总数占故障注入软差错总数的比例。）

文献[68]在评估微控制器 FT51 的容错性能时，故障注入机制采取了随机复位寄存器的值和在组合逻辑中模拟毛刺等方法，检测时将程序的运行轨迹与预期轨迹进行比较，模拟实验后对故障屏蔽检出率进行了理论建模和分析。

系统对故障响应的信息可以用来分析结构级影响可靠性的因素。有研究发现，系统失效率与每周期执行指令数（IPC）、分支预测率、Cache 命中率呈正比例关系^[69]。实际上，高性能处理器中的空闲逻辑、死程序（dead program）状态、错误猜测都会屏蔽瞬时故障的影响。实验发现，微结构级的屏蔽能力与流水线上的有效可提交指令数呈反比关系，但故障屏蔽率并不是无限制降低：当流水线上指令数接近饱和情况下，屏蔽率趋于渐近值 70%，不再低于这个值。

为了更好地在体系结构层面评估 SER，学术界提出了体系结构脆弱因子（Architectural Vulnerability Factor, AVF）的概念。AVF 定义为故障发生后在用户层面表现为差错的概率^[70]。处理器中不同部件的 AVF 差别很大，比如分支预测器出现故障几乎对结果无影响，其 AVF 近似为 0；而程序计数器中出现故障，则会影响程序路径，其 AVF 接近 100%。体系结构研究者通过 AVF 来判定各个部件是否需要加强故障保护机制。

评估 AVF 的难点在于判断发生的故障是否会影响最终的结果，其判别过程如图 7 所示。

计算单个部件或者整个系统的 AVF 时，其结果和负载类型关系很大。因此，评估过程中往往采用平均 AVF 值。要评估 AVF 值，就需要知道程序中哪些指令对结果的正确性有影响，哪些指令对结果正确性没有影响。如果该位出现故障会影响程序最后的结果，则将该位定义为 ACE（Architecturally Correct Execution）位，反之定义为 un_ACE 位。在实施的过程中，判断 ACE 位往往比较困难。一般先判断哪些指令为 un_ACE 位，剩余的就被认为均对程序正确执行有影响。

从微体系结构层面判断，如下指令将不会对程序的正确执行产生影响：

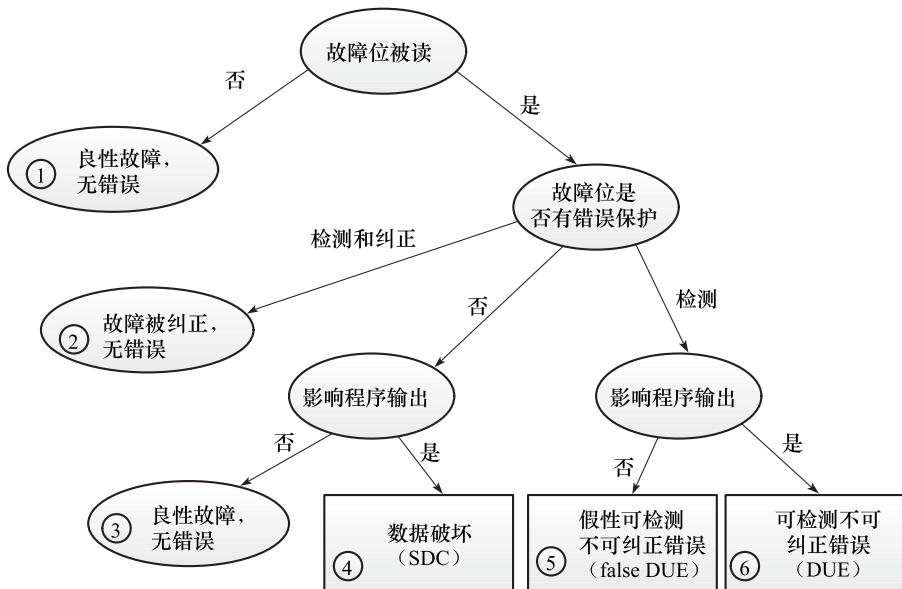


图 7 故障导致的输出分类

- NOP 指令。
- 性能提升指令：比如预取指令等。
- 错误路径上的指令。
- 动态死指令：该指令结果不被任何其他指令使用。
- 被逻辑屏蔽的指令：比如与操作中一个操作数为 0，则另一条指令的故障则会被屏蔽。

性能模型往往由 C 或者 C ++ 实现，是机器时序行为的抽象模型。该模型主要用于在设计初期评估性能。它运用性能模拟器进行模拟实验，在此基础上进行 ACE 分析，进而计算出 AVF。该方法也可以用于寄存器传输级（RTL）模型，但是很少被使用，主要是因为在设计过程中，RTL 模型的获得一般较晚。其次，基于 RTL 模型的模拟器运行速度比高级语言实现的模拟器的速度慢得多。

基于性能模型的方法也存在一些局限性。例如，在判断哪些指令会影响程序正确执行的过程中，如何界定程序正确执行有着不同标准。在通常执行情况下，很多信息用户并不可见，且这些信息出现错误并不影响用户最终所能观测到的结果，所以从用户的角度看应该界定为正确执行。但是在调试模式下，这些信息的错误是能被用户观测到的，由于此时用户并不能判断该错误是否会影响程序最终的执行结果，所以只能保守地认为程序执行出现故障，即便这不影响程序的最终结果。

AVF 方法在处理器的 RTL 模型中注入故障，也有的研究选择在 C 语言模型中注入故障。它往往需要对故障模式做一定的假设，比如，在一般情况下，假设软差错的时间和空间服从均匀分布。有研究表明，该方法和基于硬件的故障注入方法相比，有着很好的一致性^[71]。和基于原型的故障注入方法相比，该方法提供了更有效的可观测性，从而可以有效地监测故障的传播及故障对于结果的影响。学术界对此已有大量的研究，所涉及

的处理器包括 DLX、Alpha、SPARC、ARM、龙芯 1 号、控制器 FT51 等。

基于抽象模型的故障注入方法得出了很多有价值的结论，文献[72]表明试验中所选的 DLX 处理器中大约有 96% 的软差错被屏蔽，没有在体系结构层面体现出来。而且软差错导致系统错误的概率和负载类型有着非常紧密的联系，主要表现为存储器访问越频繁，软差错导致系统错误的概率越高。

近年来，随着多核多线程结构的出现，文献[73]针对多线程结构研究了软错误导致的脆弱性。结果表明，较之于超标量结构，同时多线程结构对于软差错更敏感，因为在同时多线程结构中，由于线程间对于资源的共享，资源的利用率更高了，因此结构中存在关键信息的时间比例更高。访存密集型负载比计算密集型负载对软差错更敏感，因为前者更容易遇到 Cache 访问缺失，在这种情况下，指令会在流水线中滞留更长时间。进一步的研究表明，取指策略对于软差错率也有很重要的影响，其中 Flush 策略效果最好。

AVF 表征微处理器中不同结构发生的软差错导致程序执行出错的概率，因此，在线计算不同结构的 AVF，能够指导动态选择具有不同容错能力的容错技术，在满足可靠性的前提下降低带来的性能开销。文献[74]提出了基于占用率的在线计算方法，通过分析程序运行过程中不同结构的占用率，避免了进行复杂的体系结构正确执行位分析。该方法首先记录程序运行时不同结构的占用率，然后分析并排除包含的空指令，计算当前程序段运行时的 AVF，并对后续程序段运行时的结果进行预测，根据预测结果指导选择合适的容错技术。实验结果表明，该方法能够有效计算不同结构的 AVF，与离线计算方法相比，发射队列、重排序缓存和存取队列三个结构的误差仅为 0.10、0.01 和 0.039。

研究表明，间歇故障已经成为影响微处理器可靠性的重要因素，目前缺乏量化分析间歇故障对微处理器可靠性影响的指标和方法。文献[75]首先分析导致间歇故障产生的原因，并对间歇故障建模。根据微处理器中不同结构对间歇故障的屏蔽效应，计算间歇故障脆弱因子。此外，通过改变微处理器配置参数以及选择不同程序段，计算对应结构的间歇故障脆弱因子。实验结果表明，间歇性固定为 1 故障引起程序执行出错的概率最高。此外，间歇故障脆弱因子随着待分析的结构以及运行程序的变化而发生变化，预示着设计时可以着重保护间歇故障脆弱因子高的结构，以降低对系统性能和功耗的影响。

4.4.2 系统级模型解析法

系统级模型解析法首先建立可描述时序行为的系统级模型，并用该模型描述 SEU，分析各种类型的屏蔽作用，最后计算系统级的失效率。SEU 引发系统级失效的概率可以衡量系统级正确性。SEU 可能不会影响系统输出，或可能出现延迟。因此，通过模拟方法评估系统级的软差错影响比较困难。

马尔可夫链模型可以描述时序系统的行为，并计算系统级的失效率。它用基于 BDD（二进制决策图）和 ADD（代数决策图）的符号模型描述 SEU，并处理逻辑屏蔽、电气屏蔽及锁存窗屏蔽。该模型可以很好地估计小规模电路在受到粒子撞击后的几个时钟周期中产生差错的概率。

基于电路结点签名的 SER 评估工具 AnSER^[76]既可以评价组合电路又可以评价时序

电路，但是 AnSER 没有考虑电气屏蔽对系统 SER 的影响，而电气屏蔽对于电路 SER 的影响是不可忽视的。

AVF 可以被表述为每个时钟周期中部件所包含的关键信息位数与结构中的总位数之比。根据 Little's Law，它可以表述为以下公式： $AVF_{structure} = B \times L/N$ 。其中，B 表示进入该部件的平均带宽，L 表示每个关键信息位在结构中的平均滞留时间，N 表示结构中的总位数。在多数情况下，B 和 L 可以利用硬件性能计数器来获得。有如下结论：处理器的 IPC 不能单独决定 AVF，一般情况下，可以将 IPC 视为 B 值。

4.4.3 电路原型与产品的加速测试法

通常，在制造出集成电路产品后，可通过筛选和可靠性试验估计其可靠性，并采用加速寿命试验确定产品的平均寿命。如果发现可靠性不满足要求，就要从设计和工艺角度进行分析，并加以改进。

评价器件质量和可靠性的方法主要有 3 类：

- 1) 批接收抽样检验，检验该批产品是否满足产品规范要求。
- 2) 可靠性寿命试验，评价产品的可靠性水平。
- 3) 从现场收集并积累使用寿命数据，评价相应产品的使用质量和可靠性。

加速测试方法使用电路原型或实际产品作为实施目标，采用特定的物理手段从外界对其施加影响，并且多采用的是非接触式的故障注入。所采用的外部干扰产生源包括重离子辐射、电磁干扰、激光放射等，它们可以引起目标芯片内的突发性电流。这种方法的优点在于它可以模拟目标系统所处的真实环境中所遭遇到的物理现象，缺点是由于不能精确定重离子辐射或电磁场产生的时刻，导致无法控制故障注入的确切时间和位置。

在实际工程实践中，美国军用标准 MIL-HDBK-217 的推荐方法被广泛地应用于微电子器件的失效率估计^[77,78]。假设电路是在同样环境、同样工艺下生产的一批产品，而且都处在偶发失效期，其失效率近似为常数。

近年来，VLSI 电路集成度不断提高，同时可靠性水平也迅速提高，传统的评价方法暴露出各种各样的问题，如批接收抽样检验方法因分辨能力有限已不能有效区分高水平产品质量之间的不同；可靠性寿命试验方法因要求的样本数太多而导致成本上升；基于现场数据收集的方法因存在“滞后性”而不能及时对产品质量进行评价，这就促使人们开始研究新的评估技术。研究者提出了工艺能力评价（Cpk）、工艺过程统计受控状态分析（Statistical Process Control, SPC）、产品出厂平均质量水平 PPM 考核等方法，并在产业界得到了应用。

系统级的 SER 可以通过存储器和逻辑电路的原始差错率（即无工作负载时的分析结果），及其降级因子合成得到^[79]，即 $SER_{system} = A_{mem} \times SER_{mem} + A_{logic} \times SER_{logic}$ 。其中， A_{mem} 和 A_{logic} 分别对应存储器和逻辑电路的降级因子，用以反映各自占系统差错率的比例，其取值依赖于芯片结构、版图布局和运行的应用程序。而对于原始 SER，可以通过在强辐射的实验环境下统计单个位的失效率来测量。

Swift 等人采用重离子辐射的方法研究 PowerPC750 微处理器的容错特性^[80]。文献

[79]分析了片上存储器 SRAM 的 SER。实验环境使用 Am241 和 Th232 金属薄片放射 alpha 粒子、中子，模拟宇宙射线的中子辐射，放射速度和剂量均高于真实环境。观察不同电压、频率下，比较 SRAM 和触发器在 alpha 粒子和中子两种环境下的 SER。结果发现，SRAM 单元对两种环境都没有特别的依赖性，因为其对称的版图和电路设计，敏感结点是平衡的；而对于触发器，其数据状态表现出对 alpha 粒子环境更强的敏感性。

但是，使用加速测试方法后发现，用于测试的存储单元类型及逻辑单元数目是受限的^[81,82]。这样，就需要基于模拟的故障注入方法加以补充。

5 国内外研究现状比较

5.1 可靠性设计方法现状比较

我国对可靠性设计的需求还是很强烈的，例如以神舟飞船为代表的空间技术、蛟龙号深海运载技术等都对可靠性有很高的要求。但是，出于保密的原因，国内研究相对比较封闭，导致研究成果分散，较难形成系统的可复用的核心技术。相比之下，以美国为代表的西方国家相对开放，IBM 公司的高可靠 Z 系列处理器的微结构都已经公开发表，对故障检测和恢复等核心机制都有详细描述，对可靠性设计这一研究方向提供了很有价值资料。

其次，由于国产处理器还没有在民用领域广泛应用，因此针对微处理器的可靠性设计方法的工作较少，至今没有公开系统的可靠性设计技术。但在系统可靠性领域，有不少集群软件出现，通过软件实现系统可靠性。例如 LanderCluster 集群软件、PlusWell 集群软件等。总体而言显示出“软件先于硬件”的趋势。但这些并不能弥补微处理器芯片可靠性缺陷。

5.2 可靠性理论研究现状比较

在可靠性理论研究方面，国内外在集成电路可靠性评估技术的研究上也存在显著差别。国外的工作是全方位展开的，既重视理论研究，也重视技术开发和工程应用。国内高校和研究机构的研究注重学术性，但缺乏后续明确的应用目标，成果转化慢。国内企业的工作一般围绕重点工程和国防建设展开，重视具体产品开发和工程应用，但基础理论方面的研究深度不够。高校和企业的关注点的重合性不够大。

国外研究的参与者既有高校，也有企业，分工较为明确，高校注重理论研究和工具原型开发，企业则以实用或商用工具开发和应用为主。除了与国防建设相关的高校、研究机构和企业之外，国内参与研究的普通单位的分工还不是十分明确。近年来，随着国家产学研用相结合的科技策略的推进进入实质性阶段，这种现象正在逐渐改变。

面向产品的可靠性评估是国内传统的研究方向，也取得了较好的成效，但与国外相比，工程应用水平依然存在差距。近年来，国内普通高校和研究机构也加入了集成电路可靠性研究领域，主要关注集成电路高层可靠性评估，但研究内容比较分散，研究力量也不够强，达到国际先进水平的学术成果还不多。

6 可靠性设计技术展望

6.1 近阈值电压计算模式下的可靠性设计

目前，功耗已成为芯片设计的重要瓶颈。随着片上集成度日益增加，功耗墙的限制导致片上所有晶体管工作难以维持。近阈值电压工作模式逐渐走进人们的视野。在这种模式下，供电电压略大于阈值电压，研究发现此种模式下处理器效能较高。Intel 在 ISS-CC2012 上展示了近阈值工作模式的 x86 处理器 (Claremont)，其功耗仅 2mW。这一趋势是很激动人心的。但需要注意的是虽然近阈值计算很好地平衡了能耗和性能，同时其也面临一些挑战，主要包括性能的降级与可靠性问题。我们将从以下三个方面加以论述。

6.1.1 性能降低

由于在近阈值模式下供电电压低于正常电压，故其性能相比于正常电压有所下降。在 45nm 工艺下，FO4 (4 扇出反向器) 在 400mV 和 1.1V 的工作电压下频率相差 10 倍。这也是其是否能得到广泛应用的主要挑战。因此，近阈值计算亟须结合新型体系结构设计来满足应用的新的需求，例如利用应用中的数据级并行性提高近阈值处理器的吞吐量。适合近阈值计算模式的应用需求还需积极探索。

6.1.2 性能偏差

在近阈值电压下，MOSFET 的驱动电流和 V_{th} 、 V_{dd} 及温度存在指数关系，故近阈值芯片的性能偏差更大。相比于正常电压工作模式下 30% 的偏差，近阈值模式下的性能偏差会增大到 400%。许多工作致力于减少近阈值电压处理器的性能偏差，这些工作主要包括：1) 新型工艺，例如 FinFet 技术。2) 电路设计，主要包括边缘时钟控制和电压偏置。边缘时钟控制可以使用基于双相锁存器或者触发器的设计来控制时钟，或使用一种模糊时钟边界的触发器，通过一个小的透明窗口来进行时钟调整。电压偏置是一种通过改变基极的电压，调整偏差芯片的性能与漏电功耗的技术。但通过电路设计减少性能偏差均存在开销较大的问题。3) 微体系结构级偏差补偿方案。通过增加冗余部件，弃用偏差较大的部件以补偿偏差。4) 体系结构级偏差补偿方案。通过使用双电压和系统计数器来负载均衡，减少处理器核级的性能偏差。

6.1.3 功能故障

工艺、温度和电压偏差不仅影响近阈值芯片的性能，还可能使其发生功能性故障。尤其是对于 SRAM 功能部件。标准 6T SRAM 类型单元在存在工艺偏差的情况下 P 管和 N 管的 V_{th} 值均会升高，所以存在潜在的读写故障。近阈值电压处理器的可靠性的本质问题是相比于正常电压处理器，其随着供电电压减小，工艺、温度和电压偏差的影响大大增加。故通过偏差补偿，可以部分解决功能故障问题。除此之外，还可以通过以下方法解决一定的功能故障问题：1) 电路设计。改变 SRAM 类型单元结构，例如使用 10T SRAM，减少发生功能故障的可能。2) 冗余单元设计，使用冗余单元存放校验信息，利用冗余空间进行容错。3) 体系结构设计，通过改变缓存行的映射关系，增加可用单元，提高可靠性。

总体来说，近阈值工作模式拥有更好的效能，但同时也存在着许多可靠性以及选择应用场景的问题。因此，需要探索新型的应用模式及适合近阈值模式的体系结构，同时为其提供低开销的可靠性保障。

6.2 三维芯片的可靠性设计

按照摩尔定律，芯片发展脚步正在逐步放慢，从 45nm 到 28nm 的制程经历了 33 个月，远远超过摩尔定律预测的 18 个月。而且，制程升级的难度越来越大，所需的时间将会越来越长。业界普遍认为三维（3D）芯片将会是缓解摩尔定律脚步放缓的强心剂。3D 芯片的概念在本世纪初提出^[83]，与传统的 2D 芯片相比，其优点包括更高的集成度、更好的性能、更低的互连功耗，并且支持不同工艺的晶片集成（即“异质集成”）。市场调研机构 YoleDeveloppement 预测，到 2017 年，基于 TSV 的 3D IC 市场产值将达到 400 亿美元，占半导体市场的 9%。目前，包括三星电子、高通、美光、台积电、日月光等在内的主要半导体厂商都在研发 TSV 3D IC 方面做了大量投入。

6.2.1 三维芯片互连的可靠性设计

与传统的 2D 芯片相比，针对 3D 芯片的可靠性研究尚不成熟。3D IC 引进了独特的生产工艺，例如 TSV（Through Silicon Via）的刻蚀、硅衬底研磨等，这些技术都还处于发展阶段。制造缺陷会给 3D IC 带来许多可靠性问题，电迁移效应（Electro-Migration Effect）就是其中一个突出问题。研究表明，制造缺陷会改变硅通孔的导通特性（电阻值），加剧电迁移效应。需要深入研究镀铜气泡、绑定非对齐和绑定界面尘埃沾染等制造缺陷和电迁移效应之间的关系，并提出相应的故障模型。

6.2.2 三维芯片电源噪声及热点均衡

在三维芯片中，由于电流从最底层的引脚输入芯片，并逐层送往更高层的芯核，因此电流的密度会更大，上层的芯核的电压稳定性也会受到下层芯核的影响，在垂直方向

上产生类似 IR 降和感应噪声效应。需要研究如何隔离不同层间电源噪声的耦合效应。其次，由于硅的导热性能较差，传统的散热方案很难有效控制三维芯片热分布，因此需要研究立体的冷却方案和负载均衡策略。

6.3 面向应用的可靠性设计

集成电路工艺特征尺寸的进一步细化使得晶体管变得越来越不可靠。影响可靠性的因素包括硬故障、软差错、老化和工艺偏差等。传统的容错技术（Fault Tolerance）的目标是保证系统出现故障时仍能正常工作，应用的输出结果不能出现任何差错。由于它没有考虑应用的语义（Semantics），因此通常会使得可靠性指标过设计，从而在能耗、性能和面积上引起较大的开销。但是，从应用的角度来看，并非所有应用都需要完全精确的输出结果，例如多媒体应用以及新型的 RMS（Recognition, Mining 和 Synthesis）应用。因此，最近几年近似计算（Approximate/Inexact Computing，其他类似的术语包括 Error Tolerance, Soft Computing, Application-Level Correctness 等）受到了广泛关注。

应用对差错固有的容忍能力来源于以下几个方面：1) 输入数据有噪声或冗余；2) 输出结果的轻微偏差不会影响感知（例如图像和视频中像素的轻微偏差不会影响人眼感知）；3) 计算模式具有统计或概率特性^[84]。由于放松了可靠性设计的约束，因此利用应用本身的差错容忍特性来进行容错设计能够降低设计成本。面向应用的容错技术主要针对最重要的部分（即对应用的输出结果影响最大的部分）进行可靠性设计（类似的术语还有轻量级容错（Lightweight Fault Tolerance）、选择性加固（Selective Hardening）），在 ASIC 和 CPU 中均已得到使用。下面将分别从硬件层、软件层以及软硬件协同的角度来介绍面向应用的容错技术。

6.3.1 面向应用的硬件容错

传统的存储器可靠性设计方法包括纠错码和冗余行列修复，它们对所有的数据位一视同仁。对一些应用而言，数值数据的表示有 MSB（Most Significant Bit）和 LSB（Least Significant Bit）之分，低位出错对应用的影响很小。因此，在对存储器进行可靠性设计时，考虑到不同数据位的重要性不同，可对重要位采用较强的纠错码，对非重要位采用较弱的纠错码，或者采用冗余修复方法优先修复存储重要位的存储单元。对于视频应用中的相变存储器，可以采用近似写方法（即新数据和旧数据的绝对差值不超过设定的阈值时就取消写操作）减少其写操作，从而提高其生命期可靠性^[85]。

考虑到应用的差错容忍特性，对逻辑电路同样可采用选择性加固的方法进行容错设计，例如针对受软差错影响的触发器，只选择对输出影响最大的触发器进行保护。

6.3.2 面向应用的软件容错

考虑到软差错对应用的输出结果有不同的影响，可通过轻量级检查点进行容错，即只需要保存导致程序崩溃的数据，从而减少检查点数目，降低性能开销。另一种典型的

容错方法是识别对应用影响最大的关键指令并对关键指令进行复制。类似的方法还包括在程序的关键位置部署差错检测器。除此之外，还有面向应用的算法容错。该方法利用应用本身的特性进行纠错，例如图像处理中常用的滤波算法可用来缓解图像中差错的影响，硬件神经网络可以重新训练以应对电路中出现的永久性故障。

6.3.3 面向应用的软硬件协同容错

由于软件和硬件能够发挥各自的优势，因此将二者结合在一起的软硬件协同容错相比于纯粹的硬件或软件容错具有更灵活、更广泛的应用前景。例如，Relax 扩展指令集并同时对软硬件进行修改，可以有效地在软件层进行故障恢复；ERSA 结合 RMS 应用的特性和多核不对称的可靠性（有些核可靠，有些核不可靠）并通过软件优化可以协同容忍各种差错；3D 图像传感器中的微焊点、模数转换器和 TSV 都有可能出现缺陷，为了提高其可靠性，可以首先在硬件上交错像素和模数转换器之间的连接，然后通过去噪算法纠错。

集成电路越来越不可靠的趋势以及新型差错容忍应用的出现为面向应用的容错技术提供了良好的契机。作为一类低成本的容错方法，面向应用的容错技术为上层应用和底层硬件提供了沟通的桥梁，有望获得更广泛的应用前景，其探索空间还很巨大。首先，由于面向应用的容错技术与应用密切相关，因此选择合适的指标深入理解和刻画各种应用的差错容忍特性依然是基础性但意义深远的工作。其次，在通用计算平台下，结合各种应用本身的特性以及硬件自身的不可靠性来进行软硬件协同容错是未来的发展趋势，而且协同容错可能贯穿编程模型、算法、体系结构、电路等各个层次。此外，根据应用和硬件不同阶段的需求来动态调节应用输出结果的精度和容错设计方案也是一个值得探索的研究方向。

7 结束语

高可靠微处理器芯片是构建可靠计算服务的基石，不仅在任务关键领域具有特别重要的意义，在民用领域也凸显出其重要性。在过去的十年中，学术界将可靠性与性能、功耗作为芯片的综合评价指标，但是从芯片的产品来看，业界还是更注重性能和功耗，而对可靠性的重视程度相对较弱。高可靠处理器由于涉及军用等保密领域，研发具有我国自主知识产权的相关核心技术意义更加重大。因此希望国家加大这一领域的持续投入，有效推动国内高可靠微处理器芯片的研发和应用。国内的学术界和芯片研发机构应该抓住这一历史机遇，产生出系统的具有自主知识产权的高可靠微处理器设计研发技术，在与国外的芯片厂商的竞争中取得优势。

参考文献

- [1] S Feng, S Gupta, A Ansari, S Mahlke. Shoestring: Probabilistic Soft Error Reliability on the Cheap [C]. ASPLOS'10, March 13-17, 2010, Pittsburgh, Pennsylvania, USA.

- [2] Pham D, S Asano, et al. The Design and Implementation of a Rirst-generation CELL Processor [C]. ISS-CC'05 , 2005.
- [3] L Zhang, Y Han, Q Xu, X Li, H Li. On Topology Reconfiguration for Defect-Tolerant NoC-Based Homogeneous Manycore Systems[J]. IEEE Transactions on Very Large Scale Integration(VLSI) Systems , 2009 , 17(9) : 1173-1186.
- [4] G Yan, X Li. Online Timing Variation Tolerance for Digital Integrated Circuits[C]. International Test Conference(ITC'11) , Anaheim, USA. Sep. 2011.
- [5] G Yan, Y Han, X Li. ReviveNet: A Self-adaptive Architecture for Improving Lifetime Reliability via Localized Timing Adaptation[J]. IEEE Transactions on Computers(TC) , 2011 , 60(9) : 1219-1232.
- [6] G Yan, X Liang, Y Han, X Li. Leveraging the Core- Level Complementary Effects of PVT Variations to Reduce Timing Emergencies in Multi-Core Processors[C]. 37th Annual International Symposium on Computer Architecture (ISCA '10) , Saint-Malo, France , 2010: 485-496.
- [7] T C May, M H. Woods. Alpha-Particle-Induced Soft Errors in DynamicMemories[J]. IEEE Transactions on Electron Devices , 26(1) : 2-9 , 1979.
- [8] Y-M Jiang, K-T Cheng. Analysis of Performance Impact Caused by Power Supply Noise in Deep Submicron Devices[C]. 36 thDesign Automation Conference. New Orleans, LA , USA , 1999.
- [9] S Mitra, N Seifert, M Zhang, Q Shi, K S Kim. Robust System Designwith Built-in Soft-Error Resilience [J]. COMPUTER , 2005: 43-52.
- [10] D C Bossen, J M Tendler, K Reick. Power4 System Microarchitecture[J]. IBM Journal of Research and Development , 46(1) : 5-25 , 2002.
- [11] B Sinharoy, R N Kalla, J M Tendler, R J Eickemeyer, J B Joyner. Power5 System Microarchitecture [J]. IBM Journal of Research and Development , 49(4/5) : 505-521 , 2005.
- [12] S K Reinhardt, S S Mukherjee. Transient-Fault Detection Via SimultaneousMultithreading[C]. 27th Annual International Symposium onComputer Architecture. Vancouver, BC, Canada , 2000: 25-36.
- [13] N J Wang, S J Patel. Restore. Symptom Based Soft Error Detection inMicroprocessors[C]. 35th Annual IEEE/IFIP InternationalConference on Dependable Systems and Networks, Yokohama, Japan , 2005: 30-39.
- [14] V K Reddy, S Parthasarathy, E Rotenberg. Understanding Prediction- Based Partial Redundant Threading for Low-Overhead, High-Coverage FaultTolerance[C]. 12th International Conference on Architectural Support for Programming Languages and Operating Systems. San Jose, CA , USA , 2006: 83-94.
- [15] G A Reis, J Chang, N Vachharajani, R Rangan, D I August. Swift: Software Implemented Fault Tolerance [C]. 3rd InternationalSymposium on Code Generation and Optimization, San Jose, California , 2005: 243-254.
- [16] Hu W J Wang, et al. Godson-3: A Scalable Multicore RISC Processor with x86 Emulation [C]. IEEEEMI-CRO , 2009 , 29: 17-29.
- [17] L Zhang, Y Yu, J Dong, Y Han, S Ren, X Li. Performance- Asymmetry- Aware Topology Virtualization for Defect-Tolerant NoC-based Many-core Processors[C]. IEEE/ACM Design, Automation and Test in Europe (DATE) , Dresden, Germany , March 8-12, 2010: 1566-1571.
- [18] 李晓维, 胡瑜, 张磊, 鄢贵海. 数字集成电路容错设计——容缺陷/故障、容参数偏差、容软错误 [C]. 北京: 科学出版社 , 2010.
- [19] David Bernick, Bill Bruckert, Paul Del Vigna, David Garcia, Robert Jardine, Jim Klecka, Jim Smullen. NonStop Advanced Architecture[C]. International Conference on Dependable Systems and Networks , 2005.

- [20] Y Xie, X Fu, Z He, Y Zhao, Y Hu, X Li. A self-Repairable Microprocessor [J]. ECS Transactions, 2009, 18(1): 249-254.
- [21] 绳伟光, 肖立伊, 毛志刚. 用于电路级仿真软故障注入的自动化方法[J]. 计算机辅助设计与图形学学报, 2009, 21(3): 346-353.
- [22] M Zhang, N RShanbhag. A Soft Error Rate Analysis(SERA) Methodology [J]. IEEE Trans. On Computer-Aided Design of Integrated Circuits and Systems. 2006, 25(10): 2140-2155.
- [23] R Rajaraman, J S Kim, N Vijaykrishnan, Y Xie, et al. SEAT-LA: A Soft Error Analysis Tool for Combinational Logic[C]. IEEE 19th Int'l Conf on VLSI Design, Hyderabad, January 2006: 499-502.
- [24] D Alexandrescu, L Anghel, M Nicolaidis. New Methods for Evaluating The Impact of Single Event Transients in VDSM ICs[C]. IEEE 17th IntSympOn Defect and Fault Tolerance in VLSI Systems, Vancouver, November 2002: 99-107.
- [25] G Hubert, J M Palau, K Castellani-Coulie, et al. Detailed Analysis of Secondary Ions' Effect for The Calculation of Neutron- Induced SER in SRAMs [J]. IEEE Trans. on Nuclear Science, 2001, 48 (6): 1953-1959.
- [26] G Hubert, et al. Simplified Monte-Carlo DASIE method dedicated to the MBU understanding and characterization in integrated devices[C]. IEEE Nuclear & Space Radiation Effects Conference, Seattle, June 2005: 75-91.
- [27] D Bhaduri, S Shukla, S Paul. Reliability Analysis of Large Circuits Using Scalable Techniques and Tools [J]. IEEE Transactions on Circuits and Systems. 2007, 54(11): 2447-2460.
- [28] M R Choudhury, K Mohanram. Reliability Analysis of Logic Circuits[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2009, 28(3): 392-405.
- [29] H Cha, E M Rudnick, G S Choi, et al. A fast and Accurate Gate-level Transient Fault Simulation Environment [C]. 23rd International Symposium on Fault- Tolerant Computing. Toulouse, France, 1993: 310-319.
- [30] Y S Dhillon, A U Diril, A Chatterjee, et al. Sizing CMOS Circuits for Increased Transient Error Tolerance [C]. IEEE 10th Int'l On-Line Testing Symp., Funchal, July 2004: 11-16.
- [31] Y S Dhillon, A U Diril, A Chatterjee. Soft-error Tolerance Analysis and Optimization of Nanometer Circuits [C]. Conf on Design, Automation, Test in Europe, Munich, March 2005: 288-293.
- [32] C Zhao, X Bai, S Dey. A Scalable Soft Spot Analysis Methodology for Compound Noise Effects in Nanometer Circuits[C]. Conf on Design Automation, San Diego, July 2004: 894-899.
- [33] M SonzaReorda, M Violante. Accurate and Efficient Analysis of Single Event Transients in VLSI Circuits [C]. IEEE 9th Int'l On-Line Testing Symp., Kos Island, Greece, July 2003: 101-105.
- [34] B Zhang, W S Wang, M Orshansky. FASER: Fast Analysis of Soft Error Susceptibility for Cell-Based Designs[C]. 7th Int'l Symp. on Quality Electronic Design, March 2006: 1-6.
- [35] M Zhang, N R Shanbhag. Soft-Error- Rate- Analysis(SERA) Methodology [J]. IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems. 2006, 25(10): 2140-2155.
- [36] K Mohanram, N A Touba. Cost- effective Approach for Reducing Soft Error Failure Rate in Logic Circuits [C]. International Test Conference, Texas, September 2003: 893-901.
- [37] K N Andre, J Samir, J Goran. Combinational Logic Soft Error Analysis and Protection[C]. IEEE 12th Int'l On-line Testing Symposium, Lake of Como, July 2006: 43-46.
- [38] S J S Mahdavi, K Mohammadi. SCRAP: Sequential Circuits Reliability Analysis Program [J]. Microelectronics Reliability. 2009, 49(7): 924-933.

- [39] M R Choudhury, K Mohanram. Reliability Analysis of Logic Circuits[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2009, 28(3) : 392-405.
- [40] R C Ogas. The Probability of a Correct Output From a Combinational Circuit[J]. IEEE Trans. On Computers, 1975, 24: 534-544.
- [41] J S Kim, C Nicopoulos, N Vijakrishnan, et al. A Probabilistic Model for Soft- error Rate Estimation in Combinational Logic[C]. 1st Int'l Workshop on Probabilistic Analysis Techniques for Real Time and Embedded Syems, Italy, Elsevier Science, June 2004. 25-31.
- [42] G Asadi, M B Tahoori. An Analytical Approach for Soft Error Rate Estimation in Digital Circuits[C]. IEEE IntSymp on Circuits and Systems, Kobe, John Wiley & Sons, May 2005 : 2991-2994.
- [43] S Krishnaswamy, G F Viamontes, I L Markov, et al. Accurate Reliability Evaluation and Enhancement via Probabilistic Transfer Matrices[C]. Design, Automation and Test in Europe Conference and Exhibition, Orlando, March 2005 : 282-287.
- [44] 王真, 江建慧. 基于概率转移矩阵的串行电路可靠度计算方法[J]. 电子学报. 2009, 37(2) : 241-247.
- [45] J Xiao, J Jiang, X Zhu, C-T Ouyang. A Method of Gate-level Circuit Reliability Estimation Based on Iterative PTM Model[C]. IEEE 17th Pacific Rim International Symposium on Dependable Computing, Pasadena, USA, 2011 : 276-277.
- [46] J Han, H Chen. Reliability Evaluation of Logic Circuits Using Probabilistic Gate Models[J]. Microelectron Reliability. 2011, 51(2) : 468-476.
- [47] S Ercolani, M Favalli, M Damiani, et al. Estimate of Signal Probability In Combinational Logic Networks [C]. 1st European Test Conference, Paris, 1989 : 132-138.
- [48] D Bhaduri, S Shukla, S Paul. Reliability Analysis of Large Circuits Using Scalable Techniques and Tools [J]. IEEE Transactions on Circuits and Systems. 2007, 54(11) : 2447-2460.
- [49] M R Choudhury, K Mohanram. Reliability Analysis of Logic Circuits[J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems. 2009, 28(3) : 392-405.
- [50] T B Krishnamurthy. Improved Techniques For Estimating Signal Probabilities[J]. IEEE Trans. Computers. 1989, 38(7) : 1041-1045.
- [51] A Hossein, B T Mehdi, T Chandra. Estimating Error Propagation Probabilities With Bounded Variances [C]. IEEE 22nd Int'l Symp. on Defect and Fault Tolerance in VLSI Systems, Rome, November 2007 : 41-49.
- [52] A Hossein, B T Mehdi. Soft Error Modeling and Protection For Sequential Elements[C]. IEEE 20th Int'l Symp on Defect and Fault Tolerance in VLSI Systems. Monterey, November 2005 : 463-474.
- [53] S J S Mahdavi, K Mohammadi. SCRAP: Sequential Circuits Reliability Analysis Program [J]. Microelectronics Reliability. 2009, 49(7) : 924-933.
- [54] H Jahanirad, K Mohammadi, P Attarsharghi. Sequential Circuits Reliability Analysis Using Conditional Probabilities[C]. 19th Iranian Conference on Electrical Engineering. Tehran, Iran, 2011 : 1-4.
- [55] K Lingasubramanian, S Bhanja. Probabilistic Error Modeling For Sequential Logic[C]. 7th IEEE Conference on Nanotechnology, Hong Kong2007: 616-620.
- [56] K Lingasubramanian, S Bhanja. An Error Model to Study The Behavior of Transient Errors in Sequential Circuits[C]. 22nd International Conference on VLSI Design, New Delhi, India, 2009 : 485-490.
- [57] HUGIN EXPERT. <http://www.hugin.com>. 2012.

- [58] GeNIE& SMILE. <http://genie.sis.pitt.edu>. 2010.
- [59] J T Flaquer, J M Daveau, L Naviner, et al. Fast Reliability Analysis of Combinatorial Logic Circuits Using Conditional Probabilities[J]. *Microelectronics Reliability*, 2010, 50(9-11): 1215-1218.
- [60] M Fazeli, S G Miremadi, H Asadi, et al. A Fast Analytical Approach to Multi-cycle Soft Error Rate Estimation of Sequential Circuits[C]. 13th Euromicro Conference on Digital System Design: Architectures, Methods and Tools. Lille, France, 2010: 797-800.
- [61] C Ouyang, J Jiang, J Xiao. Reliability Evaluation Of Flip-Flops Based On Probabilistic Transfer Matrices [C]. 16th IEEE Pacific Rim International Symposium on Dependable Computing, Tokyo, 2010: 239-240.
- [62] 欧阳城添, 江建慧. 基于概率转移矩阵的时序电路可靠度计算方法[J]. 电子学报, 2013, 41(1): 171-177.
- [63] 肖杰, 江建慧. 结合版图结构信息的基本门电路故障概率估计[J]. 电子学报, 2012, 40(2): 235-240.
- [64] 肖杰, 江建慧. 考虑时间因素的不同基本门故障概率计算[J]. 电子学报, 2013, 41(4): 666-673.
- [65] 肖杰, 江建慧. 面向缺陷分析的广义门电路故障概率的计算[J]. 计算机辅助设计与图形学学报, 2013, 25(4): 564-572.
- [66] P C Murley, G R Srinivasan. Soft-error Monte Carlo Modeling Program, SEMM[J]. *IBM Journal of Research and Development*. 1996, 40(1): 109-118.
- [67] 黄海林, 唐志敏, 许彤. 龙芯1号处理器的故障注入方法与软差错敏感性分析[J]. 计算机研究与发展, 2007, 43(10): 1820-1827.
- [68] 龚锐, 陈微, 刘芳, 等. FT51: 一种容软错误高可靠微控制器[J]. 计算机学报, 2007, 30(10): 1662-1673.
- [69] N J Wang, J Quek, T M Rafacz, et al. Characterizing The Effects of Transient Faults On a High-Performance Processor Pipeline[C]. IEEE Int'l Conf. on Dependable Systems and Networks, Florence, June 2004: 61-70.
- [70] X Li, S V Adve, P Bose, J A Rivers. Architecture-level Soft Error Analysis: Examining The Limits of Common Assumptions [C]. International Conference on Dependable System and Networks, Edinburgh, UK, June 2007: 266-275.
- [71] P N Sanda, et al. Soft-error Resilience of The IBM POWER6 Processor[J]. *IBM Journal of Research and Development*, 2008, 52(3): 275-284.
- [72] P Giacinto, Saggese, J Nicholas, Wang, et al. An Experimental Study of Soft Errors In Microprocessors [J]. *IEEE Micro*, 2005, 25(6): 30-39.
- [73] W Zhang, X Fu, T. Li, J Fortes. An Analysis of Microarchitecture Vulnerability To Soft Errors On Simultaneous Multithreaded Architectures [C]. IEEE International Symposium on Performance Analysis of Systems&Software, April 2007.
- [74] S Pan, Y Hu, X Li. Online Computing and Predicting Architectural Vulnerability Factor of Microprocessor Structures[C]. IEEE Pacific Rim International Symposium on Dependable Computing, China, Nov. 2009: 345-350.
- [75] S Pan, Y Hu, X Li. IVF: Characterizing the Vulnerability of Microprocessor Structures to Intermittent faults[C]. IEEE/ACM Conference on Design, Automation and Test in Europe, Germany, Mar. 2010: 238-243.

- [76] S Krishnaswamy, S M Plaza, I L Markov, et al. Signature-based SER Analysis and Design of Logic Circuits [J]. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 2009, 28(1): 74-86.
- [77] MIL-HDBK-217 F. Military Handbook Reliability Prediction of Electronic Equipment, Notice 2. Washington, DC: Department of Defense, Feb. 28, 1995.
- [78] L Gullo. The Revitalization of MIL-HDBK-217, 2008.
- [79] X W Zhu, R Baumann, R Pilch, et al. Comparison of Product Failure Rate To The Component Soft Error Rate In A Multi-core Digital Signal Processor[C]. IEEE 43rd Annual Int'l Reliability Physics Symposium, New York, April 2005: 209-214.
- [80] G M Swift, F F Farmanesh, et al. Single-Event Upset In The PowerPC 750[J]. IEEE Transactions on Nuclear Science, 2001, 48(6): 1822-1827.
- [81] X W Zhu, R Baumann, R Pilch, et al. Comparison of Product Failure Rate To The Component Soft Error Rate In A Multi-core Digital Signal Processor[C]. IEEE 43rd Annual Int'l Reliability Physics Symposium, New York, April 2005: 209-214.
- [82] H Tino. Soft Error Rates In Deep-submicron CMOS Technologies[C]. IEEE 12th Int'l On-line Testing Symposium, Lake of Como, July 2006: 41-45.
- [83] S J Sourl, K Banerjee, A Mehrotra, K C Saraswat. Multiple SI Layer ICs: Motivation, Performance Analysis, And Design Implications[C]. 37th Conference onDesign Automation (DAC'00), New York, 2000: 213-220.
- [84] V Chippa, S Chakradhar, K Roy, A Raghunathan. Analysis and Characterization of Inherent Application Resilience For Approximate Computing[C]. 50th Conference onDesign Automation (DAC'13), Austin, USA, 2013.
- [85] Y Fang, H Li X Li. SoftPCM: Enhancing Energy Efficiency and Lifetime of Phase Change Memory in Video Applications Via Approximate Write[C]. 21st Asian Test Symposium, Niigata, Japan, 2012: 131-136.

致谢

本文中 SRP 处理器部分的资料由中科院计算所胡瑜研究员提供，在此表示衷心感谢。

作者简介

鄢贵海 中国科学院计算技术研究所助理研究员，博士，CCF 会员，CCF 优秀博士论文获得者，主要研究方向为高效能计算机体系结构、容错计算。



江建慧 同济大学教授，博士，博士生导师，同济大学软件学院副院长，CCF 容错计算专业委员会副主任，主要研究方向为可信系统与网络、软件可靠性工程、VLSI/SoC 测试与容错。



李华伟 中国科学院计算技术研究所研究员，博士，博士生导师，CCF 容错计算专业委员会秘书长，主要研究方向为 VLSI/SoC 测试、设计验证、可靠设计。



李晓维 中国科学院计算技术研究所研究员，博士，博士生导师，CCF 容错计算专业委员会主任，主要研究方向为 VLSI/SoC 测试、可靠设计、容错计算。



软件测试技术研究进展与趋势

CCF 软件工程专业委员会

摘要

软件测试是软件缺陷检测和质量控制的重要手段。本文简要介绍了软件测试的基本概念和历史发展。针对测试用例设计和故障分析两个关键环节，着重分析了近年来国际、国内研究的若干热点问题及研究进展，包括基于模型的测试生成、测试用例优化选择、测试充分性评估、基于统计的故障定位和缺陷预测等技术。选择嵌入式系统和服务化软件作为两个典型的应用领域，介绍了测试技术在这两个领域的应用和发展现状。本文最后综合分析并展望了测试技术的发展趋势。

关键词：软件测试，软件工程

Abstract

Software testing is an important method for software bug detection and quality control. This report briefly introduces the basic concepts and development history of software testing. Test case design and fault localization are two key issues in software testing. The report analyzes state-of-the-art progress of testing techniques, taking several hot topics in recent years including model-based testing, test optimization and prioritization, test adequacy criteria, statistical-based fault localization, and fault prediction. In addition, the report investigates testing research and applications in two domains: embedded system and service-based software system. Finally, it analyzes the trends and roadmaps of software testing.

Keywords: software testing, software engineering

1 引言

1945 年，Grace Hopper 博士命名了 Bug（缺陷），并形成了计算机历史上第一个有记载的 Bug。此后，如何高效及时地捕捉软件缺陷、保证软件质量，就成为软件开发的一个重要任务，软件测试的研究与实践应运而生。经过近 70 年的发展，软件测试已经成为软件工程的一个重要分支，是开发过程的主要环节之一，在软件企业中举足轻重；并且形成了独立的学科专业方向，受到学术界的广泛关注，许多大学都开设相应的专业课程；从开发人员测试到独立的第三方测试，从单元测试到集成测试到系统测试，从功能测试到非功能测试，从单机应用到互联网系统，从封闭的实验室环境到在线开放平台，软件测试积累了大量的研究成果和工程实践经验，测试方法、技术与工具都在不断地发展演化。测试与软件形式化验证、程序证明等技术互为补充，基于不同的理论基础，在软件

生命周期的不同阶段，保证软件的可信性。

1972 年，Bill Hetzel 博士组织了首次正式的软件测试会议，给出软件测试的定义为“测试就是建立一种信心，确信程序能够按预期的设想运行”^[1]。Glenford J. Myers 在 1979 年提出了“测试的目的是证伪”这一思想，即“测试是为发现错误而执行的一个程序或者系统的过程”，测试的目的就是发现尽可能多的错误。Bill Hetzel 在 1983 年重新修订其定义为“软件测试是一系列活动，这些活动是为了评估一个程序或软件系统的特性或能力，并确定其是否达到了预期结果”。综上研究观点，1990 年在 IEEE std. 610. 12 标准中给出软件测试定义为：“1) 在特定条件下运行系统或组件的过程，观测或记录结果，并对系统或组件的某些特性做出评价；2) 软件分析的过程，以识别测试结果状态与需求定义的状态之间的区别（即发现缺陷），并对软件特性进行评估”。具体来讲，软件测试就是通过模拟数据输入、软件交互的使用过程、软件运行环境的变化，运行被测系统；通过比较实际输出与预期结果的一致性，判定被测系统的功能、性能等可信属性是否满足需求。

软件测试一直以提高测试的有效性和测试效率为主要目标，即选择最有效的测试用例，以发现尽可能多的缺陷，在统计规律下，辅助评价软件的可信度水平。软件测试是发现、定位并排除系统缺陷的主要手段，迭代回归测试结果的综合分析是预测软件可信性变化规律的重要依据。在核电、航空、航天等任务关键和安全关键的软件系统中，测试过程和测试结果是评价系统可信性的重要依据。例如，在核电领域，IEC 880 标准中就严格规定了安全软件开发的基本原则，要求每个开发阶段都具有遵循严格的验证过程，实施确认测试，并提供可审计的过程文档；在航空领域，适航认证 DO-178 B/C 明确提出 5 种不同安全等级软件的测试必须满足的验证过程和测试覆盖率要求。因此，软件测试一直是软件可信性工程的重要工作之一，在大型项目中可占项目总投入的 60% 以上；在关键软件中，甚至可达其他工作总投入的 3~5 倍。G. Tassey^[2]在 2002 年发布的 NIST 报告中指出，缺乏鲁棒、规范的测试技术将导致软件企业缺陷率上升、开发代价增高、开发进度延迟、市场推广代价高，进而产生重大的经济损失。报告选取制造业和金融领域为代表，估算由于测试基础设施不充分，在全美所造成的经济损失将高达 590 亿美元；如果改进测试，可预见减少 220 亿美元损失。

在软件工程发展早期，测试投入少介入晚，只在产品形成后进行。研究发现，在软件开发的早期（需求设计阶段）常常会引入缺陷，如图 1 显示的典型的缺陷累计 S 曲线。早期的缺陷如不能及时修复，会造成越来越严重的后果，如 Boehm 在“Software Engineering Economics”中所述：平均而言，如果在需求阶段修正一个错误的代价是 1，那么设计阶段将是 3~6 倍，编程阶段是 10 倍，内部测试阶段是 20~40 倍，外部测试阶段是 30~70 倍，产品阶段可达到 40~1000 倍，修复代价几乎成指数增长。因此，测试逐步发展为覆盖整个软件开发周期、与需求设计并行的过程活动，“尽早地和不断地测试”成为软件测试的一个基本原则，“测试驱动开发”更是倡导用测试推动软件开发工作，通过先编写测试用例后编写代码的方式，保证系统需求可执行可验证，持续增量测试代码的正确性，及时准确定位缺陷。

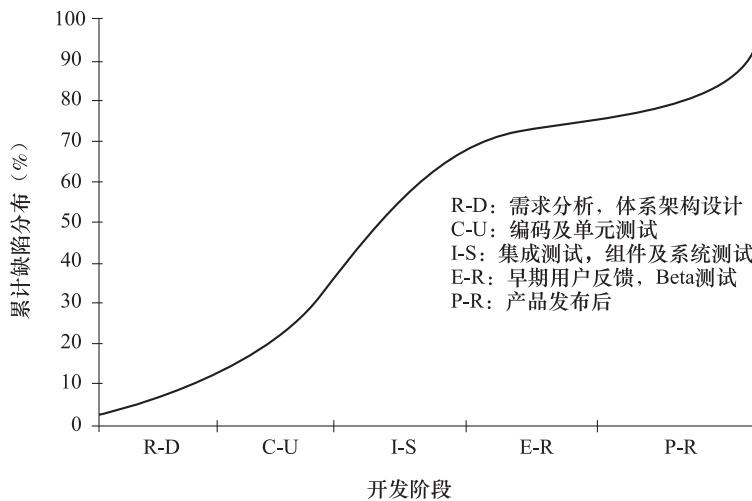


图1 典型的缺陷检测累计分布图^[2]

根据被测软件特点、测试所应用的生命周期阶段的不同，软件测试技术可以从多个角度进行分类，例如，按照基本的测试用例设计方法，可以分为黑盒测试、灰盒测试和白盒测试；按照测试的层次，可以分为单元测试、集成测试、系统测试等；按照测试目标，可以分为功能测试、性能测试、可靠性测试、安全性测试等；按照被测系统的分类，可以分为面向对象软件测试、分布式系统测试、嵌入式系统测试等。随着软件规模和复杂度的迅速增长，软件质量问题日益突出，软件测试的重要性和必要性不断提升。而随着软件形态的变化，各种新型软件范型的涌现，软件测试方法和技术也发生了重要的变化，这主要表现在以下几个方面^[3~6]：

- 软件范型演化。软件开发技术直接影响着测试技术。20世纪90年代初期，随着面向对象技术的发展，曾有学者提出模块化和软件重用等技术将大幅度提高软件质量，不再需要软件测试。事实证明，面向对象技术带来了新的风险和质量问题，例如信息封装也同时造成缺陷不易暴露；继承、多态性和动态绑定等特性都需要新的测试方法和覆盖率准则。面向对象测试随后成为一个新的研究课题。90年代末，随着Web技术、构件开发等技术发展，Web测试、构建测试也相继成为研究的专题。现在，随着Web 2.0、面向服务计算、云计算、移动计算、物联网等一系列新的技术发展变化，势必引发对新的测试方法、技术和工具的研究。
- 系统规模和复杂性增长。大型复杂软件的研制是世界关注的难题。以航空电子系统为例，美国F-22飞机航电系统软件由170万行源代码编制而成，是首次在战斗机上大规模使用软件；F-35飞机航电系统软件即达到600万行以上代码。随着被测软件的发展，软件测试规模和复杂性也相应增长，软件的质量关注点、故障模型都出现新的变化。例如，分布化、服务化、平台化成为大规模软件的主要特点，如何针对分布式环境不确定的软件行为进行测试？如何模拟规模庞大、物理

分布广泛、难以预见、复杂多样的用户行为，测试软件服务平台？为应对上述挑战，测试自动化得到了越来越广泛的关注，统计分析技术越来越多地应用于用例优化选择、充分性评估、故障定位等活动。2005年，美国国防部要求美国卡内基梅隆大学软件工程研究室（SEI）研究“在现行软件工程基础上，怎样才能构建数十亿语句的未来系统？”2006年，SEI发表了题为“超大规模系统：软件未来的挑战”的研究报告^[7]，报告分析系统的主要特征以及挑战性问题，软件质量依然是重要问题之一。

- 过程模型演化。早期的软件过程模型中，测试视为软件生命周期中代码发布后的一个单独的阶段。随后提出V模型，测试由需求设计验证、单元测试、集成测试、系统测试等一些列活动构成，分别对应软件开发的不同阶段，成为和软件开发并行的过程。随着过程成熟度模型的发展，也提出了相应的软件测试成熟度模型。极限编程等敏捷开发方法提出，测试可先于代码开发，即测试驱动开发。针对不同程序设计语言，先后涌现出一系列的单元测试框架，如JUnit，CUnit，CppUnit，Python Unit等，以辅助开发人员测试。随着开源软件、开放平台软件等架构的出现，遵循Web 2.0协同开发、群体智能等思想，测试和开发过程结合得越来越紧密，并且测试环境、测试过程越来越开放。例如，持续测试^[8]提出将测试集成在开发环境中，与代码的编译同步测试以随时发现缺陷。谷歌^[9]通过测试服务器，每天可支持2万余项目开发的50多万次Build的实时测试；在线测试（on-line testing）或被动测试（passive testing）提出在系统运行过程中实时监测系统执行，检测故障；DevOps^[10]提出构建软件开发、运营、质量保证的统一平台；uTes（www.utest.com）和Mob4 Hire（www.mob4 hire.com）等提出采用Crowdsourcing的方式，基于社区的协同测试。

本文选取若干热点技术，以点带面分析了软件测试技术的国内外研究现状，尤其是国内学者在相关领域所取得的成果，并基于Bertolino^[4]和Harrold^[5]的软件测试路线图的报告，对测试技术发展趋势进行了分析和展望。

2 国际研究现状

本文从测试用例设计、缺陷定位与预测、面向特定领域测试三个方面，调研分析了软件测试技术近年来若干研究热点问题和主要进展。在测试用例设计方面，讨论了基于模型的测试生成、测试用例优化选择、测试充分性评估等技术；在缺陷定位与预测方面，主要分析了故障定位和缺陷预测两个方面的技术进展；在面向特定领域的测试研究方面，以嵌入式系统测试和服务化软件测试为主要代表，探讨了两种背景下软件测试主要技术特点。

2.1 测试用例设计技术

测试用例指在测试执行之前设计的测试方案，说明测试所依赖的环境、操作步骤、输入数据以及每一个执行步骤之后，系统的预期状态和输出结果。测试用例设计是测试的基础和依据。Myers 提出了关于测试用例设计的重要观点，即：

- 1) 一个好的测试用例是在于它能发现以前未发现的错误。
- 2) 一个成功的测试是发现了以前未发现的错误。

大型复杂系统通常需要大量的测试用例，以检测系统的功能正确性、鲁棒性、可靠性、可用性、性能、安全性等多种功能/非功能特性。针对不同类型的软件系统，如何设计提高测试用例设计的有效性、经济性，满足不同阶段的各种测试目标，是软件测试的一个挑战性问题，也一直是研究热点。

近年来，随着软件开发技术的发展、被测软件的故障模型和质量关注点的变化，一些新的测试用例设计方法被提出并得到日益广泛的关注。

2.1.1 基于模型的测试

基于模型的测试方法（Model-Based Testing, MBT）由于可以提高测试自动化的水平，近年来得到了越来越广泛的关注。采用抽象的模型来描述系统的预期结构或行为，作为测试的需求；通过模型的分析和理解，自动生成测试用例，判定测试结果，并评价被测系统的可靠性和测试过程的质量。与 OMG 模型驱动的开发方法（Model Driven Architecture, MDA）相结合，模型驱动的测试方法（Model Driven Testing, MDT）也应运而生。MDA 强调在计算无关（CIM）、计算相关（CSM）、平台无关（PIM）、平台相关（PSM）4 个层次上建立软件模型，通过模型自动转化，实现从需求到代码的自动化的软件开发过程。MDT 提出建立测试的 PIM/PSM 模型，并实现从软件的 PIM/PSM 模型到测试的 PIM/PSM 的转化，以期达到以下目的：

- 通过自动化的模型转化技术，可以在短时间内系统地生成大量的测试用例，并保证满足一定的覆盖率准则。可以有效降低测试设计的成本，并且避免测试人员人为引入的设计错误，提高测试用例的质量。
- 建立从软件需求/设计到测试设计、测试脚本之间的关联关系，保证各个阶段模型、软件代码和测试脚本之间的完整性、一致性、可追踪性和可验证性。
- 提高测试的柔性，即当软件需求/设计变化时，可以根据从软件模型到测试模型以及到测试脚本的映射关系，识别变化的影响范围，有效选取、修改和配置相应的测试集进行回归测试。
- 提高测试的可重用性。得益于测试模型（如 U2TP）的定义，使得测试可重用的粒度大大降低，不再是测试脚本的文件，而是可以细化到每个测试行为、测试数据。
- 提高测试的可管理性和可维护性。测试模型为系统的组织、管理和维护不同粒度

的测试资产提供了依据。

MBT 的研究最早可追溯到 1976 年, Ramamoothy 等人^[11]提出了一种基于 Fortran 语言程序结构图的方法来识别测试路径并生成测试用例。近年来, 先后提出了多种软件建模和基于模型的测试方法, 如目前比较典型的模型有扩展的状态机模型 EFSM、规约描述语言 SDL、ESTELLE 和 UML 等。Uutting 等人^[12]根据模型的特征、测试用例生成技术和选择标准、测试执行的在线/离线方式, 从 7 个方面对 MBT 方法进行了分类, 如图 2 所示。

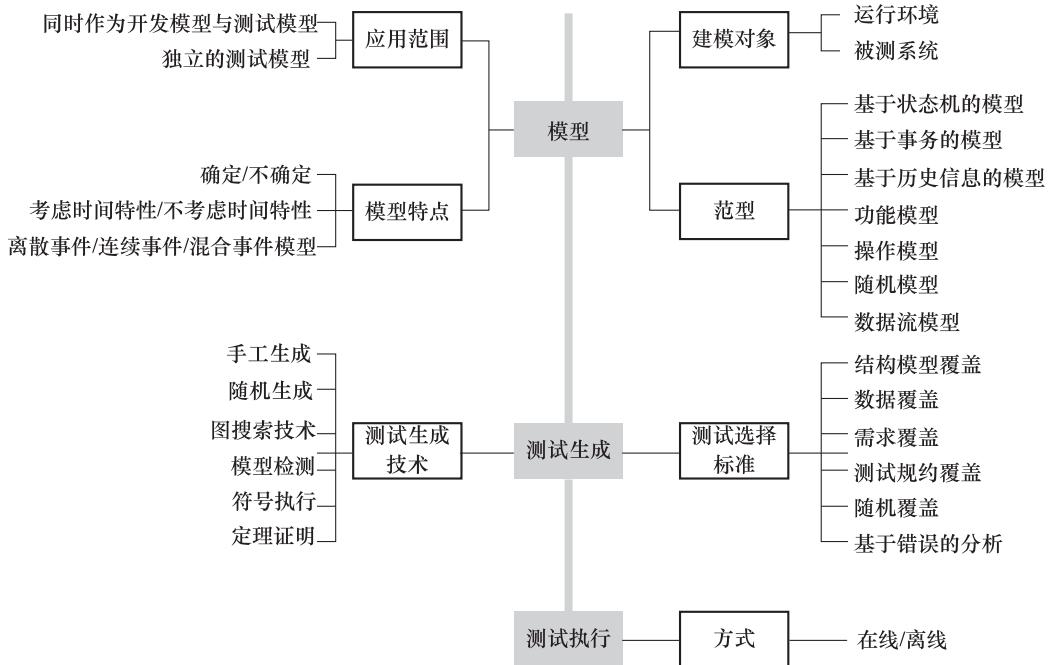


图 2 基于模型的测试方法分类图^[12]

建模对象和建模方法是基于模型测试的基础。模型可以分为不同的抽象层次和角度, 从建模的视图方面, 可以刻画被测系统的预期行为、结构或者运行环境; 从抽象层次上, Prenninger 和 Prestschner^[13]从功能、数据、通信、QoS 等几个方面分析了模型的不同抽象层次。根据开发过程不同, 测试可以与开发共享模型, 或者构建单独的测试模型。基本模型以描述被测软件的功能、确定性特性为主, 但实际软件系统还需考虑多种不确定因素(如时间、事件等), 对系统不确定性和动态性的建模是一个研究难题。Walton 等人^[14]提出了迁移概率模型, Berkenkotter 等人^[15]分析了实时和连续/离散混合系统测试的测试方法。Lamsweerde^[16]将建模的形式化方法分为 7 类: 基于状态的模型、基于转移的模型、基于时间序列的模型、功能模型、运行模型、随机模型和数据流模型。有限状态机 FSM 作为一种形式化方法, 广泛地应用于多种软件建模、分析和测试中, 例如 D. Xu 等人^[17]探讨了采用状态机模型对面向刻面的程序进行建模与测试的方法, 提出了面向结构和面向属性两种策略生成刻面测试的测试用例, 以有效检测刻面设计和

植入的错误。

基于模型的测试自动生成算法是当前的一个研究热点。典型的生成方法有随机生成^[18]、统计分布^[19]、图搜索^[20]、约束求解^[21]、符号执行^[22]等。测试数据生成问题通常可以通过建模转化为搜索问题，可以利用目前已经成熟的搜索算法来求解和生成测试数据。常用的搜索算法包括爬山算法、模拟退火算法、遗传算法和蚁群算法等。McMinn^[20]分析了搜索算法在结构测试、功能测试、非功能测试等方面的应用，以及未来的主要发展趋势。目前的搜索算法还多用于基于结构的测试数据生成，受限于搜索问题定义、搜索空间规模等问题，在功能和非功能测试中的有效应用还有待于进一步研究。

基于模型检测的软件测试研究是一个新兴的关注点^[23,24]。静态分析和形式化验证的方法采用状态机模型表示系统的行为，用模态/时序逻辑公式描述系统的性质，通过搜索系统的状态空间来确认系统是否具有所希望的性质（或不具有不希望的性质）。Callahan 和 Engels 最早提出了采用模型检测工具自动生成测试用例的方法，其基本思想是采用反例（counterexample）作为测试用例，而如何系统、高效且自动化地生成反例，就成为基于模型检测测试生成的关键问题。近年来，有学者从多个角度研究基于模型检测的测试问题，如测试用例生成、评估、覆盖率分析等，取得了大量的研究成果。

与形式化方法的理论研究相对应，UML 作为工业界的事事实标准，基于 UML 半形式化模型的测试方法得到了广泛的关注和实际应用，例如基于 UML 用例图、序列图、状态图、协作图等模型，在单元测试、集成测试不同测试层次上，针对对象、构件、接口等不同被测对象的测试用例生成技术。2004 年，国际标准化组织 OMG（Object Management Group）正式推出了 U2 TP（UML 2.0 Testing Profile）测试建模语言，用于描述测试的主要概念模型及相互关系，如测试体系架构、测试行为、测试数据和测试时间等。基于 UML 的测试也是国内学者的一个研究热点，例如，X. Li 等人^[25]提出采用 UML 2.0 的 IOD（Interaction Overview Diagram）和序列模型（Sequence Diagram），从异常处理一致性（Exceptional Consistency）和强制一致性（Mandatory Consistency）两个方面，建模和验证 Java 程序的运行时特性。

基于模型测试的主要目标是通过自动化的测试生成技术，大幅度降低测试代价；但由于建模的代价可能过高，且需要特殊的专业背景，直到近些年来才见诸工具和工业应用，而与实际问题结合的推广应用也成为基于模型测试所面临的主要挑战^[26]。在几十年的研究中，提出了多种模型，如迁移系统、前置/后置条件规约、场景模型等，如何统一不同风格的模型成为一个 important 问题。Grieskamp^[27]以微软软件测试为背景，提出了多范型融合的一种方法，构建一体化环境，可以集成来自于不同阶段、不同语言的模型。

2.1.2 测试用例优化选择

测试通常需要耗费大量的人力和时间，在被测系统庞大而测试资源有限的情况下，可以采用测试用例排序等方法，优化选择测试用例子集，达到测试目标。测试用例优化选择主要包括对测试用例集最小化、排序和选择。测试用例集最小化以去除冗余的测试用例为目标，期望在保证覆盖率的前提下，缩减测试用例集规模。测试用例排序通过对

测试用例排列优先执行顺序，以便在测试中途停止情况下能获得最好的测试效果。

测试用例集最小化关键研究内容主要集中于如何定义测试覆盖率目标以及如何排除冗余的测试用例；测试用例排序主要研究内容主要集中在如何建立测试用例重要程度的度量；回归测试时进行的测试用例选择研究关键在于如何度量以往版本与当前版本程序的不同处，并据此选择测试用例进行覆盖。

测试用例的覆盖率、错误检测能力、测试代价和资源约束是优化的主要依据。测试覆盖率可以分别针对黑盒（如接口覆盖率、需求覆盖率）和白盒（代码覆盖率）测试选取；错误检测能力可以根据不同的错误类型和严重程度进行分类；测试代价包括选择算法时间、测试执行时间等；资源约束包括时间和软硬件资源的可用性等。通常有两种主要的分析方法：基于测试执行的历史数据和基于静态分析。

1) 基于静态分析，对测试用例进行排序。H. Mei 等人^[28]提出了一种基于程序调用图的方法，估算 JUnit 中测试用例的覆盖率，并对测试用例进行排序。研究针对 4 个 2000 ~ 80000 行代码 Java 程序的 19 个不同版本进行实证分析，选取了计算策略、覆盖率标准、测试级别三个控制条件，与基于动态覆盖的测试用例排序方法进行比较，取得了较好的优化效果。L-J Mei 等人^[29]针对 Web 服务黑盒测试，采用服务接口描述 WSDL 中 XML 的标志覆盖率作为评价准则，对测试用例进行排序。

2) 基于动态测试执行的历史数据，选择测试用例子集。Y-C Huang 等人^[30]提出了一种改进的代价模型，根据测试历史数据判定测试用例的故障检测能力，采用遗传算法按照检测能力的高低对测试用例进行排序。Z-Y. Wang 等人^[31]提出了一组组合测试覆盖率准则，结合组合权重和测试代价，综合评价和排列测试用例。K. Zhai 等人^[32]针对位置服务，提出了基于 Point-of-Interest (POI) 覆盖率的测试用例选择算法。C. Zhang 等人^[33]提出了一种根据测试用例的执行剖面对其进行聚类分析的方法，不同类别的测试用例具有不同的错误检测能力，从而指导用例选择过程。L. Zhang 等人^[34]分析了测试用例的时间代价，采用整数线性规划算法选择测试用例。

S. Huang 等人^[35]结合两种方法，提出了一种基于端到端程序行为的回归测试用例选择方法。该研究根据分析测试用例对原始版本的调用关系获取其代码覆盖图，比较新的待测程序与原始版本并标识发生变化的程序代码，评估并选择对修改部分代码覆盖率高的测试用例集用于回归测试。Y. Zou 等人^[36]提出了一种结合 Web 界面元素和代码相结合的混合覆盖准则。动态 Web 应用的界面元素通过服务端脚本动态生成。单一覆盖准则难以引导有效的测试用例优化选择。实验表明前端 Web 界面元素与后端代码的混合覆盖能够比单一覆盖更加有效。S. Sampath 等人^[37]提出一种通用的混合覆盖框架，该框架将已有的混合策略分为三种类型，并且通过若干 GUI 程序的测试优化选择验证了其有效性。

在用例选择上，多数采用各类搜索算法，S. Li 等人^[38]通过模拟生成测试用例集合，以 APRCI (Average Percentage Requirement Coverage Improved) 作为评价标准，分析了主要的搜索算法在回归测试用例选择中的应用效果。D. Hao 等人^[39]提出了一种基于统计分析的方法，将优化目标定义为：以 $c\%$ 的置信度保证缺陷发现率不低于 $L\%$ ，而在这一置信度条件下，寻找能够最大程度减少用例数目的测试用例子集。C. Fang 等人^[40]通过分

析 8 种现有逻辑覆盖准则引导测试用例优化选择结果，验证了在实际应用中这些逻辑覆盖准则之间的强弱关系^[41]，同时指出了传统测试优先级度量 APFD 可能存在的缺陷。

2.1.3 测试充分性评估

测试充分性是指软件在有限测试数据时的表现能够代表软件在所有输入数据上的表现的性质。当测试充分性达到一定要求时，测试可以停止，从而有效地控制软件测试的质量和开销。对于测试集测试充分性的度量是相对于某个具体准则的。测试充分性准则的谓词形式定义如下：

设 D 是一个可数集合，表示软件的输入、输出数据集合； P 是 D 上的可计算函数集合的子集，表示以 D 中的数据为输入和输出的程序集合； S 是 D 上二元关系的一个子集，表示软件的功能规约的集合； $T = 2^D$ 是测试数据集合的集合。那么软件测试的充分性准则 C 是一个定义在 $P \times S \times T$ 上的谓词，即 $C: P \times S \times T \rightarrow \{\text{true}, \text{false}\}$ ， $C(p, s, t) = \text{true}$ 表示用测试数据集 t 对功能规约 s 测试程序 p 是充分的；否则该测试是不充分的。

常用的测试充分性准则有控制流充分性准则、数据流充分性准则、基于程序文本的准则、基于故障的充分性准则等。覆盖率是用于定量度量测试充分性的一种重要方法。传统的覆盖率标准有白盒测试中的代码覆盖率、黑盒测试中的功能点覆盖、模型驱动的状态机图覆盖等。

近年来，随着被测软件的多样性和测试技术的发展，提出了一些新的充分性准则，以便更好地评估和优化测试，典型的研究成果有针对被测软件体系架构和程序设计语言特点，定义特征集的覆盖目标。例如针对软件图形化用户界面测试，L. Zhao 等人^[42]提出了一种基于事件处理的覆盖准则，它要求覆盖所有的事件处理及其之间的数据交互。L. Mei 等人^[43]提出针对 WSDL 接口描述 XML 的标签覆盖。

变异测试是一种基于故障的测试技术，经常用作评价测试充分性的手段。传统上的变异测试是一种白盒测试技术，生成若干称为变体的程序源代码。但变异测试也可以用在软件接口和规约上，用于评价黑盒测试充分性。针对软件的故障模型，研究有效的变异算子，是测试充分性评估的一个关注点。例如，C. Ji 等人^[44]提出了 Java 异常处理测试变异分析方法，根据语义异常层级和异常调用层级，定义了 5 种变异操作，构造出引发最多异常类型的高质量测试用例。W. Xu 等人^[45]提出了针对 XML 的数据变异方法。S. Lee 等人^[46]提出了针对语义 Web 服务，基于 OWL-S 规约的变异测试方法。Jia 和 Harman^[47]从变异算子选择和优化、变异测试应用、实验评估等方面详细分析了变异测试 30 余年的技术发展。

软件测试使用覆盖准则存在一个隐含假设：高覆盖的测试数据集比低覆盖的测试数据集具有更强的故障检测能力。然而，无论在学术界还是工业界，覆盖作为测试充分性仍具有争议。最近的一个实验^[48]表明，在安全攸关软件中，即使满足严格的 MC/DC 准则，仍存在较多的故障未被检测。鉴于覆盖准则和变异分析难以克服的缺点，研究人员开始尝试寻找一些经济高效的测试充分性评估方法。早期一些研究人员发现测试数据多样性与系统容错能力的关系^[49]。近年来，测试数据多样性开始得到研究人员的重视和尝

试^[50,51]。测试数据多样性是指测试数据分散在程序行为空间中的均匀程度，通过提高数据多样性可以达到提高故障检测能力的目的。如何通过测试数据多样性制定测试充分性准则仍需要进一步的研究。

2.2 缺陷定位与预测技术

分析测试结果，准确定位软件故障发生的位置，预测系统的质量，是软件测试研究的重要内容。

2.2.1 故障定位

故障诊断的目标是根据观测到的系统症状，快速、准确地推理出故障假设。通常可以分为两种类型：一种是确定性故障检测，故障发生必然表现出相关症状，典型的研究如基于规则的专家系统故障诊断方法；另一种是不确定性故障诊断（概率性故障诊断），即故障以一定的概率引发存在依赖关系的症状异常，典型的研究如基于神经网络、不确定推理等技术的故障诊断方法。

基于大量的测试覆盖率信息，采用概率统计模型和数据挖掘算法进行故障定位，是近年来研究的一个热点。给定故障程序和测试用例集，故障定位根据测试用例的覆盖等信息，给出可能出错的程序体的怀疑度排序。典型的方法就是对比程序的成功运行和失效运行的测试用例所覆盖的语句集合，发现失效运行中偏离的语句集合，计算和度量其可能出错的概率（即可疑度）。A. Gonzalez-Sanchez 等人^[52]使用贝叶斯推理对语句包含错误的可能程度进行概率建模，提出了一种信息增益的测试用例优先排序技术（IG）。进而提出了一种通过模糊分组（ambiguity group）约简的 RAPTOR 技术^[53]，根据覆盖率信息进行语句分组，对每个组中包含错误的程度进行概率建模。L. Gong 等人^[54]使用线性模型对语句在测试用例选择过程中排名的变化进行拟合。

假设检验是一种广泛应用的分析方法，即假设程序特征服从特定的概率分布（如高斯分布），这些假设的正确性常常会影响到故障定位的准确性。Z. Zhang 等人^[55]提出了基于谓词的使用无参数假设检验的故障定位框架，并选择 Mann-Whitney 检验和 Wilcoxon Signed-rank 检验作为无参数假设检验方法，取得了较好的定位效果。

程序特征谱分析及故障定位（Spectrum-based Fault Localization）是一种轻量级的自动化故障定位方法。程序特征谱是在软件运行过程中，通过插桩等方式，收集到的一组运行剖面数据，为软件动态行为分析和比较提供依据。M. J. Harrold^[56]、R. Abreu^[57]等人分析了多种程序特征谱技术，如分支图谱（branch spectra）、路径图谱（path spectra）、完全路径图谱（complete-path spectrum）、数据依赖图谱（data-dependent spectra）、输出图谱（output spectrum）、执行跟踪图谱（execution-trace spectrum）等，并通过实证分析了图谱在揭示程序动态行为、辅助故障分析中的有效性。基于特征谱的故障分析，就是通过比较正确和异常执行测试用例的图谱，计算每个程序单元发生故障的概率。通常采用数据聚类分析的方法，计算图谱向量之间相似性，给出可疑度排序准则。可疑度计算

是这种方法的一个关键问题，典型的计算方法有 Taratula、Ochiai 和 Jaccard 等，目前已经有多个开源工具支持程序特征谱分析，如 Pinpoint、Tarantula 和 AMPLE 等。

Naish 等人^[58]调研了 30 多种相似系数并将这些系数作为概率统计模型用于故障的定位，分析了两种故障与程序逻辑和测试数据的两种相关模式——必然条件触发和偶然条件触发，并提出了两种理想度量可疑度公式。K. Yu^[59]提出了使用多程序谱模型进行错误定位的技术 LOUPE，利用多种程序特征谱模型，考虑控制与数据依赖关系捕捉程序的异常行为，针对两种缺陷类型分别建立可疑度计算模型，整合计算可疑度大小。X. Xie 等人^[60]进一步将故障定位的风险评估函数归为 ER1、ER2、ER3、ER4、ER5、ER6 六类等价的风险评估函数和一些非等价类别，并理论证明了 ER1、ER5 是单故障环境下故障定位效果最好的。Y. Miao 等人^[61]分析了偶然正确测试用例给故障定位带来的分析，并提出了一种基于测试聚类的偶然正确识别方法，能够在一定程度改进故障定位效果。

互联网的动态性和不确定性，给网络软件的故障诊断带来了更多的挑战性问题，也成为一个新的研究热点。

2.2.2 缺陷预测

缺陷预测是一种主动式的质量保障方法，能够在软件开发的早期预测出软件模块出错的倾向，从而找到相应的解决方案。近期一些研究表明，如果采用鲁棒的预测模型，采用故障预测方法，可以达到 71% 的故障识别概率，而采用软件评审方法，故障识别概率仅为 60%。因此，缺陷预测有可能获得更佳的成本效益。C. Catal^[62]分析了 1990—2009 年间 90 篇相关论文，指出了缺陷预测的两个主要发展趋势。一是当缺少缺陷的历史数据（例如新项目启动）时，如何根据已有软件的故障数据，对新系统进行预测。Catal 等人提出了模块聚类方法，采用 X-means 聚类、阈值控制、模糊聚类、K-means 聚类等方法，对实际样本数据进行分析，取得了较好的预测效果。二是当信息不够完整，例如项目中部分模块缺少缺陷数据，如何根据已有模块数据推算未知的模块缺陷信息。Catal 利用朴素贝叶斯算法，提出了半监督学习的缺陷预测模型。基于机器学习的故障预测方法也在国内学者研究中得到了关注，例如，M. Li 等人^[63]分析比较了三种缺陷预测方法：基于普通机器学习方法的随机取样，基于半监督学习方法的随机取样，以及基于半监督学习方法的主观取样方法。T. Hall 等人^[64]对 2000—2010 年间 208 篇软件缺陷预测相关论文做了详细的分析，一方面发现简单的机器学习方法（如朴素贝叶斯和 logistic 回归）通常具有较好的预测效果，另一方面发现特征选择有助于提高缺陷预测模型的性能。据此，他们认为预测模型建立方法对缺陷预测模型的准确性具有较大的影响。

2.3 面向特定领域的测试

2.3.1 嵌入式系统测试

根据 IEEE（国际电器和电子工程师协会）的定义，嵌入式系统是“用于控制、监视

或者辅助操作机器和设备的装置”，即嵌入式系统是软件和硬件的综合体，是应用于特定环境下，针对特定用途设计的系统。现代嵌入式系统在航空、航天、工业控制、民用电子等领域已得到广泛应用。嵌入式软件广泛存在于嵌入式系统各设备平台之上。软件调用平台硬件、基础软件资源，进行平台内、平台间的交互与协同，构成嵌入式软件系统。嵌入式软件与硬件紧密结合，软件特性与硬件特性紧密相关。一般系统的硬件资源（如处理器、存储器等）非常有限，具有系统内核小、专用性强、系统精简、实时可靠等特点。例如，嵌入式软件任务的执行必须满足严格的时间约束，例如时间点、执行区间、时间同步、执行速率、触发周期等。违反时间约束可能会带来严重的灾难和损失；嵌入式软件的性能极大地受制于嵌入式系统平台的可用资源状况，任务软件之间通常通过同步、异步、并发等关系，共享硬件和平台资源，交互协同，资源分配和调度的性能与合理性将直接影响任务的可用性和可靠性。

嵌入式系统经过 50 余年的发展，其发展轨迹呈现出硬件和软件交替发展的双螺旋式，先后经过了无操作系统阶段、简单操作系统阶段、实时操作系统阶段和分布式互联网阶段。随着硬件性能的提高，嵌入式软件规模越来越大，软件担负的任务和完成的功能越来越多。在现代嵌入式系统中，软件负责完成的任务在总任务中所占比例越来越高，且多是复杂、核心任务，如算法执行、系统调度等。平台硬件日趋同质化，硬件的设计与开发技术相对成熟。因此，软件的设计与开发已成为嵌入式系统设计与开发的主要问题。然而，嵌入式软件开发仍缺乏有效的方法和工具支持，主要采用以程序员为核心的手工作坊方式，受限于人员的经验和水平，使得开发工作量大，周期长，且软件质量难以保证。例如，据统计，在我国核电领域软件开发中，平均每 100 行代码需要 4 人天的工作量，而代码缺陷数仍高达 80 个/千行。验证和确认是软件缺陷检测的主要手段，在嵌入式软件中，尚存在以下挑战性问题：

- 系统规模大，可信性要求高。嵌入式系统正逐渐由硬件器件密集型转向软件密集型，软件在嵌入式系统中发挥着越来越重要的作用，软件规模和复杂度大幅度增长。例如美国 F-22 飞机航电系统软件由 170 万行源代码编制而成，是首次在战斗机上大规模使用软件；F-35 飞机航电系统软件更是达到 600 万行以上。这些代码具有极高的安全性和可靠性要求，例如，根据 DO178B 适航认证要求达到 100% 代码 MC/DC 覆盖。如果没有行之有效的验证和确认，软件开发将会面临极其严峻的挑战。
- 环境相关度高。嵌入式系统是一个软硬件的统一整体，强调软硬件的协同和集成性。在特定应用中，嵌入式软件对嵌入式硬件具有较强的依赖性。这就使得我们在对嵌入式软件进行设计和验证时，必须综合考虑嵌入式硬件，而硬件的连续性使得建模和验证变得复杂。嵌入式软件通过嵌入式硬件感知和改变环境，在考虑嵌入式软件的可信性时，嵌入式硬件则成为嵌入式软件的环境的一部分。环境的不确定性（如电磁干扰、极端环境）增加了嵌入式软件建模和验证的难度。
- 验证和确认效率低。从代码、模块、模块组合到系统集成，嵌入式软件验证和确认分为多个阶段，每个阶段的工作量都很大。很多时候搭建一个组件的测试环境比开发这个组件花费更大。此外，在嵌入式系统中有很多错误只有当某些特定的

事件按照一定的序列发生后才会出现，而且不可重复。如果用测试的方法来找这类错误就需要尝试很多测试用例，由于不可重复性，这些测试用例的价值很有限。但目前缺乏有效的工具支持，主要采用人工方式，成本高，效率低下。

Ebert 和 Jones^[65]采用 SPR 数据库中案例，分析了嵌入式软件缺陷防范、验证和确认的典型开发模式，如表 1 所示。实际工程数据显示，60% 的系统缺陷来自于 20% 的组件，但根据组件的特点分布不同；90% 的失效来自于 10% 的代码；20% 的缺陷需要 60% ~ 80% 的修复代价。在集成测试和系统测试阶段，往往要耗费 30% ~ 40% 的嵌入式系统开发资源，占整个项目周期的 15% ~ 50% 的时间。在今后数年中，嵌入式软件将保持每年 10% ~ 30% 的年增长率，复杂度的增长势必导致更加严峻的系统质量问题。对于嵌入式系统而言，防危性（safety）、安全性（security）和性能（performance）是测试的重点问题，并且通常情况下，三者相互制约 - 密切相关，难以独立地设计或者测试，需要统一考虑。

表 1 嵌入式系统测试与验证工作量统计表^[65]

测试活动	任务分配（功能点）	工作效率（功能点每月）	缺陷移除效率（%）	修复引入的缺陷率（%）
人工评审				
设计审查	1 000	160	85.0	4.0
代码审查	200	60	85.0	4.0
质量功能部署	1 000	200	82.0	3.0
测试计划审查	750	125	80.0	5.0
测试脚本审查	300	175	78.0	4.0
文档评审	3 500	1 000	77.0	2.5
结对编程	2 500	200	75.0	5.0
缺陷修复审查	300	90	70.0	3.0
质量保证评审	2 500	750	45.0	7.0
手工测试				
子程序测试	5	100	50.0	2.0
组件测试	1 250	150	40.0	3.0
系统个测试	2 000	200	40.0	7.0
新功能测试	125	110	35.0	5.0
回归测试	150	150	30.0	7.0
单元测试	50	90	25.0	4.0
自动测试				
静态代码分析	15 000	10 000	87.0	1.0
系统测试	500	200	40.0	8.0
回归测试	500	175	37.0	7.0
单元测试	500	250	35.0	4.0
新功能测试	500	200	35.0	5.0
认证				
可用性测试	5 000	2 000	65.0	4.0
试运行测试	15 000	500	45.0	5.0
接收测试	5 000	3 000	40.0	7.0

模型驱动的方法，可以在系统开发的早期引入系统规范的质量控制手段，有助于尽早发现和排除错误，是一个重要的发展方向。国际上，在嵌入式领域，SCADE 是一个应用于安全关键领域的典型的模型驱动开发平台。采用 Lustre 同步数据流描述语言，并可支持 C 语言程序的代码自动生成。A. Lakehal 和 Parisis^[66]探讨了 Lustre 程序覆盖率与 C 语言结构覆盖率之间的对应关系，如基本路径覆盖、基本条件覆盖、多条件覆盖等，并进一步探讨了针对 Lustre 的操作算子（如 when 和 current 时间算子）的覆盖率标准。在此基础上，建立了测试环境，构建了 Lustre 测试工具 Lutess，可自动生成测试数据和测试用例。

2.3.2 服务化软件测试

软件服务化已经成为互联网时代软件发展的主要趋势。基于服务的软件开发更加强调基于标准接口、具有一定自治计算能力的大颗粒度功能组件的重用和组装，从以编码为核心转化为侧重于建模及模型自动转化的开发过程。与传统软件相比，基于服务的软件通过标准服务接口及服务访问协议，屏蔽了软件内部的复杂性和软件之间在程序设计语言、运行平台、数据结构等多方面的异构性，使得软件之间的动态协同成为可能。基于面向服务体系架构 SOA，软件系统的构建逐步转化为模型驱动的自动化过程，从业务模型自动转化为软件组合的架构、行为模型，再通过服务（数据）的发现和动态绑定技术，实现模型的实例化过程，并最终形成可运行的软件系统。在由服务发布者、服务中介、服务使用者、基础设施提供者等多方协同构成的体系架构中，“服务”可以在开放的环境中，动态地发布、发现、组装和绑定，不依赖于任何特定的平台。在诸如 Mash-up、动态协同、服务总线等新技术的支持下，软件工程的重点不再是编码，而是在更高抽象层次上的建模及组装技术。

开放性、动态性为基于服务的软件系统带来巨大潜能，也对系统的可信性保障机制带来新的需求。服务在开放的环境中发布、运行，服务的开发过程与使用过程相分离，一方面，服务开发者难以预测各种使用场景和应用环境；另一方面，服务的使用者难以控制服务开发的质量，软件的失效风险较高且失效影响范围广。例如，Gmail 在 2008 年 11 月 8 日发生失效，1.5 个小时邮件服务不可用，数以百万人受到影响。开放性使得系统的故障原因更加难以预测，难以建立有效的管理控制机制；动态性使得系统的失效模式更加复杂，难以恢复。

对服务化软件可信性的分析和软件测试的研究始于 2002 年。早期研究中，Bloomberg^[67]将服务软件测试分为三个主要阶段：2002—2003 年，侧重于基于服务规约的单元测试；2003—2005 年，服务的发布、发现、绑定等能力测试；2004 年之后，关注于 Web 服务的动态运行时能力测试。Canfora 和 Di Penta^[68]根据测试类型，从功能测试、非功能测试、回归测试、集成测试等几个方面，分析 Web 服务的测试研究现状。Bozkurt 等^[69]在此基础上进一步全面分析了服务测试的主要特点和挑战性问题，已有的测试方法如何应用到 Web 服务测试当中，以及新兴的测试技术。

标准服务接口是服务之间编制、编排、协同的基础。接口建立了服务之间的契约关

系，服务的接口、组合和语义规约可视为服务的预期行为描述，因此，各种协议分析和建模技术，以及基于模型的测试技术广泛地应用于服务化软件的测试。典型的研究成果有：

- 模型检测。采用模型检测技术验证 Web 服务的相关协议和描述，主要是将 Web 服务的规约（如 WSFL、BPEL4 WS）转化成形式化的模型（如有限状态机模型，Petri 网模型等），借助模型检测器（Model Checker）等自动化工具，验证协议和服务描述的完整性和一致性。其中，服务规约的建立及特性的描述、形式化模型的转换及选择模型检测器对特定特性的验证是模型检测技术实现的关键。通过选用不同的服务规约及模型检测器，基于模型检测技术的服务验证取得了一定的研究成果，如 Nakajima 采用模型检测器 SPIN^[70]，Foster 等提出了一种基于有限状态机的服务流描述语言和验证机制^[71]，Narayanan 和 McIlraith^[72]采用 Petri 网模型描述服务流，M. Koshkina 等人^[73]引入了基于 BPEL4WS 的 BPE 演算。H. Huang 等人^[74]将模型检测与测试技术相结合，一方面，将 OWL-S 规约转化成类 C 代码的描述，采用扩展的 BLAST 模型检测器对其并发性进行检测；另一方面，将部分需要检测的特性嵌入到转化的 C 代码描述中，并根据模型检测结果自动生成测试用例。
- 契约模型。契约作为软件的一部分，定义了功能的输入、输出等约束条件，不仅有助于提高程序设计质量，还能提高软件测试的自动化程度。基于契约的设计思想已被广泛应用于面向对象的软件设计以及基于组件的软件开发技术中。将契约设计用于软件测试，可以从可观察性（错误能够被发现的概率）和可诊断性（定位错误以及隔离错误的便捷性）两个方面提高系统的可测试性。在 SOA 中，服务提供与服务应用即可定义为契约关系。Bertolino^[75]将 UML2 中的协议状态机描述引入 Web 服务描述 WSDL 中，增加对服务前置/后置条件以及数据/进程约束的描述，增强了 Web 服务的可验证性。Heckel^[76]指出在 SOA 中契约描述存在多种表达方式：实现级、XML 级和模型级契约描述。Bruno^[77]提出将测试用例作为服务提供商和用户之间的契约，令用户能够利用提供商提供的测试用例对服务进行回归测试，验证 Web 服务新、旧版本的一致性。X. Bai 等人^[78]提出了扩展 UDDI 中间件模式和 Test Broker 模式，在多个层次上建立测试契约关系，并探讨了基于契约的测试用例生成技术。
- 接口及组合测试。以 Web 服务协议栈为 SOA 典型实现，先后有学者在传统的数据流测试、控制流测试的基础上进行改进，提出了基于 WSDL 接口模型、基于 BPEL 的服务组合模型，以及基于 OWL-S 的语义服务模型的测试方法。Tsai^[79]提出进一步扩展接口描述以增强其可测试性，扩展机制，如输入/输出依赖、服务调用序列、并发序列，以及功能层次关系等。X. Bai 等人^[80]分析了 WSDL 的服务接口，从 XML-Schema 数据类型和功能依赖两方面，讨论了测试数据和测试流程的生成方法。B. Li 等人^[81]提出了一种扩展 BPEL 流图 XBFG，在服务组合路径上附着消息数据，分析服务流程、绑定和接口的多种变化模式，提出了一种多版本

BPEL 测试回归测试用例选择算法。C. Sun 等人^[82]提出了一种基于 BPEL 的控制流程分析的面向场景的测试方法。S-S Hou 等人^[83]提出了从 BPEL 的数据流分析中提取并生成消息序列图的方法。L-J Mei 等人^[84]提出了基于 XPath 的数据流分析方法。X. Bai 等人^[85]进一步提出了基于 OWL-S 的语义服务测试方法，探讨了本体的分区测试数据生成、基于 OWL-S 的 IOPE 分析的测试路径生成等技术。C. Ma 等人^[86]提取语义 Web 服务的操作的前置、后置条件，采用 Stream X-machine 建立测试模型。

基于故障的测试也是一种典型的测试方法，根据故障模式分析，将特定类型的故障植入软件中，以检测测试用例集对故障的检测能力和有效性。故障植入可以通过人工植入、变异测试或者增加扰动的方式。W. Xu 等人^[45]分析了服务中间件层的故障模型，如服务挂起、服务中断、数据崩溃、消息重复、消息丢失和延迟等各种可能故障，通过 SOAP 编程接口挂接故障注入器，将错误植入通信层，检验在不可靠的网络通信中服务的可靠性。S. Lee 等人^[46]分析了语义 Web 服务的故障，提出 OWL-S 的变异测试方法，设计了针对 OWL-S 规约的变异算子。

针对服务动态、多样、协同的特点，W-T. Tsai 等人^[87~89]提出了协同验证与确认（CV&V, Collaborative Verification and Validation）的测试框架，在服务注册中扩展服务质量管理机制，引入检入/检出测试流程；提出基于群测试理论，根据测试结果动态评估测试用例的故障检测能力，并进行优先排序；提出基于投票（voting）机制，设定测试判定（Test Oracle）。服务的使用与服务的运行环境、最终用户等非功能因素密切相关，因此服务质量评价应综合考虑各种定性、定量的质量特性。

服务软件测试的另一个重要的发展趋势是基于云平台的服务测试。基于云计算平台的系统结构提出了多层次的资源共享模式，如基础设施即服务（IaaS）、平台即服务（PaaS）、数据即服务（DaaS）、软件即服务（SaaS）等，使得软件系统从买卖方式转化为依托大量开放资源的租赁方式。云平台软件具有高可伸缩性、多租户体系架构等特点，资源租用、远程部署的方式造成服务可测试性较低，系统运行环境多样且多变，这些特点引发了新的云测试技术的研究，例如，基于云平台的分布式并发测试执行和分析、基于 SLA 的元数据驱动测试生成、大规模动态负载模拟等。J. Gao 等人^[90,91]分析了云测试的主要问题、挑战、需求和主要的实践，并提出了基于云平台 TaaS 系统以用于 SaaS 的性能测试与评估。Riungu 等人^[92,93]从应用、技术、管理、费用等方面分析了云计算中软件测试的若干研究问题，并通过对典型测试服务商的访谈，分析了实际应用中的云测试的机遇与挑战。

利用云平台资源虚拟化、共享服务等优势，可构建新型的测试系统。Ciortea 等人^[94]提出了一种基于云平台的符号执行系统 Cloud9，Parveen 等人^[95]将 JUnit 测试框架移植到云平台构成 HadoopUnit 系统，Oriol 等人^[96]构建了基于 MapReduce 的自动化随机测试系统 YETI。Grossman 等人^[97]构建了开放云测试床 OCT，截至 2009 年已连接北美 4 个数据中心的 120 个结点，集成了包括 Eucalyptus、CloudStore、Hadoop、Sector/Sphere 和 Thrift 在内的多种异构云平台。L. Yu 等人^[98]分析了基于云平台的测试任务管理和调度问题，

提出构建基于 Xen 的 TaaS 平台。X. Bai 等人^[99]调研了云测试工具的发展现状，并构建了 Vee@ Cloud 测试系统以支持动态可伸缩测试资源配置^[100]。J. Wu 等人^[101]构建了一种集成多种异构资源的云测试平台 Agaric，并提出测试控制流程协议。

基于云平台的服务测试也催生了新的测试模式，如测试即服务（Testing- as- a- Service, TaaS）、众包测试（Testing by crowdsourcing）。TaaS^[102,103]指基于云平台的测试系统构建和测试服务提供方式。从用户角度看，TaaS 允许测试群体以协同方式开发和共享各类测试资产（如测试用例、脚本、数据、环境、缺陷报告等），利用云平台计算和存储资源在分布式环境下部署测试，可以支持大规模动态可伸缩的测试运行和结果分析。TaaS 系统的构建也可以采用 SaaS 软件多租户体系架构、元数据驱动等技术。众包测试，如 uTes（www.utest.com）和 Mob4 Hire（www.mob4hire.com），采用众包开发的方式，基于 Web 2.0 群体智能的思想，将测试设计、脚本开发、测试执行等任务分别打包并发布到测试社区，引入新的开放社区测试流程和评价方法，管理和控制测试质量。众包软件测试也开始得到学术界的关注^[104,105]，并尝试用来解决 GUI 测试和测试 Oracle 等疑难问题。

服务计算和云计算是近十年来新兴的计算范型，尚未形成一致的概念，其故障模型和检测方法还处于早期探索研究阶段。

3 国内研究进展

软件测试一直是国内较为活跃的一个领域，在中国计算机学会软件工程专委会成立了软件分析与测试学组（<http://software.nju.edu.cn/zchen/sat/>），工业界也有大量的第三方测试机构和从业人员，各类学术研讨会和技术论坛等活动也十分活跃，形成了较好的业界与学术界互动的科研环境，测试研究和实践得到了越来越广泛的关注。近年来，国内在软件测试技术的研究方面十分活跃，国内学者先后在软件分析和测试重要的国际会议（ICSE, ECOOP, ICST, ISSTA, FSE 等）和期刊（TSE, TOSEM 等）上，发表了大量的论文，除了前述部分代表性成果外，本节选列了国内在相关研究方向上的部分研究工作。

基于模型的测试是一个关注的热点，国内上海大学、南京大学、清华大学等学者，分别采用状态机及其扩展、契约模型、语义模型等建模方法，探讨了 Web 应用、构件软件、面向刻面的软件、面向服务的软件等不同软件范型的测试方法。上海大学刘攀等人^[106]分析了典型的基于 FSM 的测试方法，包括图遍历、W、W_p、UIO、DS 和 T 方法等，提出了基于前置条件和后置条件分析的测试序列冗余约简算法，保证在不降低测试序列的错误探测能力的前提下，减少测试用例的数量。研究进一步从测试成本、错误探测能力和冗余约简效率几个方面，对各种测试生成算法进行了实验验证比较，给出了基于 FSM 测试的经验建议。研究人员进一步探讨了应用状态机模型，对 Web 浏览器的交互过程进行分析、建模和测试^[107]。南京大学雷斌等人^[108]针对构建软件开发，研究了基于

FSM 的构件健壮性测试方法，在自动机的状态转移中，引入契约定义给出健壮性构件接口的卫式设计，提出了测试路径和测试用例的生成方法，并实现了原型工具 RoTesCo，提高对于系统健壮性错误的错误发现能力。浙江理工大学舒挺等人^[109]研究了基于扩展有限状态机（EFSM）模型，测试序列生成算法。由于 EFSM 模型包含丰富的变量和谓词条件，它们之间的冲突可能导致自动生成的测试序列不可执行。研究在 EFSM 变迁及变迁之间的关联关系分析的基础上，定义了一个邻接变迁关联图，提出了一种自适应 EFSM 可执行测试序列生成算法，基于自适应预测搜索函数启发式引导可达性分析树扩展生成可执行的测试序列。清华大学侯可佳等人^[110]提出了接口契约模型，以及基于本体的测试数据生成技术。借鉴语义 Web 服务及契约设计的思想，从两个方面增强和扩展服务操作的语义表达能力：利用本体建立数据和服务统一的概念模型；借助语义规则，建立数据和服务的约束模型。并采用搜索算法，探讨了基于本体及约束的测试数据生成技术。

在测试用例优化选择方面，北京大学学者基于静态分析、测试用例及故障数据挖掘等技术，形成了一系列的研究成果。H. Mei 等人^[28]提出了一种基于程序调用图的方法，取得了较好的优化效果。D. Hao 等人^[39]关注到测试用例集约简带来的错误检测能力降低风险，提出按需约简策略。通过分析和评估测试用例与测试需求之间的关系，对给定错误检测能力要求进行预测，并基于 ILP 进行最优的测试用例选择。L. Zhang 等人^[111]提出两种模型统一全局 (total) 和额外 (additional) 的测试用例优先级策略，通过建立全局策略与额外策略直接的谱系，以达到两种策略互补的目的。实验结果显示这种结合能够优于单一使用全局策略和额外策略。D. Hao 等人^[112]关注到在软件演化中进行测试用例维护的困难。当测试用例在软件新版本失效时，通过训练回归测试中的失效实例，构建分类器进行失效原因分类：程序错误还是测试错误。该方法能够帮助工程师进行有效的测试维护。

南京大学学者在基于动态信息的测试用例优化方面做出了系列工作。S. Yan 等人^[113]通过测试数据的执行剖面（如函数调用序列）刻画数据特征，基于执行剖面定义测试用例之间距离（如欧氏距离）对所有测试数据进行聚类，通过类簇完成测试数据多样性分析。C. Zhang 等人^[114]将聚类分析技术引入回归测试选择中以进一步提高测试用例选择的精度。他们对测试执行剖面进行距离度量和刻画完成聚类，然后针对每个类簇进行抽样（如 5%），若某一抽样测试用例能够检测故障，则选择该类簇中的所有测试用例，最终构成一个较小的回归测试用例集。上述方法的结果很大程度依赖于测试聚类结果对多样性的刻画精度。Z. Chen 等人^[115]采用语句块覆盖记录执行剖面进行测试用例的聚类选择。他们采用程序切片技术来大幅度减少收集的语句块数目，从而有效降低测试用例的属性维度，提高聚类效率。实验结果表明该方法能够同时提高选择精度和故障检测测试用例的召回率。S. Chen 等人^[116]引入半监督聚类提高聚类精度，通过少量的测试数据成对标记信息 Must-link 和 Cannot-Link 生成新的距离空间，能够提高测试数据与故障检测的刻画精度。

组合测试是一种经典的测试数据生成技术。中科院软件所严俊等人^[117]深入分析了组合测试的原理和方法。南京大学聂长海等学者，在组合测试方面发表了一系列的研究成果。C. Nie 等人^[118,119]对组合测试的模型、测试生成技术、基于约束的组合测试、故障诊断、组合优化技术、度量评估、测试过程及应用等方面进行了深入分析。

另外，武汉大学潘伟丰等人^[120]采用复杂网络模型分析 Java 软件、测度类结点的复杂性来评价类发生和传播错误的概率，结合测试用例对类的覆盖率，对测试用例的重要性进行排序。

在测试充分性度量方面，中国科学院 B. Chen 等人^[121]提出一种元组密度（tuple density）的新型组合测试度量方法。与传统 N – 因素覆盖准则不同，元组密度提供一种更加灵活的高阶覆盖度量，在给定组合强度下追求更优的错误检测能力。清华大学安金霞等人^[122]综合考虑 10 余种测试覆盖率，提出了一种基于多维度测试覆盖率的软件测试动态评价方法，从测试监测信息、动态分析和评价模型、测试优化策略几个方面展开讨论，给出了综合测试覆盖率的定义和经验公式。上海大学刘攀等人^[123]依据程序中错误分布的 2 – 8 定律，将测试分为两个阶段：第一阶段产生覆盖有限状态机 FSM 中所有迁移的最小测试成本的测试用例集合（完备的、无冗余或冗余较少、测试用例数目少、测试用例总长度最短）；第二阶段针对测试第一阶段中发现的错误进行重点测试。

故障定位是近几年来国内研究的一个热点，国防科技大学、南京大学、中科院软件所等多所院校在此方面，产生了一系列的研究成果。

国防科技大学 Y. Lei 等人^[124]提出一种结合程序切片和频谱统计的故障定位模型。通过定制一种近似后向切片，刻画故障传播路径，能够进一步提升程序实体和故障之间的相关性分析。Y. Qi 等人^[125]反思了当前自动故障定位的实际作用，提出一种采用自动程序修复的评价新思路。采用自动修复的候选补丁包数量（NCP）作为自动故障定位技术的评估，并采用 11 个待测程序评估了 15 种自动故障定位技术，验证了该方法的有效性。Y. Zhang 等人^[126]提出一种新颖的符号执行技术 SSE，通过优化约束求解器调用来提升符号执行效率。SSE 只有在路径探索累计足够分支条件数量时才调用约束求解器。

南京大学 Y. Wang 等人^[127]引入测试执行频谱信息对测试执行剖面进行加权处理，能够进一步提高测试的聚类精度。在实际应用中，一个程序通常包含多个故障且同一测试有可能同时检测到多个故障。Y. Feng 等人^[128]引入多标签学习方法，能够精准刻画同一测试用例同时检测到多个故障的动态行为。基于此构建多标签聚类和分类方法，能够适应真实的多故障应用场景，为工业应用奠定基础。谭德贵^[129]提出了改进的错误定位算法，将失败测试用例的边际权重引入到可疑度计算的过程中。

中国科学院软件所 Z. Zhang 等人^[130]提出一种仅依赖失效测试的故障定位方法 FOnly。在当今在线系统收集大量错误报告的情景下，FOnly 能够建立有效的统计故障定位模型，并进一步降低计算成本。他们还提出了基于谓词的使用无参数的假设检验的故障定位框架，并选择 Mann-Whitney 检验和 Wilcoxon Signed-rank 检验作为无参数假设检验方法，取得了较好的定位效果^[55]。针对假设检验模型，北京航空航天大学叶钢等人^[131]分析西门子测试集中程序失败用例的规模和谓词评估偏差的分布，实验表明中心极限定理的前提和参数假设检验的假设条件无法得到普遍满足。因此提出了基于非参数假设检验定位程序缺陷的方法，比较了基于 2 种常用的非参数假设检验模型（Kolmogorov-Smirnov 检验和 Cramer-von-Mises 准则）的缺陷定位方法的效率。

此外，上海交通大学 H. Shen 等人^[132]提出一种称为 EFindBugs 的静态分析新方法，

采用两阶段错误排序方法降低静态分析的误报率，并在常用的 Java 程序中验证了 EFind-Bugs 的错误检测效率。解放军理工大学战伟等人^[133]探讨了将特征谱分析技术用于整数溢出错误定位，构建了 INTRank 模型，通过收集程序特征谱、计算分支和定义使用对的可疑度、计算错误传播率、建立语句可疑度之间的线性关系等步骤，取得了较低的漏报率。北京邮电大学褚灵伟等人^[134]提出采用动态贝叶斯网络对依赖关系进行概率建模的方法，针对瞬态故障、间歇故障、持续故障等动态故障集，设计相应的故障诊断算法。

在缺陷预测方面，中国科学院软件所王青等人^[135]详细深入地分析了缺陷预测技术的主要起源、发展和当前所面临的挑战，并对主流的缺陷预测技术进行了分类讨论和比较。北京航空航天大学何智涛等人^[136,137]对此软件测试过程及其发现缺陷能力呈现一定的时序变化规律进行了案例研究，并指出软件测试过程的基本特征和约束，影响其缺陷发现规律的主要因素（测试者发现缺陷的能力、被测项目规模和复杂度及其版本演化程度、测试周期等），以及软件缺陷发现时序过程具有的周期性、震荡性和衰减性等规律，提出了一种描述软件缺陷发现时序过程特征的叠加双阻尼振荡模型（Accumulative Bi-Damped Oscillation Model，ABDOM），并展示了其对潜在缺陷的预测能力。

在嵌入式系统测试方面，以 NASA、Boeing、ESA、Airebus 为代表的组织和机构带动了该领域技术的发展，国内起步较晚，主要应用于国防电子领域。1992 年开始的载人航天工程在国防电子领域率先推行软件工程化，制订了一系列软件工程技术标准，重点加强了软件测试和第三方评测工作，成立了专门的第三方软件评测结构，形成了自查自测与第三方独立评测相互补充的有效体系^[138]。

随着各个行业加强对电子产品关注，嵌入式测试技术的研究近年来也得到了一定规模发展。北京航空航天大学在国防领域的密切合作过程中，积累了多年的实践经验和技术创新基础，形成了一系列的研究成果。刘辉、徐福等人^[139~141]提出了一种基于仿真器的嵌入式软件单元测试方法，引入交互式脚本执行机制以及分布式嵌入式系统仿真测试技术。殷永峰等人^[142]设计和实现了嵌入式软件测试描述语言的执行引擎，并给出基于 SBRMS (Segment Based Rate Monotonic Scheduling) 调度策略的测试执行过程。周庆等人^[143]针对综合模块化航电软件测试，提出了基于软件故障注入的灰盒仿真测试环境方案。

面向服务的软件测试是国内较早介入的研究领域之一。互联网软件的测试是一个重要的研究领域。梅宏等人^[144]分析了基于互联网的计算范型和应用模式的演变，回顾了从资源共享、人机交互、智能计算等各不同角度所提出的各种范型（如网格计算、普适计算、服务计算、语义网、Web 科学），并从软件技术角度提出了网构软件的概念和一种新型软件形态。计算范型和软件形态的发展，促使软件模型、运行支撑平台、软件开发方法、质量目标和保证技术，均发生相应的技术演变。清华大学白晓颖等人针对软件的自治性、协同性、动态性等特点，研究提高测试的自动化和智能化的方法和技术。在面向服务的体系架构中，有效引入基于契约的服务在线协同测试机制^[78]，从基于规约的测试自动生成^[46,80,85]、基于群测试理论的测试选择^[89]、基于多 Agent 技术的分布式测试^[145]等几个方面，探讨有效的服务测试体系架构、方法和技术。邓小鹏等人^[146]提出了面向 QoS 的 Web 服务组测试方法，将功能、性能、代价等满意度指标作为 QoS 的度量指

标，对一组具有同样功能的 Web 服务进行排序选择。成都电子科技大学廖军等人^[147]提出了基于 Pi - 演算的 Web 服务组合的描述和验证。东南大学李必信等人研究了组合服务的测试用例生成和选择技术^[81,148]，研究提出了基于扩展 BPEL 流图 XBFG、OWL-S 模型的回归测试方法和测试用例选择算法。服务组合测试是一个研究的关注点，北京科技大学 C. Sun 等人^[82]、北京大学 S-S Hou 等人^[83]也从不同规定角度，提出了多种 BPEL 的数据流和控制流分析方法，以及测试用例生成技术。清华大学白晓颖等进一步开展了基于云服务的软件测试技术研究^[90,91,99,100]。

4 国内外研究进展比较

软件测试技术研究的根本动力源于软件技术及其应用的发展。在软件技术快速发展的初期，随着软件规模和复杂度的快速增长，国外学者敏锐地发现了软件测试面临的根本性问题和挑战，即在理论上不存在能够确保软件正确无误或发现任何隐藏缺陷的理想方法，并且在实践上软件质量的提升与软件开发的进度和成本控制之间存在着矛盾和冲突。针对软件开发面临的这种困扰，以 Myers 等人为代表的一代学者明确指出了软件测试在软件开发中的重要作用和意义，及其局限性，并且提出了一系列软件测试的基本准则，以及必要的测试过程和测试方法，从而构建了软件测试研究和实践的基础。长期以来，一些国内外学者围绕着一些经典的软件测试问题开展了长期而深入的研究，比如测试用例生成和优化、组合测试、回归测试等。国内学者在这些经典问题的研究中也取得了一些重要的研究进展。

伴随着计算机软硬件技术、网络技术、移动通信技术、信息处理技术等的蓬勃发展，新的软件形态也不断出现。面向对象技术引入了类、继承、封装、消息等新概念和新的实现机制；构件化技术强调基于契约的软件构件定义及其复用机制；各种网络化应用改变了人们对软件的传统认识，出现了软件即服务（SaaS）、软件作为服务（SaaS）等概念，展现出软件的新形态和新特征，并引出了服务质量（QoS）等概念；移动通信、物联网、云计算、大数据等的出现，更是从多个不同层面拓展了软件的外延和内涵，赋予其更加丰富的特质和更加强大的能力。所有这些新技术和新应用赋予软件的新特点，都给软件测试带来了更多的挑战。尽管这些问题的提出大多出自于国外的学者，他们首先注意到软件发展的新趋势，指出由此带来的新问题，提出基本的解决方法，并通过持续深入的研究，发展出新的实用方法和支持工具。与此同时，国内学者也密切地关注国际上的新热点，同步地开展相关研究。

此外，随着国内软件与信息产业的发展，各种网络化应用、嵌入式软件、无线移动终端应用已进入公众的生活，自主研发的数百万乃至上千万行规模的大型复杂软件已不鲜见，从而直接促进了国内软件技术的研究和发展，特别是实用技术的研究。在软件测试技术研究方面亦如此。例如，面向网络化应用的软件可信性、信息安全性（security）、性能、可用性等关键属性的测试问题；面向航空、航天等众多领域中安全攸关软件的测试问题，特别是面向特定领域的嵌入式软件仿真测试技术，可靠性、防危性（safety）、实时性等关键质量属性的测试技术。围绕这些问题，国内的许多学者和科研人员有针对

性地开展了大量研究，突破了工程实践中面临的大量关键技术问题，开发了相应的软件测试工具，促进了应用的发展。在国内的各种学术期刊和会议上可以看到许多相关的论文。这些研究植根于自身的应用需求，蕴藏着持续发展的原动力，因此，将有力地促进我国软件测试技术的深入研究和发展。

5 进展趋势与展望

2000 年，ICSE 会议组织了软件工程未来发展方向的研讨，M. J. Harrold 报告了软件测试的发展路线图^[5]，如图 3 所示。路线图首先明确了与工业实践密切结合，是软件测试研究的一个重要特点。基础研究所形成的软件方法、技术，将在工业实践中进行实证分析和检验，并最终形成实用的测试方法、工具和过程，为开发高质量的软件提供支持。

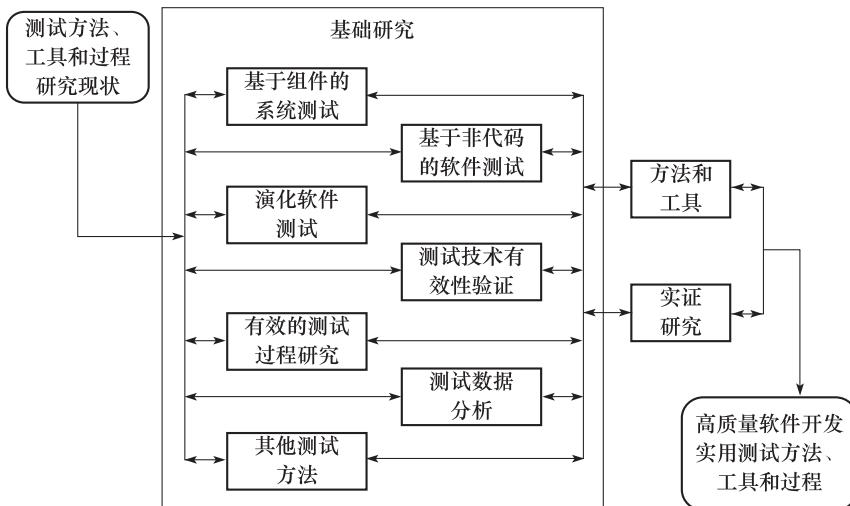


图 3 软件测试发展路线图^[5]

在基础研究方面，报告着重分析了 7 个方面的主要研究方向：

- 基于组件的系统测试。随着软件系统规模和复杂度的增长，基于组件的分布式系统已成为主流。组件（如货架软件）的生命周期不同于传统软件产品，具有组件发布和使用两个不同的阶段，组件提供者和使用者具有不同的需求和测试目标。由于组件可测试性和可控性难以保证，对组件测试方法、充分性评估准则等技术都提出了新的挑战。
- 基于规约的测试。Harrold 指出，随着测试工作的前移，测试设计可更多地依据代码之前的软件规约，如需求、设计和软件体系架构。基于形式化、半形式化甚至自然语言描述的规约，可以改进测试过程，提高测试用例设计质量。例如，Eickelmann、Richardson、Bertioli、Horrold、Stafford、Wolf 等学者就先后探讨了体系架构的可测试性，基于体系架构的单元、集成和回归测试方法，基于体系架

构的覆盖率标准，以及体系架构切片等技术。

- 演化系统的测试。在系统开发和维护的整个生命周期，需要不断地进行修改、完善、升级等演化工作，回归测试是验证和确认演化系统的重要手段，常常是代价最高的测试活动之一。如何针对系统变化选择有效的回归测试子集，提高回归测试的效率，降低测试代价，是测试研究的重要问题。
- 测试有效性分析评价方法。Goodenough 和 Gerhart 最早提出了测试充分性评估准则，之后多年有大量的学者从多个角度提出多种基于覆盖率的测试充分性评估方法。但研究表明，还需要进一步采用分析、统计和实证等方法，证明测试充分性准则之间及其与测试用例的缺陷检测能力之间的关系。
- 建立测试的有效过程。测试活动需要和软件开发、维护、分析以及其他质量保证活动结合，以保证高质量的软件开发。随着软件过程研究的开展，先后提出了多种软件测试过程模型，如 V 模型、软件测试过程成熟度模型、持续测试、测试驱动开发、每日构建测试等。对已有过程模型的分析验证，以及新模型的提出，都是测试研究的方法。
- 测试数据分析。测试过程中产生大量的数据，这些数据对于故障定位、软件质量评估与预测、测试过程改进多具有重要的意义。新的数据分析方法的应用与研究，有助于从海量的测试数据中，进一步挖掘有价值的信息和知识，成为测试研究的一个新的热点。

2007 年，A. Bertolino 在 FOSE (Future of Software Engineering)^[4] 研讨上，分析报告了软件测试的成果、挑战性问题和目标，如图 4 所示。

报告分析了测试的 4 个终极目标（“梦想”），即普适测试理论、基于测试的建模、100% 测试自动化，以及测试工程效用最大化。

在普适测试理论方面，从早期的数据流控制流分析、测试充分性准则研究开始，测试人员就一直努力寻求测试的基本理论。Bertolino 分析了测试理论所面临的几个挑战性问题：1) 明确测试假设 (explicit test hypothesis)。穷尽测试是不可能的，启发式方法是测试中常用的方法。启发式方法中常基于特定的前提假设，如等价类测试中假设针对每个等价类中的输入，软件具有相同的行为。借助形式化等方法，将测试中隐含的条件明确地表达出来，将有助于我们确定测试的边界，评价测试结果。2) 测试有效性评价。如 Harrold 报告，Bertolino 也同样认为有效性准则，以及不同准则之间的比较是一个重要问题。3) 组合测试 (compositional testing)。组件 + 组合的开发模式带来了新的质量问题，如何根据已有的组件测试结果，测试组件之间组合的可信属性，是近年来研究的一个关注点。例如，Blundell 等人^[149] 提出了基于假设 - 检验推理的构件组合验证，这种形式化验证方法可延伸至组合测试方法，即根据独立组件的行为推理组件组合行为。Bijl、Rensink 和 Tretmans^[150,151] 提出 ioco 测试理论，采用标记迁移系统分析组件之间的并行组合。4) 实证分析。实验验证在软件工程研究中起着重要的作用，实证研究也在软件工程的各个研究方向上得到了越来越广泛的关注。Juristo 等学者^[152] 从实验设计、数据分析、结果可重复性、领域应用 4 个方面，对随机测试、功能测试、控制流测试、数据流测试、变异测试、回归测试等测试技术进行了分析比较，认为测试仍缺乏形式化基础，超过一般的测试知识是基于直觉。

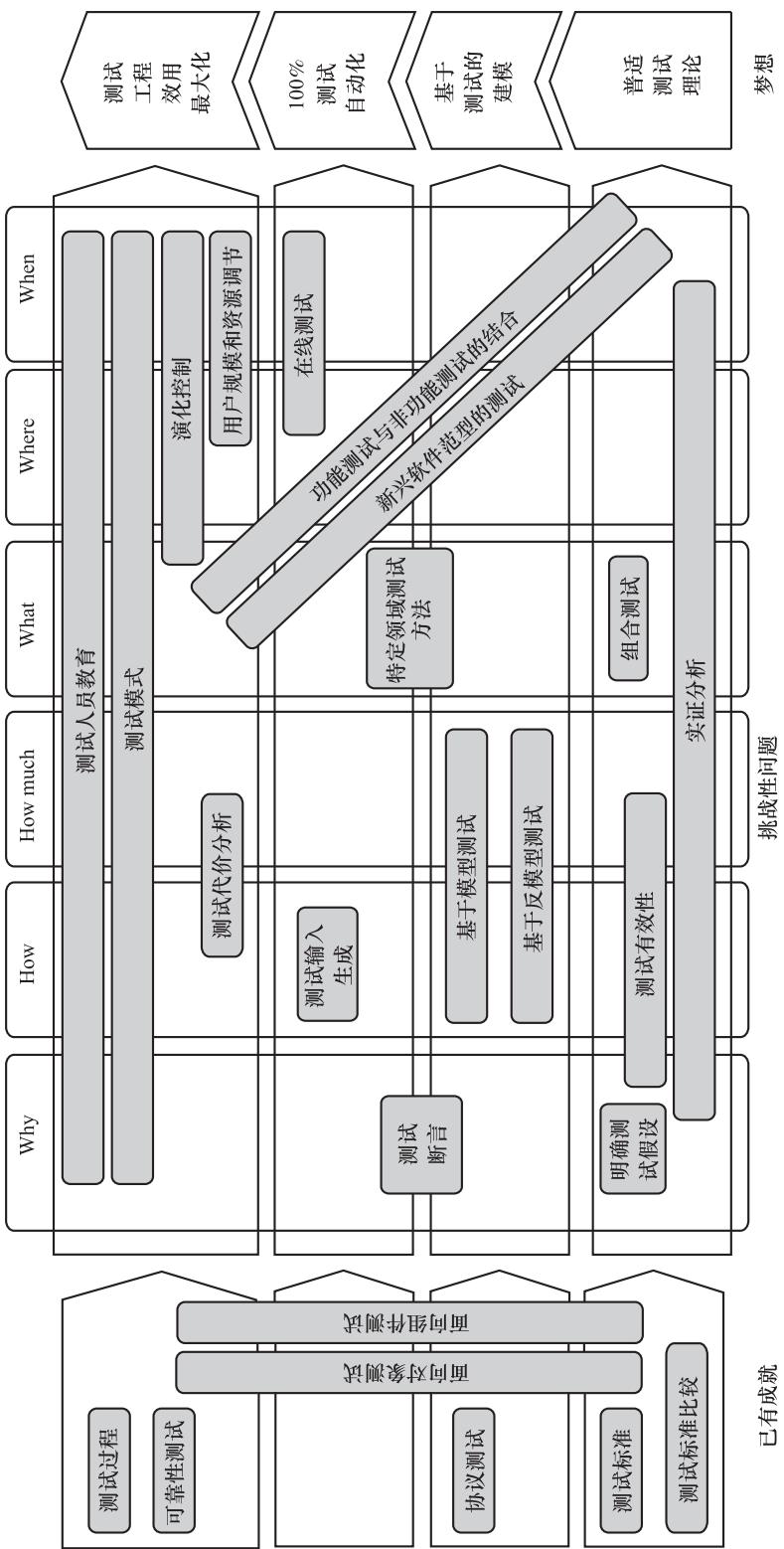


图4 软件测试已有成就、挑战性问题以及梦想 [4]

在基于测试的建模方面，现有的模型驱动测试方法多基于已有的开发模型，从测试驱动开发的角度，可以根据测试需求对开发模型提出更多的要求，促进提高系统的可测试性（design-for-testability），如输入可控性（controllability）和输出的可见性（observability）。断言（assertion）、契约测试等都是典型的测试建模方法。这方面的挑战性问题包括基于模型的测试、基于反模型的测试（anti-model based testing）以及测试预言（test oracle）。基于模型的测试为测试用例设计提供了相对完善的理论基础，但是如何构建模型，如何降低构建模型的代价以便于应用于实际工业领域，如何将不同风格的模型有效融合，如何将基于模型的自动化测试过程应用于特定的过程模型并与其他软件过程活动有效结合，是目前研究所面临的挑战性问题。测试断言是目前较少关注的一个领域，在实际测试中，多数情况下，测试结果依靠测试人员人工判定。在测试用例数量较多、测试工作量加大的情况下，自动化的产生对测试结果的判定将有助于降低测试代价。

在100%全自动测试方面，报告指出自动化测试主要面临着三个挑战性问题：测试输入自动生成、特定领域测试自动化和在线测试。输入是测试的基础，现有的基于模型的测试中多缺少对数据的建模和分析，一个可能的研究趋势就是将数据模型和数据生成算法与状态机等模型相结合，例如，Frantzen等人^[153]就提出了将符号转移系统（symbolic transition systems），在迁移系统中增强数据和数据依赖控制流的建模与分析。符号执行的研究始于70年代，曾由于一些难以克服的困难搁置了许多年，但近年来随着强类型程序设计语言、约束求解工具的发展，再次得到关注和应用^[153,154]。基于搜索的测试生成将测试数据和用例生成转化为搜索问题，采用智能搜索算法求解，是测试生成算法一个新兴的研究点^[20]。由于被测系统不同（如Web应用、嵌入式系统、数据库系统等），常常需要领域特定的测试技术（如测试驱动），因此面向特定领域的测试技术和工具研究是应用实践的基础。

在测试工程效用最大化方面，报告分析了将测试研究成果应用于实际软件工程，面临着一系列的挑战性问题。例如，回归测试如何能满足持续在线的软件演化的需求，在线测试中如何平衡用户规模和资源，如何根据需求选择适合的测试技术、测试费用和风险控制、测试专业教育与人员培训等。

除上述挑战问题外，软件测试还面临着另外两方面问题：1) 新兴软件范型的测试技术研究。随着软件开发技术的发展，不断涌现出新的开发方法和技术，测试需能适应新技术的发展。2) 功能测试与非功能测试的有效结合。目前对软件非功能特性的测试，如大规模负载性能、安全性等研究较少，也是需要关注的一个重点。

综上所示，软件测试的发展呈现出以下几方面的特点：

1) 测试自动化已成为必然趋势，自动化工具可以在测试设计、用例生成、运行部署、环境模拟、数据采集、结果分析、质量评估等各个环节提供支持，以提高测试的效率和质量，辅助测试知识和经验的积累与重用。目前，以第三方测试为代表的国内测试行业，多以人工为主，依赖于程序员的经验和能力，工作量大，且难以保证一致的和可信的测试结果。测试自动化已成为国际关注的一个重要课题，2007年美国海军曾委托OMG组织向全球发出自动化测试的方案征集书，探讨建立规范标准，推广最佳实践，以

降低测试费用和认证时间。

2) 大规模测试数据分析成为一个关键技术，数据分析的实时性、准确性、合理性直接影响测试质量。输入数据、执行序列、测试结果等，大量数据是驱动测试过程、诊断系统故障、评估测试结果的重要依据。例如，在淘宝、京东等电商系统中，要分析系统的性能瓶颈，需分析大量的用户行为和性能日志。近年来，越来越多的学者开始关注将数据挖掘、机器学习等技术应用到软件测试，把传统的准确建模和故障定位等问题转化为采用统计分析方法基于概率模型的估算问题，以解决越来越普遍的软件不确定性问题。

3) 系统测试问题突出，亟待新的方法和技术突破。早期软件测试技术主要针对模块级的数据流、控制流分析和覆盖率标准。随着开放式体系架构的深入发展，组件+组装的开发模式已经逐步成为主流，典型的如服务化的互联网软件和可组合嵌入式系统。受组件的异构性、环境的多样性、组装的动态可重构性等因素影响，集成测试和系统测试需面临新的故障模型和故障检测方法。

4) 测试系统的构建可借助于软件开发技术的发展，可有望突破传统的规模限制。例如，符号执行（symbolic execution）是 20 世纪 70 年代提出的一种经典的软件测试技术，但由于计算量较大，需耗费大量的内存和 CPU 等资源，难以应用到实际的工业实践中进行大规模程序分析。近年来，Ciortea 等人^[94]采用 Map Reduce 编程模型重构符号执行工具，将其任务分解和并行化，引入负载均衡等性能优化技术，并租用 Amazon EC2 云平台上虚拟机资源部署运行，初步实验结果显示性能平均提高 47 倍，最多可提高 250 倍。

6 结束语

测试是软件质量保证的重要手段，测试工作本身也需要具有较高的质量保证。国内已经形成了一定的业界标准和指南，指导系统规范地开展测试活动。但还需要更有效的测试技术和工具的支持，以提高测试效率，降低测试代价，保证测试结果的可信性。目前，国产测试工具的研发和推广还很有限，技术研究成果还需进一步结合业界的实际问题，以进一步推动业界测试工程化技术水平。

另一方面，开放平台、服务组合、众包等软件开发方式发生了较大的变化，测试的规模、层次、角度、质量关注点等也发生了相应的变化，相应地，也产生了新一类的测试技术。近年来，智能搜索、机器学习、并行编程、形式化建模等技术，越来越多地应用在测试研究中。软件开发技术的发展，势必为测试研究提出更多、更新的挑战性问题，提供更广阔的发展空间。

参考文献

[1] 朱少民. 软件测试方法和技术 [M]. 2 版. 北京: 清华大学出版社, 2012.

- [2] G Tassey. The Economic Impacts of Inadequate Infrastructure for Software Testing. National Institute of Standards and Technology , RTI Project, 2002 , 7007.
- [3] T Kurokawa, M Shinagawa. Technical Trends and Challenges of Software Testing. Science and Technology Trends, 2008.
- [4] A Bertolino. Software Testing Research: Achievements, Challenges, Dreams. Proc. Future of Software Engineering (FOSE) , 2007 , 85-103.
- [5] MJ Harrold. Testing: a Roadmap. Proc. Future of Software Engineering (FOSE). ACM, 2000 , 61-72.
- [6] L Baresi, M Pezze. An Introduction to Software Testing[J]. Electronic Notes in Theoretical Computer Science, 2006 , 148(1) : 89-111.
- [7] Ultra- Large- Scale Systems: The Software Challenges of the Future. Software Engineering Institute , Carnegie Mellon University. June 2006.
- [8] D Saff, M D Ernst. Reducing Wasted Development Time via Continuous Testing. Proc. IEEE International Symposium on Software Reliability Engineering (ISSRE) , 2003 , 2003 : 281-292.
- [9] J Penix. Panel Discussion. in 33rd International Conference on Software Engineering, 2011.
- [10] S Hosono. A DevOps Framework to Shorten Delivery Time for Cloud Applications[J]. International Journal of Computational Science and Engineering, 2012 , 7(4) : 329-344.
- [11] C Ramamoorthy, S Ho, W Chen. On the Automated Generation of Program Test Data[J]. IEEE Transaction on Software Engineering, 1976 , SE-2(4) : 293-300.
- [12] M Utting, A Pretschner, B Legeard. A Taxonomy of Model-Based Testing. 2005.
- [13] W Prenninger, A Pretschner. Abstractions for Model-Based Testing. Electronic Notes in Theoretical Computer Science, 2005 , 116 : 59-71.
- [14] G H Walton, J H Poore. Generating Transition Probabilities to Support Model-Based Software Testing [J]. Software: Practice and Experience, 2000 , 30(10) : 1095-1106.
- [15] K Berkenkotter, R Kirner. Real-Time and Hybrid Systems Testing. in Model-based Testing of Reactive Systems. M. Broy, et. al. (Eds.) , Springer-Verlag, 2005 , 355-387.
- [16] A Lamsweerde. Formal Specification: a Roadmap. Proc. Future of Software Engineering (FOSE) . ACM, 2000 : 147-159.
- [17] D Xu, O El-Ariss, W Xu, L Wang. Testing Aspect-Oriented Programs with Finite State Machines , Software Testing, Verification and Reliability, 2012 , 22(4) : 267-293.
- [18] K Claessen, J Hughes. QuickCheck: a Lightweight Tool for Random Testing of Haskell Programs. ACM SIGPLAN notices, 2011 , 46(4) : 53-64.
- [19] A Arcuri, L Briand L A Practical Guide for Using Statistical Tests to Assess Randomized Algorithms in Software Engineering. Proc. International Conference on Software Engineering (ICSE) , 2011 , 1-10.
- [20] P McMinn. Search - Based Software Test Data Generation: a Survey[J]. Software Testing, Verification and Reliability , 2004 , 14(2) : 105-156.
- [21] S Jehan, I Pill, F Wotawa. SOA Grey Box Testing--A Constraint-Based Approach. Proc. Software Testing, Verification and Validation Workshops (ICSTW) , 2013 , 232-237.
- [22] C Cadar, P Godefroid, S Khurshid, C S Păsăreanu, K Sen, N Tillmann, W Visser. Symbolic Execution for Software Testing in Practice: Preliminary Assessment. Proc. International Conference on Software Engineering. ACM, 2011 , 1066-1071.
- [23] P Amman, P E Black, W Ding. Model Checkers in Software Testing, NIST-IR 6777 , National Institute of

- Standards and Technology, 2002.
- [24] G Fraser, F Wotawa, P E Ammann. Testing with Model Checkers: A Survey[J]. *Software Testing, Verification and Reliability*, 2009, 19(3) : 215-261.
 - [25] X Li, X Qiu, L Wang, X Chen, Z Zhou, L Yu, J Zhao. UML Interaction Model-Driven Runtime Verification of Java Programs[J]. *IET Software*, 2011, 5(2) : 142-156.
 - [26] M Utting, B Legeard. Practical Model-Based Testing: a Tools Approach. Morgan Kaufmann, 2010.
 - [27] W Grieskamp. Multi-Paradigmatic Model-Based Testing. Formal Approaches to Software Testing and Runtime Verification. Springer Berlin Heidelberg, 2006, 1-19.
 - [28] H Mei, D Hao, L Zhang, L Zhang. A Static Approach to Prioritizing JUnit Test Cases[J]. *IEEE Transactions on Software Engineering*, 2012, 38(6) : 1258-1275.
 - [29] L Mei, WK Chan, TH Tse, R G Merkel. XML-Manipulating Test Case Prioritization for XML-Manipulating Services[J]. *The Journal of Systems and Software*, 2011, 84(4) : 603-619.
 - [30] Y-C Huang, K-L Peng, C-Y Huang. A History-Based Cost-Cognizant Test Case Prioritization Technique in Regression Testing[J]. *The Journal of Systems and Software*, 2012, 85(3) : 626-637.
 - [31] Z Wang, L Chen, B Xu, Y Huang. Cost-Cognizant Combinatorial Test Case Prioritization[J]. *International Journal of Software Engineering and Knowledge Engineering*, 2011, 21(06) : 829-854.
 - [32] K Zhai, W K Chan. Point-of-Interest Aware Test Case Prioritization: Methods and experiments, *IEEE 10th International Conference on Quality Software (QSIC)*, 2010, 449-456.
 - [33] C Zhang, Z Chen, Z Zhao, S Yan, J Zhang, B Xu. An Improved Regression Test Selection Technique by Clustering Execution Profiles. *IEEE 10th International Conference on Quality Software*, 2010, 171-179.
 - [34] L Zhang, S-S Hou, C Guo, T Xie, Hong Mei. Time-Aware Test-Case Prioritization using Integer Linear Programming. *ACM Proceedings of the 18th International Symposium on Software Testing and Analysis*, 2009, 213-224.
 - [35] S Huang, Z J Li, J Zhu, Y Xiao, WWang. A Novel Approach to Regression Test Selection for J2EE Application. *27th IEEE International Conference on Software Maintenance (ICSM)*, 2011, 13-22.
 - [36] Y Zou, C Fang, Z Chen, X Zhang, Z Zhao. A Hybrid Coverage Criterion for Dynamic Web Testing, *Proc. SEKE*, 2013, 201-213.
 - [37] S Sampath, R Bryce, A Memon. A Uniform Representation of Hybrid Criteria for Regression Testing. *IEEE Transactions on Software Engineering*, preprints, 2013.
 - [38] S Li, N Bian, Z Chen, D You, Y He. A Simulation Study on Some Search Algorithms for Regression Test Case Prioritization, *10th International Conference on Quality Software*, 2010, 72-81.
 - [39] D Hao, L Zhang, X Wu, H Mei, G Rothermel. On-Demand Test Suite Reduction. *Proceedings of the International Conference on Software Engineering*, 2012, 738-748.
 - [40] C Fang, Z Chen, B Xu. Comparing Logic Coverage Criteria on Test Case Prioritization[J]. *Science China Information Sciences*, 2012, 55(12) : 2826-2840.
 - [41] Z Chen, T Y Chen, B Xu. A Revisit of Fault Class Hierarchies in General Boolean Specifications[J]. *ACM Transaction on Software Engineering and Methodology*. 2011, 20(3) : 13.
 - [42] L Zhao, K-Y Cai. Event Handler-Based Coverage for GUI Testing. *IEEE 10th International Conference on Quality Software (QSIC)*, 2010, 326-331.
 - [43] L Mei, W K Chan, T H Tse, R G Merkel. Tag-Based Techniques for Black-Box Test Case Prioritization for Service Testing. *IEEE 9th International Conference on Quality Software (QSIC)*, 2009, 21-30.

- [44] C Ji, Z Chen, B Xu, Wang Z. A New Mutation Analysis Method for Testing Java Exception Handling. IEEE 33rd International Computer Software and Applications Conference (COMPSAC), 2009, 556-561.
- [45] W Xu, J Offutt, J Luo. Testing Web Services by XML Perturbation. 16th IEEE International Symposium on Software Reliability Engineering (ISSRE), 2005 , 266.
- [46] S Lee, X Bai, Y Chen. Automatic Mutation Testing and Simulation on OWL- S Specified Web Services. Proc. IEEE the 41th Annual Simulation Symposium (ANSS) , 2008. 149-156.
- [47] Y Jia, M Harman. An Analysis and Survey of The Development of Mutation Testing[J]. IEEE Transactions on Software Engineering, 2011 , 37(5) : 649-678.
- [48] R Baker, I Habli. An Empirical Evaluation of Mutation Testing for Improving the Test Quality of Safety-Critical Software. IEEE Transaction on Software Engineering, 2013.
- [49] P Ammann, J C Knight. Data Diversity: An Approach to Software Fault Tolerance[J]. IEEE Transactions on Computers, 1988 , 37(4) : 418-425
- [50] T Y Chen, F-C Kuo, R G Merkel, T Tse. Adaptive Random Testing: The Art of Test Case Diversity [J]. Journal of Systems and Software, 2010 , 83(1) : 60-66.
- [51] H Hemmati, A Arcuri, L Briand. Achieving Scalable Model-Based Testing Through Test Case Diversity [J]. ACM Transaction on Software Engineering and Methodology, 2013 , 22(1) : 6.
- [52] A Gonzalez-Sanchez, E Piel, H G Gross, A J Gemund. Prioritizing Tests for Software Fault Localization. IEEE 10th International Conference on Quality Software (QSIC) , 2010, 42-51.
- [53] A Gonz' alez-Sanchez, R Abreu, H-G Gross, A J C van Gemund. 26th IEEE/ACM International Conference on Automated Software Engineering (ASE) , 2011 , 83-92.
- [54] L Gong, D Lo, L Jiang, H Zhang. Diversity Maximization Speedup for Fault Localization. Proc. of the 27th IEEE/ACM International Conference on Automated Software Engineering (ASE) , 2012 , 30-39.
- [55] Z Zhang, W K Chan, T H Tse, Y T Yu, P Hu. Non-Parametric Statistical Fault Localization[J]. Journal of Systems and Software, 2011 , 84(5) : 885-905.
- [56] M J Harrold, G Rothermel, R Wu, L Yi. An Empirical Investigation of Program Spectra[J]. ACM SIGPLAN Notices, 1998 , 33(7) : 83-90.
- [57] R Abreu, P Zoeteweij, A J C van Gemund. On the Accuracy of Spectrum-based Fault Localization. IEEE Testing: Academic and Industrial Conference Practice and Research Techniques- MUTATION, 2007, 89-98.
- [58] L Naish, H Lee, K Ramamohanarao. A Model for Spectra-Based Software Diagnosis[J]. ACM Transactions on Software Engineering and Methodology, 2011 , 20(3) : 1-32.
- [59] K Yu, M Lin, Q Gao, H Zhang, X Zhang. Locating Faults Using Multiple Spectra- Specific Models. Proc. ACM Symposium on Applied Computing (SAC) , 2011 , 1404-1410.
- [60] X Xie, T Y Chen, F C Kuo, et al. A Theoretical Analysis of the Risk Evaluation Formulas for Spectrum- Based Fault Localization. ACM Transactions on Software Engineering and Methodology (TOSEM) , 2013.
- [61] Yi Miao, Zhenyu Chen, Sihan Li, Zhihong Zhao, Yuming Zhou. Identifying Coincidental Correctness for Fault Localization by Clustering Test Cases. Proc. SEKE, 2012 , 267-272.
- [62] C Catal. Software fault prediction: A Literature Review and Current Trends, Expert Systems with Applications, Elsevier, 38(2011) , 4626-4636.

- [63] M Li, H Zhang, R Wu, Z-H Zhou. Sample-Based Software Defect Prediction with Active and Semi-Supervised Learning. *Automatic Software Engineering*, 2012, 201-230.
- [64] T Hall, S Beecham, D Bowes, D Gray, S Counsell. A Systematic Literature Review on Fault Prediction Performance in Software Engineering [J]. *IEEE Transactions on Software Engineering*, 2012, 38(6) : 1276-1304.
- [65] C Ebert, C Jones. Embedded Software: Facts, Figures, and Future[J]. *IEEE Software*, 2009, 42(4) : 42-52.
- [66] A Lakehal, I Parassis. Structural Coverage Criteria for LUSTRE/SCADE Programs [J]. *Software Testing, Verification and Reliability*, 2009, 19(2) : 133-154.
- [67] J Bloomberg. Web Services Testing: Beyond SOAP. <http://searchsoa.techtarget.com/>.
- [68] G Canfora, M D Penta. SOA: Testing and Self-Checking. Proc. of International Workshop on Web Services Modeling and Testing (A. Bertolino and A. Polin, eds.) 2006, 3-12.
- [69] M Bozkurt, M Harman, Y Hassoun. Testing Web Services: A Survey [J]. Department of Computer Science, King's College London, Tech. Rep. TR-10-01, 2010.
- [70] S Nakajima. Model-Checking Verification for Reliable Web Service. Proc. OOPSLA Workshop on Object-Oriented Web Services, 2002.
- [71] H Foster, S Uchitel, J Magee, J Kramer. Model- Based Verification of Web Service Compositions. Proc. 18th IEEE International Conference on Automated Software Engineering, 2003, 152-161.
- [72] S Narayanan, S McIlraith. Simulation, Verification and Automated Composition of Web Services. Proc. of the 11th ACM International Conference on World Wide Web, 2002, 77-88.
- [73] M Koshkina, F van Breugel. Modeling and Verifying Web Service Orchestration by means of the Concurrency Workbench. *ACM SIGSOFT SEN*, 2004, 29(5) : 1-10.
- [74] H Huang, W T Tsai, R Paul, Y Chen. Automated Model Checking and Testing for Composite Web Services. 8th IEEE International Symposium on Object- oriented Real- time distributed Computing (ISORC), 2005, 300-307.
- [75] A Bertolino, G. De Angelis, L Frantzen, A Polini. Model- Based Generation of Testbeds for Web Services. *Testing of Software and Communicating Systems*. Springer Berlin Heidelberg, 2008, 266-282.
- [76] R Heckel, M Lohmann. Towards Contract- Based Testing of Web Services. *Electronic Notes in Theoretical Computer Science*, 2005, 116: 145-156.
- [77] M Bruno, G Canfora, M Di Penta, G Esposito, V Mazza. Using Test Cases as Contract to Ensure Service Compliance Across Releases. In *Service-Oriented Computing (ICSOC)*, Springer Berlin Heidelberg, 2005, 87-100.
- [78] X Bai, Y Wang, G Dai, W T Tsai, Y Chen. A Framework of Contract- Based Collaborative Verification and Validation of Web Services. The 10th International ACM SIGSOFT Symposium on Component- Based Software Engineering, 2007, 258-273.
- [79] W T Tsai, R Paul, Y Wang, C Fan, D Wang. Extending WSDL to Facilitate Web Services Testing. 7th IEEE International Symposium on High Assurance Systems Engineering, 2002, 171-172.
- [80] X Bai, W Dong, W T Tsai, Y Chen. WSDL-Based Automatic Test Case Generation for Web Services Testing. Proc. of IEEE Workshop on Service Oriented System Engineering (SOSE), 2005, 207-212.
- [81] B Li, D Qiu, H Leung, D Wang. Automatic Test Case Selection for Regression Testing of Composite Service Based on Extensible BPEL Flow Graph[J]. *The Journal of Systems and Software*, 2012, 85(6) : 1300-

- 1324.
- [82] C Sun, G Wang, B Mu, H Liu, Z Wang, T Y Chen. Metamorphic Testing for Web Services: Framework and a Case Study. IEEE International Conference on Web Services, 2011, 283-290.
- [83] S-S Hou, L Zhang, Q Lan, H Mei, J-S Sun. Generating Effective Test Sequences for BPEL Testing. International Conference on Quality Software, 2009, 331-340.
- [84] L Mei, W K Chan, T H Tse, F C Kuo. An Empirical Study of The Use of Frankl-Weyuker Data Flow Testing Criteria to Test BPEL Web Services. IEEE 33rd Annual IEEE International Computer Software and Applications Conference (COMPSAC), 2009, 1: 81-88.
- [85] X Bai, S Lee, W T Tsai, Y Chen. Ontology-Based Test Modeling and Partition Testing of Web Services. The 6th International Conference on Web Services (ICWS), 2008, 465-472.
- [86] C Ma, J Wu, T Zhang, F Hu. Web Services Testing Based on Stream X-machine. International Conference on Quality Software (QSIC), 2010, 232 - 239.
- [87] W T Tsai, Y Chen, R Paul, N Liao, H Huang. Cooperative and Group Testing in Verification of Dynamic Composite Web Services. Proc. of 28th Annual Int. Computer Software and Applications Conference, 2004, 170-173.
- [88] W T Tsai, R Paul, Z Cao, L Yu, A Saimi. Verification of Web Services Using an Enhanced UDDI Server. Proc. of 8th Int. Workshop Object-Oriented Real-Time Dependable Systems, 2003, 131-138.
- [89] X Bai, Z Cao, Y Chen. Design of a Trustworthy Service Broker and Dependence-based Progressive Group Testing, International Journal of Simulation and Process Modeling, 2007, 3(1): 66-79.
- [90] J Gao, X Bai, W T Tsai. Cloud Testing - Issues, Challenges, Needs and Practice. Software Engineering: An International Journal (SEIJ), IGI Global, 2011, 1(1): 9-23.
- [91] J Gao, K Manjula, P Roopa, E Sumalatha, Xiaoying Bai, Wei-Tek Tsai. Tadahiro Uehara: A cloud-based TaaS infrastructure with tools for SaaS validation, performance and scalability evaluation. CloudCom, 2012, 464-471.
- [92] L Riungu, O Taipale, K Smolander. Research Issues for Software Testing in the Cloud. IEEE Second International Conference on Cloud Computing Technology and Science (CloudCom), 2010, 557-564.
- [93] L Riungu, O Taipale, K Smolander. Testing in the Cloud: Exploring the Practice. IEEE Software, 2012, 29(2): 46-51.
- [94] L Ciortea, C Zamfir, S Bucur, V Chipounov, G Candea. Cloud 9: A Software Testing Service. ACM SIGOPS, 2010, 43(4): 5-10.
- [95] T Parveen, S Tilley, N Daley, P Morales. Towards a Distributed Execution Framework for JUnit Test Cases. IEEE International Conference on Software Maintenance, 2009, 425-428.
- [96] M Oriol, F Ullah. YETI on the Cloud. Third International Conference on Software Testing, Verification, and Validation Workshops, 2010, 434-437.
- [97] R Grossman, Y Gu, M Sabala, C Bennet, J Seidman, J Mambratti. The Open Cloud Testbed: A Wide Area Testbed for Cloud Computing Utilizing High Performance Network Services. arXiv preprint arXiv: 0907.4810, 2009.
- [98] L Yu, X Li, Z Li. Testing Tasks Management in Testing Cloud Environment. Proc. of 35th Annual IEEE International Computer Software and Applications Conference (COMPSAC), 2011, 76-85.
- [99] X Bai, M Li, B Chen, J Gao, W T Tsai. Cloud Testing Tools. Proc. IEEE International Symposium on Service Oriented System Engineering (SOSE), 2011, 1-12.

- [100] X Bai, M Li, X Huang, W T Tsai, J Gao. Vee@ Cloud: the Virtual Test Lab on the Cloud. Proc. of the 6th International workshop on Automation of Software Test (AST), 2013.
- [101] J Wu, C Wang, Y Liu, L Zhang. Agaric — A Hybrid Cloud Based Testing Platform. International Conference on Cloud and Service Computing (CSC), 2011, 87-94.
- [102] G Candea, S Bucur, C Zamfir. Automated Software Testing as a Service, Proc. 1st ACM symposium on Cloud computing, 2010, 155-160.
- [103] L Yu, W Tsai, X Chen, L Liu, Y Zhao, L Tang, W Zhao. Testing as a Service over Cloud. Proc. Fifth IEEE International Symposium on Service Oriented System Engineering, 2010, 181-188.
- [104] E Dolstra, R Vliegendaal, J Pouwelse. Crowdsourcing GUI Tests. Proc. ICST, 2013.
- [105] F Pastore, L Mariani, G Fraser. Crowdoracles: Can the Crowd Solve the Oracle Problem? Proc. ICST, 2013.
- [106] 刘攀, 缪淮扣, 曾红卫, 刘阳. 基于 FSM 的测试理论、方法及评估[J]. 计算机学报, 2011, 34(6): 965-985.
- [107] B Song, H Miao. Modeling Web Applications and Generating Tests: A Combination and Interactions-guided Approach. Third IEEE International Symposium on Theoretical Aspects of Software Engineering, 2009, 174-181.
- [108] 雷斌, 王林章, 卜磊, 李宣东. 基于状态机模型的构件健壮性测试[J]. 软件学报, 2010, 21(5): 930-941.
- [109] 舒挺, 刘良桂, 徐伟强, 李文书. 自适应 EFSM 可执行测试序列生成[J]. 计算机研究与发展, 2012, 49(6): 1211-1219.
- [110] 侯克佳, 白晓颖, 陆皓, 李树芳, 周立柱. 基于接口语义契约的 Web 服务测试数据生成[J].《软件学报》.
- [111] L Zhang, D. Hao, L. Zhang, G Rothermel, H. Mei. Bridging the gap between the total and additional test-case prioritization strategies. Proc. ICSE, 2013, 192-201.
- [112] D Hao, T Lan, H Zhang, C Guo, L Zhang. Is This a Bug or an Obsolete Test? Proc. ECOOP, 2013, 602-628.
- [113] S Yan, Z Chen, Z Zhao, C Zhang, Y Zhou. A Dynamic Test Cluster Sampling Strategy by Leveraging Execution Spectra Information. Proc. ICST, 2010, 147-154.
- [114] C Zhang, Z Chen, Z Zhao, S Yan, J Zhang, B Xu. An Improved Regression Test Selection Technique by Clustering Execution Profiles. Proc. QSIC, 2010, 171-179.
- [115] Z Chen, Y Duan, Z Zhao, B Xu, J Qian. Using Program Slicing to Improve the Efficiency and Effectiveness of Cluster Test Selection. IJSEKE, 2011, 21(6): 759-777.
- [116] S Chen, Z Chen, Z Zhao, B Xu, Y Feng. Using semi-supervised clustering to improve regression test selection techniques. Proc. ICST, 2011, 1-10.
- [117] 严俊, 张健. 组合测试: 原理与方法[J]. 软件学报, 2009, 20(6): 1393-1405.
- [118] C Nie, H Leung. A survey of combinatorial testing. ACM Computing Surveys (CSUR), 2011, 43(2): 11.
- [119] C Nie, H Leung. The minimal failure-causing schema of combinatorial testing, ACM Transaction on Software Engineering and Methodology (TOSEM), 2011, 20(4).
- [120] 潘伟丰, 李兵, 马于涛, 刘婧. 基于复杂软件网络的回归测试用例优先级排序[J]. 电子学报, 2012, 40(12): 2456-2465.

- [121] B Chen, J Zhang. Tuple density: a new metric for combinatorial test suites. Proc. ICSE, 2011, 876-879.
- [122] 安金霞, 王国庆, 李树芳, 朱纪洪. 基于多维度覆盖率的软件测试动态评价方法[J]. 软件学报, 2010, 21(9): 2135-2147.
- [123] 刘攀, 缪淮扣, 曾红卫, 梅佳. 确定性有限状态机的最小测试成本迁移覆盖准则[J]. 软件学报, 2011, 22(7): 1457-1474.
- [124] Y Lei, X Mao, Z Dai, C Wang. Effective Statistical Fault Localization Using Program Slices. Proc. COMPSAC 2012, 1-10.
- [125] Y Qi, X Mao, Y Lei, C Wang. Using automated program repair for evaluating the effectiveness of fault localization techniques. Proc. ISSTA, 2013, 191-201.
- [126] Y Zhang, Z Chen, J Wang. Speculative Symbolic Execution. Proc. ISSRE, 2012, 101-110.
- [127] Y Wang, Z Chen, Y Feng, B Luo, Y Yang. Using Weighted Attributes to Improve Cluster Test Selection. Proc. SERE, 2012, 138-146.
- [128] Y Feng, Z Chen. Multi-label software behavior learning. Proc. ICSE, 2012, 1305-1308.
- [129] 谭德贵, 陈林, 王子元, 丁晖, 周毓明, 徐宝文. 通过增大边际权重提高基于频谱的错误定位效率[J]. 计算机学报, 2010, 33(12): 2335-2342.
- [130] Z Zhang, W K Chan, T. H. Tse. Fault Localization Based Only on Failed Runs. IEEE Computer, 2012, 45(6): 64-71.
- [131] 叶钢, 李重文, 余丹, 马世龙. 基于非参数假设检验的程序缺陷定位方法[J]. 北京航空航天大学学报, 2012, 38(8).
- [132] H Shen, J Fang, J Zhao: EFindBugs: Effective Error Ranking for FindBugs. Proc. ICST, 2011, 299-308.
- [133] 战伟, 黄松, 嵇孟雨. 基于程序特征谱整数溢出错误定位技术研究[J]. 计算机学报, 2012, 35(10): 2204-2214.
- [134] 褚灵伟, 邹仕洪, 程时瑞, 田春岐, 王文东. 一种动态环境下的互联网服务故障诊断算法[J]. 软件学报, 2009, 20(9): 2520-2530.
- [135] 王青, 伍书剑, 李明树. 软件缺陷预测技术[J]. 软件学报, 2008, 19(7): 1565-1580.
- [136] 何智涛, 晏海华, 刘超. 软件缺陷发现时序过程的叠加双阻尼振荡模型[J]. 软件学报, 2010, 21(12): 2999-3010.
- [137] 何智涛, 晏海华, 刘超. ABDOM 的参数规范化和离散化改进[J]. 北京航空航天大学学报, 2012, 38(10): 1415-1419.
- [138] 张建伟, 叶东升. 嵌入式系统测试的发展[J]. 单片机与嵌入式系统应用, 2011, 02: 5-7.
- [139] H Liu, M Jin, C Liu. Construction of the Simulating Environment for testing distributed embedded software. 5th International Conference on Computer Science and Education (ICCSE), 2010, 97-101.
- [140] H Liu, M Jin, C Liu. A Simulator Based Testing Framework for Embedded Assembly Language. Proc. International Conference on Power Electronics Engineering and Computer Technology, 2012, 344-347.
- [141] 许福, 金茂忠, 刘辉, 陈琦. 基于仿真器的嵌入式软件单元测试方法[J]. 北京航空航天大学学报, 2012, 38(12).
- [142] 殷永峰, 刘斌, 王晨. 实时嵌入式软件测试执行引擎的设计与实现[J]. 北京航空航天大学学报, 2010, 36(6).
- [143] 周庆, 刘斌, 余正伟, 冯时雨. 综合模块化航电软件仿真测试环境研究[J]. 航空学报, 2012, 33(4): 722-733.

- [144] 梅宏, 刘譞哲. 互联网时代的软件技术: 现状与趋势[J]. 中国科学, 2010, 55(13): 1214-1220.
- [145] X Bai, B Chen, B Ma, Y Gong. Design of Intelligent Agent for Collaborative Testing of Service-Based Systems, Proc. International workshop on Automation of Software Test (AST), 2011, 22-28.
- [146] 邓小鹏, 邢春晓, 张勇, 蔡莲红. 一种面向 QoS 的 Web 服务组测试方法[J]. 计算机研究与发展, 2009, 46(8): 1285-1293.
- [147] 廖军, 谭浩, 刘锦德. 基于 Pi - 演算的 Web 服务组合的描述和验证[J]. 计算机学报, 2005, 28(4): 635-643.
- [148] B Li, S Ji, D Qiu, J Cai, J Generating Test Cases of Composite Services Based on OWL-S and EH-CPN. International Journal of Software Engineering and Knowledge Engineering (IJSEKE), 2010. 20(07): 921-941.
- [149] C Blundell, D Giannakopoulou, C S Păsăreanu. Assume-Guarantee Testing. ACM SIGSOFT Software Engineering Notes. ACM, 2005, 31(2): 1.
- [150] Van Der Bijl M, A Rensink, J Tretmans. Compositional Testing with Ioco. Formal Approaches to Software Testing, Springer Berlin Heidelberg, 2004, 86-100.
- [151] J Tretmans. Model Based Testing with labelled Transition Systems. Formal Methods and Testing, Springer Berlin Heidelberg, 2008, 1-38.
- [152] N Juristo, A M Moreno, S Vegas. Reviewing 25 Years of testing Technique Experiments[J]. Empirical Software Engineering, 2004, 9(1-2): 7-44.
- [153] L Frantzen, J Tretmans, T A C Willemse. A symbolic Framework for Model-Based Testing. Formal Approaches to Software Testing and Runtime Verification. Springer Berlin Heidelberg, 2006, 40-54.
- [154] G Lee, J Morris, K Parker K, G A Bundell, P Lam. Using Symbolic Execution to Guide Test Generation [J]. Software Testing, Verification and Reliability, 2005, 15(1): 41-61.

作者简介

白晓颖 博士, 副教授, 清华大学计算机科学与技术系, 主要研究方向为模型驱动测试自动化以及面向服务软件工程。



刘超 博士, 教授, 北京航空航天大学计算机学院, 软件工程研究所所长。计算机学会高级会员和软件工程专业委员会副主任, 中国软件行业协会软件工程分会主任和系统与软件过程分会副会长。主要研究领域: 软件质量、软件测试、软件过程改进。



徐宝文 博士，南京大学计算机科学与技术系，教授、博士生导师，国家杰出青年科学基金获得者。分别获武汉大学、华中科技大学、北京航空航天大学学士、硕士、博士学位，国防科大博士后。1990年与1994年分别破格晋升副教授与教授，1995年被评为博士生导师。主要从事程序设计语言、软件工程（软件方法论、软件分析、度量与测试）、Web技术等方面的研究。主编系列丛书（教材）三套，出版著译作10余部，发表论文300余篇，被SCI、EI、ISTP收录300余篇次。作为第一获奖者获得省部级科技进步（自然科学）一等奖3次、省部级科技进步（自然科学）二等奖6次、三四等奖若干次。已指导博士后10人、博士生40余人（其中已毕业30余人）、各类硕士生100余人。所指导的学生10多万人次获得全国百篇优秀博士论文、微软学者奖、江苏省优秀博士论文、江苏省优秀硕士论文、江苏省优秀本科毕业论文一等奖。



计算机存储技术的研究进展与趋势

CCF 体系结构专业委员会

摘要

存储的本质是信息在时间维度上传输。信息存储涉及多个科学技术领域，从底层物理存储机理到大规模并行存取过程，从物理的原始编码到亿万用户对于各种数据理解，从瞬时信息保持过程到上万年的长期保存，从嵌入式设备几个字节到超大规模系统 PB 级的容量。考虑到大数据的基础就是数据存储，随着大数据时代的到来，最近几年存储技术在各个领域获得巨大的发展。从存储介质、存储结构、存储系统和存储服务应用四个层次分别讨论国内外对存储技术的研究现状，揭示一些具体的技术思想；之后，对于国内外研究进展和差异进行比较和分析；最后给出发展趋势与展望的说明。

关键词：存储系统，系统结构，数据组织

Abstract

Data storage, in nature, is information transmission along the time dimension, which involves multiple areas of science and technology, from the underlying physical storage mechanisms to massive parallel access processes, from physical data coding to a variety of data expression forms to be understood by millions of users, from instantaneous memory to long-term preservation for thousands of years, from a few bytes used in embedded devices to petabyte scale in large storage clouds. Considering data storage is the basis of Big Data, storage systems and technologies for Big Data in recent years have obtained tremendous development. This article attempts to conclude the recent storage research status and progress Both in China and abroad, from four aspects such as storage medium, storage architecture, storage system, and storage service and application. First, some methodologies of these technologies are presented and analyzed. Then, a comparison of research status between overseas and domestic is made to indicate their difference. Finally, the future research progress, development trend and prospect are discussed and suggested.

Keywords: Storage system, Computer architecture, Data organization

1 引言

人类社会进步依赖于前人的经验和知识的积累。人类很早就开始记录各种信息，但只有当纸张这种存储介质出现后，更多的信息才被记录下来，并且长期保存，实现知识的积累和传播，使人类社会得到极大的发展。以计算机和互联网为代表的信息技术，日益深入到人类社会活动的方方面面，不断加深的数据化过程越来越深刻地影响人类社会，使现代社会高度依赖于对于数据的收集、存储和处理。在对世界基本数据化之后，人类

进入大数据时代。大数据开启了一次重大的时代转型，正不断变革我们的生活以及理解世界的方式，成为新发明、新服务和新经济的源泉。

大数据的基础就是数据存储，存储的本质是信息在时间维度上传输，也是人类知识和信息得以广泛传播、共享、继承和发展的技术基础。大数据要求在更大范围、更长时间内保存各种数据，从而为当代或者未来的研究者、经济活动提供重要信息和新知识。与传统科学采取的基于局部数据和样本分析的方法不同，大数据科学更依赖于全局数据模式，越大、越长、越繁杂的数据全集是各种相关性分析和知识挖掘的基石。这迫切要求存储系统能够更加高效地存储更多的数据。各种研究报告表明：2010 年人类社会具有的数据总量为 1.2 ZBytes(1 ZBytes = 1000 EBytes = 100 万 PBytes = 10 亿 TBytes = 1 万亿 GBytes)，而 2020 年人类数据总量将超过 35 ZBytes。这些数据都需要存放于计算机存储系统中，因此数据存储需求也将呈指数级增长。可见，研究具有大容量、高性能、安全、绿色和持久的数据存储原理和技术是大数据时代迫在眉睫的关键科学问题之一。

现有信息存储涉及多个科学技术领域，从底层物理存储机理到大规模并行存取过程，从物理的原始编码到各种数据语义表述方式，从瞬时信息保持过程到上万年的长期保存，从嵌入式设备几个字节的容量到超大规模系统 PB 级的容量。而最近几年存储技术在各个领域均获得巨大的发展。本文将从存储介质、存储结构、存储系统和存储服务应用四个方面分别讨论国内外存储研究的现状，并对国内外研究进展进行比较，给出发展趋势与展望的说明。

2 国际研究现状

2.1 存储介质

2.1.1 磁存储

到目前为止，世界上超过 95% 的数据都存储在磁硬盘中。在基于磁存储机制的硬盘技术下，数据存储在磁盘表面的二维平面内，受现有工艺技术的限制，其物理存取性能在过去十年内仅仅提高两倍。硬盘容量的提高主要依赖于存储密度的提高，密度的提高必然会带来信噪比的下降。为保证在增大存储密度的前提下使读出信号的信噪比提高，必须保证记录位拥有足够数量的磁畴。因此增大存储密度只能减少磁畴的体积，但这会导致热稳定性的减小。为保持热稳定性不变就必须增加各向异性常数，从而导致写磁场大于写头可提供的磁场；如果要增加写头提供的磁场，就需要增加写头的体积，导致记录密度和速度的下降。因此，在目前的硬盘技术领域，提高密度的几个因素相互制约，导致密度的提高受到限制。

为摆脱上述困境，磁记录界先后提出了热辅助磁记录技术（HAMR）、图案化介质（Pattern Media）等技术，2009 年还提出了过渡技术——瓦记录技术。这些都是在磁记录的技术框架内的改进技术。

热辅助磁记录技术（HAMR）着眼于解决写头磁场能量不足的问题。其主要原理是利用激光将热量注入记录位，将记录介质温度快速提高到居里点附近，以降低记录介质的矫顽力，从而实现使用较小的写磁场完成写入，写完后介质冷却又恢复高的矫顽力，保证稳定性。热辅助磁记录技术以增加磁头复杂度和牺牲访问性能为代价来获得记录密度的提高，但实现该方案面临诸多技术问题，如散热问题目前还没有很好解决，故这一技术还停留在实验室阶段，没有进入商用阶段。

图案化介质技术则主要通过改变介质来解决三难问题。该技术方案较 HAMR 具有更大的技术潜力，但是实现难度比 HAMR 更大，因而被认为是 HAMR 的下一代技术。Pattern Media 的主要问题在于盘片的制备。目前，西部数据公司所收购的日立数据已经在 Pattern Media 方面取得了重要进展，可望将存储密度在 2016 年前提高一倍。

瓦记录（Shingle Recording）技术通过如瓦片叠盖的方式获得了更大的容量，但牺牲了硬盘随机写的性能，并导致性能显著下降。因此，该技术只适合某些特定场合，如类似磁带备份的应用场景。

除存储密度之外，磁硬盘系统把存储介质和驱动器始终封装在一起，驱动器只能读本地磁介质上的数据，即使仅读写一个字节，磁盘驱动器也必须完全工作。磁盘失效率和寿命不仅受到磁介质（温度、湿度和读写次数等）的影响，还受到驱动器机械部件（存取强读、开关次数等）和负载强度的影响。在实际运行中，“每隔两年”硬盘上的数据就要被重写一次以避免数据的丢失。可以通过系统的方法把大量存储设备高效组织起来构成大规模存储系统，保证快速提升系统整体的容量、性能和可靠性。磁盘阵列系统就是这种思路的具体体现。

2.1.2 固态存储

近几年固态存储技术的研究和应用发展迅猛。固态存储介质能耗较低且性能较高，已经成为高性能存储的主要介质，但缺点是其单位存储的成本是硬盘的 10 倍以上，而且其存储寿命比磁盘还低。固态存储介质主要包括基于闪存的大容量固态盘、相变存储器（Phase Change Memory，PCM）、磁存储（Magnetic Memory，MRAM）、忆阻器（Memristor）等。表 1 给出了几种典型固态存储器的性能对比。

表 1 几种典型固态存储器的技术比较

指标	当前存储技术				新型存储技术			
	SRAM	DRAM	DISK	NAND Flash	PCRAM	RRAM Memristor	MRAM STT-RAM	FRAM
成熟度	产品	产品	产品	产品	样品	早期开发	样品	样品
单元尺寸（特征尺寸 F^2 ）	$>100F^2$	$6 \sim 8F^2$	$(2/3)F^2$	$4 \sim 5F^2$	$8 \sim 16F^2$	$>5F^2$	$25F^2$	$10 \sim 15F^2$
读延迟	$<10\text{ns}$	$10 \sim 60\text{ns}$	8.5ms	$25\mu\text{s}$	48ns	$<10\text{ns}$	$<10\text{ns}$	$<30\text{ns}$
写延迟	$<10\text{ns}$	$10 \sim 60\text{ns}$	9.5ms	$200\mu\text{s}$	$40 \sim 150\text{ns}$	$\sim 10\text{ns}$	12.5ns	50ns
每位存取能耗	$>1\text{pJ}$	2pJ	$100 \sim 1000\text{mJ}$	10nJ	100pJ	2pJ	0.02pJ	0.01pJ
静态功率	是	是	是	否	否	否	否	否
读写寿命	$>10^{15}$	$>10^{15}$	$>10^{15}$	10^4	10^8	10^5	$>10^{15}$	$>10^{14}$
寻址单位	字节	字节	扇区	页	字节	字节	字节	字节
非易失性	否	否	是	是	是	是	是	是

目前基于 NAND 的闪存技术最为成熟，已经广泛应用到移动存储到高性能加速器等各个方面。现阶段其主要考虑是每个存储单元从单层、双层到更高层发展，到 2013 年，每个单元可以容纳 4 层（16 个信息位）；而到 2015 年，预计采用 450 毫米晶圆生产，使得闪存存储密度进一步提高。但是增加单元的信息位将使读取过程更加复杂，读取器件分辨率更高，层级和单元之间串扰更大。这都是需要技术解决的问题。

Wesel^[1]给出了 4 层噪声模型，并给出一种动态电压分配方法；Jeon^[2]分析了单元间干扰问题，提出两种基于信号处理的补偿方法，并设计最小能量和调制迁移编码；Hambrey^[3]使用原始误码率变化来优化 ECC（错误检测和纠错编码）性能，增加 P/E 次数；Qin^[4]为了减少单元间串扰，禁止部分单元向量编码，并且针对读使用平衡编码，针对写使用一次写编码；Jiang^[5]提出了 bit-fixing 编码，顺序纠正二进制表示的错误；Motawani^[6]使用数字信号处理的方法提高固态盘的耐久性。Sharon^[8]使用数据 I/O 编码（RIO），把 M 层数据分割为 $M - 1$ 个数据组，若编码第 j 个数据组，当传感器测试到电量在 $j - 1$ 和 j 时，对应所有单元为开状态。Sala^[10]给出 MLC（多级单元）中的动态阈值电压方案，使得选择阈值在同一级内保证单元数量相同。传统的纠错 BCH 码对于 SSD 越来越无效，因为其容量不断增长，可靠性不断下降。低密度校验检测码（LDPC）可明显提高 SSD 的可靠性，虽然 LDPC 已经在硬盘中成功使用，但在 SSD 中使用时需要非传统的细粒度 Flash 传感器，导致内存读的延迟，Zhao^[17]提供三种技术能降低 LDPC 导致的延迟。

相变存储（PCRAM）的原理是根据材料处于晶态和非晶态时电阻的差异进行信息存储，一般晶态低电阻表示 0，非晶态高电阻态为 1。由于晶态转换到非晶态需要大电流冲击，虽然状态转换速度快，但是产生热量大，对于材料磨损也大；而从非晶态转换到晶态延迟较长。由于非晶态到晶态变化过程中存在多种稳定电阻态，因此可以像多层闪存那样针对每个存储单元设计多个层级（MLC）。在相变存储器中，写磨损导致硬件错误，而电阻漂移（Resistance Drift）导致软错误。针对不同错误模式，一般通过减少写数据量（写前读，仅仅写不同的信息位）和在所有存储单元中均匀分布写来减少单个单元的写次数。在编码方面，一般使用 6 错纠正 BCH 编码，但是这种编码方式在校验位上有较大的固定开销，并且解码复杂。在分配方面，先使用记录硬错误的位置（Error-Correcting Pointers, ECPS），在此基础之上设计按需分配存储（Pay-As-You-Go, PAYG）。Zhang^[7]提出了一种 2 错纠正的 BCH 码。但是由于材料技术不成熟，目前镁光生产的相变存储器寿命只有 10^6 次。另外，由于非晶化过程中的热量问题，使得相变存储的存储密度较难提升。

磁内存（MRAM）是另一种固态高性能非易失存储方案，它利用磁致电阻效应和隧道效应，使用自旋扭矩转换（Spin-Torque-Transfer, STT）技术改变磁致电阻。一般而言，磁通道关节（Magnetic Tunnel Junction, MTJ）中自由层磁性同向为低电阻状态，相反为高电阻状态。相对于相变存储，磁内存方案具有更高的密度，更低的读写延迟，对于读写次数几乎没有限制。因此，MRAM 能够适应 cache 的要求。目前使用氧化镁作为通道栅栏（Tunneling Barrier），硼铁化钴作为自由层材料。Sugii^[11]提出了一种 STT-RAM 的制备方案。Everspin 技术公司开发出基于 DDR3 接口的 64MB 样品。

铁电存储（Ferro-Electric RAM, FRAM）也是一种有前景的固态非易失存储技术，具有较好的写性能，但目前存储密度较低，样品为 64KB。FRAM 存储单元中包含锆钛酸铅的铁电晶体，通常被称为 PZT。当一个电场加在存储单元中的铁电晶体上时，中间的原子在电场方向上移动，穿过能量壁，导致电荷击穿。内部放大器感受到击穿，设置单元状态为 0/1。状态能够随电场的方向翻转。如果电场移开，原子保存其位置，也就保证其状态。FRAM 的写次数超过 10^{14} ，8 字节写延迟为 50ns，写电流大于为 300 微安，而 Flash 写电流为 35 毫安^[12]。

忆阻器（Memristor）是一种新型存储机制。虽然蔡绍棠在 1971 年就提出忆阻器的概念，并于 2011 年拓宽了这一概念，将几乎所有基于可变电阻切换都视为是忆阻器，包括 MRAM、PCRAM 等，但其理论严谨性一直受到其他学者的质疑。目前其研究大体集中在三个领域，一个是材料领域，主要考虑忆阻材料的制备，包括二氧化钛、聚合材料、铁电材料、磁阻材料等；另一方面是结构设计，研究基于忆阻器的内存阵列，在阵列中高阻值定义为 0，低阻值定义为 0。Zidan^[14] 正在研究忆阻交叉阵列中 Sneak-Path 限制问题；最后一个方面是应用领域的研究，根据忆阻器的功能特性，研究人体的神经网络^[13]。

除了上述新型存储介质外，三维堆叠技术（3D Stacks Of Memory）也能够进一步提高固态存储的密度。IBM 已经开展对于三维混合离电导体（3-D Crosspoint Access Devices Mixed Lonic Electronic Conduction, MIEC）的研究^[15]，利用三维多层交叉点内存技术（3D Multi-Layer Cross-Point (XPT) Memory）实现，三维交叉点内存具有高性能、高密度和非易失的特点，可以采用 STT-MRAM、RRAM 和 PCM 介质。除材料问题之外，3D XPT 会产生新的问题，在开状态下，它要求非常高的电流密度，对于 PCM 的 RESET 而言，需要大于 10MA 每平方厘米，而 RRAM 更低一点。另外，它还要求低关状态漏电流，要求开关电路比大于 10^7 ，而宽低漏电力小于 100PA，并且要求增加电压范围去容纳大阵列中的半选单元。而 X-Caliber 架构预计在 2016 年以后实现，其中每个堆使用 32GB DRAM 和 256GB PCM，逻辑芯片在堆底，有 8 个双工连接。

2.1.3 光存储

光存储技术在过去十年发展较慢，但较传统的磁存储及电存储技术在特定领域也具有技术优势，主要表现在低碳节能、长存储寿命和高存储可靠性等方面。特别是人们发现光存储在超长期信息保存方面具有独特的优势。在经过近十年的研究沉寂之后，随着纳米技术和材料的发展，以多维光存储为代表的新型高密度存储原理也取得重大进展。

体全息存储是一种超越现有串行面式存储的并行体式存储技术，利用体光栅的布喇格选择性，能够将二维数据页在体全息存储材料内通过复用技术实现三维存储。如果引入偏振、相位等信息，可以实现多维存储，光盘的存储容量可达 TB 量级^[19]。

双光子激发具有更高的空间局域性，开辟了利用介质三维物理空间的新存储模式，将每张光存储光盘的理论密度提升到 TB 数量级。由于光的衍射极限限制，因此记录光斑尺度已经接近理论极限，因此限制了三维光存储的继续发展。为进一步提高光存储的密度，科学家们先后提出了利用记录光的物理参量空间（例如偏振和波长）进行编码存

储。利用纳米光子的波长和偏振态选择激发性，顾敏率先开展了基于纳米颗粒材料的超高容量多维光存储的研究，提出了基于金纳米棒的五维光存储概念，同时利用记录光的3个波长及2个偏振态，突破了现有高密度存储技术的瓶颈，实现了单盘1.6TB的超高密度以及能耗为同容量硬盘技术的15%的低能耗光信息存储技术^[18]。

2.1.4 新型存储机理

1. 玻璃存储

2011年，英国南安普敦大学的科学家使用激光实现玻璃存储。该技术的原理是首先让一束激光聚焦，随后将名为三维像素（Voxels）的小点铭刻进纯净的石英玻璃内，使玻璃变得有点模糊，光通过玻璃时会发生极化。极化过程改变了光通过玻璃的方式，制造出了极光漩涡，从而将信息记录于玻璃内。玻璃存储器内的信息阅读方式与光纤内数据的阅读方式一样，而且，对于存储的数据也可以利用激光进行清除、重写等操作。

日立公司于2012年9月宣布已开发出一种全新的石英玻璃数据存储技术。使得存储的数据在1000℃的环境下能够存放两个小时，而在常温情况下能够保存数亿年。存储单元由边长2厘米、厚度2毫米的正方形石英玻璃组成，每平方英寸可存储40MB数据，数据是通过石英玻璃上的激光作用点按照四层结构以二进制的格式进行刻录的，可使用普通的光学显微镜读取。同时，可以通过在存储单元上添加更多的层来增加存储容量。

2. DNA 存储

作为地球所有生物信息遗传的载体，DNA双螺旋结构在1953年由沃森和克里克发现。DNA由腺嘌呤（A）、鸟嘌呤（G）、胞嘧啶（C）和胸腺嘧啶（T）四种基本碱基组成。由于碱基互补配对原则要求，也就是A和T配对、G和C配对，因此四种基本碱基对仅能表示两位信息，可以编码成二进制0和1。而且由于配对关系和双螺旋链式结构，DNA链具有信息的冗余和纠错能力。这使得用DNA存储数据在原理上是可行的。

2012年初，斯坦福大学研究人员将数据编入噬菌体的DNA片段并植入大肠杆菌的DNA，从而实现在活体细胞内DNA中重复擦写数字信息。2012年，哈佛大学乔治·丘奇（George Church）发明了一种DNA信息归档系统^[20]，完全不需要利用活体细胞进行信息存储。其原理是写数据时使用UTF-8进行二进制编码，并转化为碱基序列（A, C表示0；G, T表示1），每12个字节加上19位地址，然后通过仪器进行DNA合成；读取信息时，通过DNA测序仪提取链上的碱基序列，将其转化为0/1位流，并用计算机将所有片段按序重新组合起来，进而转换为UTF-8格式。由于数据块能够被复制成很多份，因此通过设计高效容错和纠错机制能够保证数据的可靠性。根据测算，每克DNA能够存储 4.5×10^{20} 个字节。

2013年1月，尼克·戈德曼（Nick Goldman）将739KB的数据，包括154首莎士比亚十四行诗、一张欧洲生物信息研究所的JPG照片、一篇沃森和克里克论文的PDF版和马丁·路德·金著名演讲的26秒片段，通过3位ASCII进行DNA存储。通过合成、测序和重构之后获得100%存储正确率^[21]。

目前，DNA存储的成本极高，远没有达到商业化地步，而且读写延迟非常长，不适

合在线存储，但是可以用在大量数据的长期归档存储的场合下。

2.2 存储结构

传统计算机存储结构主要涉及两个方向，第一个方向是计算机系统的层次性存储结构，也就是由 CPU、Cache、内存和外存组成的多级存储结构；第二个方向主要讨论数据持久存储，涉及外存储系统结构，包括磁盘阵列和各种 I/O 接口等。前者主要关注优化数据存取过程，进而提高计算机系统整体处理性能；而后者在关注 I/O 性能的同时，还重点考虑数据存储的可靠性、扩展性、一致性、复杂性和管理性等问题。近几年，新型固态存储器的引入和发展，使新型非易失存储器性能逐渐接近 DRAM，这对于现在的计算机系统软硬件结构提出巨大的挑战。

2.2.1 磁盘阵列

在传统磁盘阵列方面，目前的研究主要集中在下述几个方面：大规模磁盘系统的新型组织方式，包括动态扩容；磁盘阵列的重构优化方法，包括前端 I/O 优化，例如 IDO^[25]；考虑各种失效事件，包括扇区失效的可用性等问题；基于固态盘的磁盘阵列或者非易失缓存的磁盘阵列等。例如，Plank^[26]针对 SSD 磁盘阵列更容易发生扇区失效的问题，提出 SD 编码来容忍磁盘和扇区混合错误。相比于传统纠删码，SD 使用的存储空间更少。

伽罗华域（Galois Field）运算构成里德所罗门和其他纠删码的技术基础。大部分伽罗华域算法依赖于乘法表和离散对数执行。但是 Intel 流式 SIMD 扩展指令具有 128 位操作能力，运行伽罗华域运算更快。Plank^[28]利用商用处理器的 128 位扩展指令集增加可变域计算的处理性能，使得计算延迟不再成为高冗余磁盘阵列或纠删码的瓶颈，并且这个计算软件是开源的。

由于云存储环境下，应用被强迫使用同一台计算机和共享物理磁盘资源，因此可能会出现成千上万用户并发访问磁盘阵列的情况，而磁盘对于 I/O 竞争是极其敏感。随机 I/O 能够使磁盘阵列吞吐率降低数个量级，并损害运行在同一阵列上的多个应用性能。日志结构的存储设计使用顺序化方式减轻写竞争，但在应用程度读时，垃圾回收会触发读写竞争问题。Gecko^[27]是消除读写竞争的新日志结构设计，靠链接小数量驱动器装载单个日志，有效分离日志主体和尾部，尾部写操作和日志读操作分别在不同磁盘上进行。

由于新型高性能固态存储和云存储的迅猛发展，加之磁盘阵列已经有 20 多年的应用历史，因此主流学术领域研究单纯磁盘阵列的热情在下降。相比之下，基于新型存储机理和器件的系统结构则受到更多的关注。

2.2.2 固态存储设备

以磁盘为代表的外存储设备的 I/O 性能一直是计算机整体性能的瓶颈。固态存储设备的引入，为性能改善提供了巨大空间，但是要真正发挥出其性能优势还需要做大量研

究工作。

基于 Flash 固态硬盘的高 IOPS 特性彻底改变存储系统的设计方式，针对闪存的固态硬盘是近年工业和学术界研究的热点。但是 NAND 闪存的擦除性能和磨损问题限制了其应用场景，因此引入闪存转换层（Flash Translation Layer, FTL）设计、缓冲区管理算法设计以及开发闪存的并行性能来改善上述问题。FTL 算法能够实现地址映射、垃圾回收和磨损均衡等机制。但其复杂的内部机制导致优化应用程序、文件系统和内部 SSD 等方面尚需做大量研究。DFTL^[22] 将整个映射表存储在闪存中，将闪存从逻辑上分为数据块区域和映射块区域，分别用于存储数据和映射表信息。然后根据需求动态加载部分映射表（Cached Mapping Table, CMT）到 RAM 中，从而处理访问频繁的请求。由于商业设备的封闭性，大部分针对 SSD 内部的研究（例如 FTL 的改进）只能使用仿真器和模拟器，没有真实系统实现，很难判断设计方法的实际优势。Saxena^[24] 描述了两个 SSD 设计方案，分别改变设备的内部工作和主机操作系统接口，并在真实硬件上加以验证。

FTL 虽然在块设备层减少了闪存擦除次数，但由于对于上层文件系统甚至应用的 I/O 行为语义无法传递到块设备。导致文件系统的写放大，包括日志、元数据同步、页对齐更新，会产生额外写操作，加重 SSD 的磨损。Lu^[23] 提出基于对象的 FTL 设计——OFTL，靠利用页元数据的局部性，保证索引元数据的懒惰持久，在保证一致性的同时消除日志；通过粗粒度块状态维护减少持久空闲空间管理的开销；使用字节单元的存取接口；并且能够压缩和合并小更新及元数据。还有部分研究使用 Trim 机制，把文件系统语义传递给 SSD 设备，以优化其性能。

为了进一步直接利用闪存提升应用程序的性能，充分利用闪存芯片间的并发性，避免复杂 I/O 栈和 FTL 的延迟，以 FusionIO 为代表的研究直接利用 PCI-e 作为闪存存储接口，加速器中设计上百个并发通道连接闪存芯片，板上硬件控制器仅负责多通道管理和调度，应用程序可以利用并发通道优化读写过程，所有优化功能通过主机上的软件驱动实现（包括类似 FTL 的功能）。对于典型的 Web 或者数据库应用能够提供百万级 IOPS。

目前也开始考虑使用 PCM 芯片替代闪存芯片开发使用 PCI-e 的 PCM 加速卡。但是研究表明，虽然 PCM 的 PCIe 加速卡相对于基于闪存的 PCIe 卡读延迟减少一半，但是写延迟、读写吞吐率和闪存相比，没有明显的优势，其中 60% 的延迟是由 PCIe 块接口和 I/O 栈软件驱动产生的。修改应用程序和文件系统实现操作系统旁路是优化研究的重点。另外，研究发现，当 IOPS 超过 800K 时，I/O 驱动程序的粗粒度锁将成为系统的瓶颈。多核所产生的锁机制也是制约固态存储潜力的重要因素，因此目前针对每个核建立独立的 I/O 请求队列以减少不必要的锁开销是一个重要的研究方向^[31]。

由于大数据存储和处理的需要，多级分布式计算和存储结构为每个存储部件增加处理单元，从而将数据和 I/O 密集的计算分布到存储部件，允许应用程序挖掘数据的局部性，实现数据的就近处理，避免数据集中处理所导致主机和存储的重复数据传输。Minerva 就是一种典型的计算就近存储的方案^[29]，Minerva 提出的编程模型是设计一种通用命令状态机制用于分配计算，确保存储中的数据和运行在主机上应用数据的一致性视图。存储处理器生成中断，保证主机协助干预解决。主机软件驱动负责应用程序的安装、管

理、监控和释放存储处理器。使用 Moneta 能够减少 77% 的软件开销，0.5KB 的 I/O 请求时产生 30 倍 IOPS 提升。QuickSAN^[30] 在 SAN 环境下可实现分布式存储，通过 SAN 将连接每个客户端上的 SSD，尽量使用局域性实现本地存储计算，从而减小延迟开销；SSD 直接使用 NIC 减少块传输开销，提高效率。

2.2.3 存储级内存

过去一般使用动态内存作为低速外部持久存储的高速缓存。随着固态存储技术的发展，其非易失的特点以及接近甚至超过 DRAM 的性能越来越受到关注。因此，将固态内存高性能和鲁棒性与传统硬盘的持久存储能力和低成本结合起来，构造存储级内存（Storage Class Memory，SCM）成为当前研究的热点。

SCM 的需求是每 GB 的价格不超过磁盘的 3~5 倍，读延迟小于 200ns，写和更新延迟小于 1μs。每秒的 IOPS 大于 10 万次，带宽大于 1GB/s，写次数为 $10^9 \sim 10^{12}$ 之间，能耗小于硬盘功耗的十分之一。目前 SCM 的候选对象包括 NAND 闪存、PCM、MRAM 和 RRAM。

最后一级 cache（Last-level cache，LLC）的性能对于多核系统的性能和能耗有决定性的影响。目前片上内存主要采用 SRAM、STT-RAM 和 eDRAM，其中 eDRAM（embedded dynamic random access memories）已应用于 IBM Power7 处理器中，Chang^[32] 分析了 STT-RAM 作为 LLC 的设计。

使用闪存作为 cache 已经研究了多年。但是，cache 替换策略依赖于内存的全索引， N 个索引对于每个关键字需要额外 $\log N$ 的位，对于大容量 flash 缓存（数十亿个 key-value 对象），元数据开销是非常大的。另外，使用闪存做 cache 存在写放大问题，而传统 RAM 不会产生的写，包括块擦除和页写，这个问题也称为 cache 层写放大（Cache Layer Write Amplification，CLWA）。TBF^[33] 使用 key-value 和 SSD 作为磁盘的 cache，使用两个内存 bloom 子过滤器维护最近信息，并进行命中检查，使用 clock 模拟 LRU 算法，从而减少元数据的额外开销。

由于 SCM 的容量一般较大，在存储服务器上使用这些大缓存需要预热（warm-up），所以可能需要几小时或几天时间。Bonfire^[34] 是一种加速 cache 预热的机制，它监控存储服务器负载，记录重要的预热数据，具有高效的预装载存储级 cache，能在数据中心块级 trace 的分析基础之上提供启发式预热机制。

混合内存把固态介质（PCM）和 DRAM 集成起来，用户仍然使用 loads/stores 存取内存空间。传统操作系统把运行库加载到 DRAM，而混合内存会把重复使用的程序和数据放置在固态内存区域，并映射到应用程序的地址空间，这些只读数据包括系统、用户空间库、系统服务程序和内核数据等。日志技术广泛应用于现代文件系统，可使用 PCM 和 STT-MRAM 设计新的缓存和日志机制^[37]，提供可靠性和系统失效后的快速恢复。

典型应用中会经常出现对同一位置的多次重复写的问题，例如重新使用最近释放的数据，造成重复写的另一个原因是元数据会嵌入到分配/回收块中，元数据操作也会频繁写一个位置。而固态内存的数据丢失来源于单元磨损和错误写。元数据损坏导致数据丢

失，而错误写会有引起虚拟内存保护。多次分配同一个块会造成硬件的磨损，并且导致更新写频繁的应用性能下降。这些问题都是传统内存分配算法造成的。因此需要创建面向 NVRAM 的新分配算法和数据结构^[35,36]，从而减少 NVRAM 的元数据更新数量；保证元数据在异常情况下的健壮性，避免持久数据的丢失。一种解决方法是建立不分配空闲列表，在一个时间段内不进行分配，使得所有块分配的频率小于 $1/T$ ；另外一种方案是尽量保持元数据在内存中，DRAM 中的位图记录所有块的状态；针对 DRAM 中的空闲列表加速分配；Checksum 放在每个块的头部；并通过虚拟内存保护保证内存中持久数据不会被修改。由于嵌入应用环境对于存储器件尺寸和电量有限制，非易失性内存能够取代 DRAM，但是 NVM 的写延迟和次数限制其广泛应用，对此 Hu^[38] 提出一种启发式串联调度，其能够减少平均内存存取时间 9.99%，并增加寿命 26.66%。

2.3 存储系统

2.3.1 结点系统

过去通常使用磁盘阵列作为存储结点，但由于磁盘阵列使用块接口，不支持文件语义，也不能对于数据进行处理，因此目前越来越多地使用存储服务器提供存储服务。这样可以实现存储服务的多样性，并且使得数据能够得到就近处理，这也是 MapReduce 这种计算模式的结构基础。针对存储结点的设计和优化就成为后续大规模存储系统的基础。存储结点对于数据的管理和存取是通过文件系统实现的，而数据从处理器到存储介质的过程中穿越文件系统和驱动等多个操作系统层次，称为存储协议栈。研究存储结点的重点是文件系统。

服务器对于数据管理依赖于文件系统。Lu^[39] 研究了文件系统代码的演进，分析了 8 年内 Linux 文件系统中 5079 个补丁。研究发现，50% 左右的补丁以软件维护为目的，反映代码必须简单和可维护；40% 补丁是针对 bug 错误，并且 bug 错误补丁数量没有逐年减少。在 bug 补丁中，50% 以上是语义错误，这是因为文件系统语义不易理解，很难被检查出来；而并行 bug 占 20%，其中原子性破坏和死锁类 bug 占大多数；剩余的是内存 bug（内存泄漏和空指针）和不正确错误代码处理。

在存储结点容量不断增加的情况下，即使采用日记机制或者其他预防技术，硬件错误或者软件崩溃失效都有可能导致文件系统破坏，甚至数据丢失。一致性检查 (fsck) 能够检查这些文件系统损坏，将已破损的镜像恢复到可用的状态，但它通常在事后运行，并且执行很慢，对于大存储系统会导致明显的性能下降。针对这一问题，Ma^[40] 认为检查器是整个文件系统的关键部件，因此设计了 rext3 支持快速文件系统检测，rext3 文件系统自标识它的元数据并和数据存放在一起，除去在检查时昂贵定位和树遍历的开销，结果表明，对于文件系统的修改使得运行 ffscck 扫描和修复文件系统带宽接近存储设备的峰值性能。

当前存储结点面临更加复杂的应用行为，数据量不断增长，数据查询更加复杂，数

据存取模式更加复杂和随机化。传统存储系统在优化部分负载的同时，也会影响其他负载性能，原因在于现有存储系统数据结构的限制。Shetty^[41]设计了新的负载独立数据结构——VT-tree，它扩展了 LS-Mtree，能有效处理顺序和文件系统负载。VT-tree 提供并行存取机制，为文件系统、数据库 API、事务保证、大小数据项提供模式无差别的性能。

存储结点通过网络文件系统扩展自身的存储空间。通过网络存取数据有可能产生数据一致性问题。目前的研究考虑主机端采用闪存作为远程存储的缓存，但是采用写穿和读缓冲策略虽然能保证严格的一致性，但会导致写性能下降；采用写回策略，既牺牲一致性，还有乱序排出的问题。Koller^[42]因此提出两种写回 cache 策略：次序写回和日志写回，利用闪存缓存改善网络存储性能。

虚拟机被广泛应用于数据中心，虚拟机通常使用附网存储（NAS）保存数据。但对于 NAS 服务器，应用从物理转换到虚拟客户端会动态地改变 I/O 负载行为，原因在于虚拟磁盘镜像会封装客户文件系统，并且来自多个虚拟机 I/O 请求流会多路复用物理通道。遗憾的是，现在 NAS 负载产生器和 benchmark 产生的负载仅仅典型地反映物理主机的 I/O 行为。Tarasov^[43]开发了面向虚拟机的 NAS 负载 benchmark，首先研究虚拟化扩展怎样明显地改变现有的 NAS 负载行为，包括文件系统元数据操作在 NAS 层的消失，改变 I/O 大小，增加存取随机性；其次产生一组多样化的 benchmark 去合成虚拟化负载，这样无需安装完全的虚拟化环境和工作行为，就能够产生正确的虚拟化负载。

当前存储应用需求对于传统复杂的存储栈提出新的可靠性挑战。在电源失效时，需要研究存储部件表现出来的行为，这是构建健壮系统的基础。Zheng^[44]提出了新方法去揭示在电源失效时的块设备可靠性问题，其框架包括通过硬件特殊设计直接注入电源失效事件到设备，通过发生负载压迫存储部件，和检测各种错误的技术。应用该测试框架，已测试了 5 个厂商的 15 个商业 SSD。

2.3.2 大规模系统

当前数据中心通过成千上万的计算机结点提供巨大的存储空间和计算能力。分布式文件系统把分散到每个结点中的存储空间和数据管理起来，提供高性能、高可靠的存储服务。在分布式存储系统方面，已经不存在通用型的结构，目前研究者针对不同的应用类型设计新型存储系统及其分布式文件系统，包括集群存储、存储处理一体机和大规模网络数据服务等。在大规模分布式存储方面，涉及高效纠删码存储系统、列存储、键值对存储、P2P 存储系统和云存储等方面的研究。

目前，大量数据中心采用 Hadoop 分布式文件系统作为其存储集群系统的管理平台。HARDFS^[45]加强 Hadoop，防止内存或者软件错误所产生的失效静默行为（fail-silent），提出选择性轻量级版本机制。使用这种方法，HDFS 重要子系统执行的行为被第二个子系统实现所检查，这个实现使用轻量级和近似的数据结构。HARDFS 能够检测和恢复广泛存在的失效静默行为。

分布式文件系统通常为大规模数据分析应用而建造，而企业级存储具有不同的功能需求，通常基于 SAN 存储平台。为了确保在企业级数据上进行数据分析需要引入分离分

析存储 silo，但无论何时数据都需要跨 silo 的归档、拷贝和迁移，这会产生额外代价和无效率的数据管理。MixApart^[46] 使用集成数据 caching 和调度方案通过 MapReduce 计算能够分析企业级存储上的数据。通过前端 cache 层确保数据分析本地存储性能，而后端共享存储方案能够简化数据管理。

分布式文件系统的存储栈较长，处理环节较多，会导致 I/O 响应时间增加。扁平的数据中心存储（FDS）是一个高性能、容错、大规模、倾向本地的二进制大对象存储系统（Blob）^[47]。使用完全二等分带宽网络、数据和元数据分条及其流控，FDS 能够把应用程序产生的大量 I/O 分散到集群内有可用吞吐率和延迟预算的磁盘上，FDS 就可以围绕数据非局域性做出很多优化，磁盘也能够和其他磁盘进行全带宽通信，使得磁盘失效快速恢复。FDS 是专为大规模数据中心而设计的，完全分布元数据操作。FDS 应用能够实现超过 2GB/s 的单进程读写性能。FDS 的主要目标是挖掘集群存储磁盘到应用的全部带宽。大对象被划分为束，从而能够对于所有数据实现高度并行化存取，而元数据也能得到很大程度的分布处理。FDS 创造 2012 年磁盘到磁盘排序时间最短的世界纪录。

AOL 为 Web 应用提供可扩展的对象存储 HSS^[48]，对象存储的目标是消除针对每个应用的分离存储，实现计算和存储的一体化及局域化，避免发送数据到另外的服务器。HSS 具有下述功能：物理配置可知，通过文件复制的数据保护，背景数据检测和恢复，内部文件系统布局，基于高性能数据库的元数据管理，开源软件部件，系统扩展性和简单界面，网络复制平衡器，基于 IP 地址的权限控制，基于大小重定向和基于代理的文件服务，通用文件缓存，文件过期机制，文件存储应用元数据，文件审计，系统整体监控。

Lustre 系统一直是高性能计算平台所常用的分布式文件系统。在实际部署中，采用多级 I/O 处理方式，能够在一定程度上吸收突发输出。由于超级计算环境下 I/O 负载中 75% 是写操作，但是实际写带宽达不到峰值带宽，因此 xie^[50] 研究了共享 Lustre 系统中的写吸收行为，发现滞后结点造成写性能的下降。Henschel^[49] 测试具有 3500 公里距离、100Gbps 带宽的广域网平台上 Lustre 的性能，发现能够保证数据带宽达到 90% 的效率。

当前科学发现依赖大规模超级计算仿真和与之相匹配的数据分析。这些数据分析可以在小规模集群上离线进行，也可以在超级计算机上进行。遗憾的是，因为在计算机和存储子系统之间的增加额外的数据移动，所以这些技术会导致性能和能耗的低效率。ActiveFLASH^[52] 一种现场科学数据分析技术，使数据分析能够在数据驻留的 SSD 上进行，性能和能耗模型显示 ActiveFLASH 有效地提高了 HPC 的仿真性能。

2.3.3 云存储系统

云计算已经成为社会必要的信息基础设施，是信息时代社会化大分工的必然，能够解决越来越复杂的技术和个人日常需求的矛盾。云存储是云计算中最必要也是最容易实现的一种基础服务。云存储构建于大规模存储系统之上，但它更关注资源调度、性能隔离、数据安全和服务质量保证等问题。

云存储系统往往以 SLA (Service-Level Agreement) 协议的方式对用户或企业承诺服务质量，而可靠性和性能是 SLA 核心内容。即使少量的数据丢失或者很短时间内不可用，以及性能下降都会导致存储服务提供商丧失数据存储安全可靠性方面的信用，甚至面临严重的法律诉讼或经济损失。

共享存储服务在商用云中广泛应用。大部分系统提供弱的性能隔离，很难保证所有租户间的公平性。不当行为或者高需求租户可能导致共享服务产生过载，破坏其他正常用户，导致不可预测的性能波动并且破坏 SLA。Pisces^[53]是一个面向数据中心规模的基于共享键值对的存储系统，它保证租户性能隔离和公平性。当前对于多租户资源分配的方案是基于虚拟机分配，或者基于硬速度限制，这种限制必须假设相同负载需要相同的使用率。在共享服务的全局资源上，Pisces 针对每个租户保证权重公平共享机制或者最低吞吐率。即使当不同租户的分区物理上在一起，或者请求对于不同分区的不均衡，又或不同服务器剩余资源随时间变化产生的瓶颈，Pisces 都能实现目标。Pisces 靠分解这些公平共享问题进行四个补充机制的组合，这些机制包括分区位置、权重分配、副本选择和权重公平队列。最终在不同时间尺度，Pisces 组合上述机制提供系统级的最大最小边界公平。

Cake^[54]是为了达到吞吐率和延迟服务要求，针对共享分布式存储环境而实现的协调多资源调度器。Cake 使用两级调度策略实现高层服务质量保证 (SLO)，第一级调度器控制磁盘和处理器资源，具体功能包括提供差异调度、将大请求分布到多个 chunk、限制过度设备请求。第二级调度器协调多个一级调度器，映射高层服务质量保证目标到实际物理设备的调度变量，这些变量能够随着负载的变化而进行自动调整。

云存储服务商能够提供虚拟无限存储空间，数据用户为了避免服务商的锁定，或者增加可用性和时效，力图在多个云存储上保存关键数据。Scalia^[51]是一个云代理解决方案，可以根据服务质量、价格和存取模式自动调整数据存放位置。

2.4 存储应用

当前大数据存储的应用多种多样，未来还会不断产生各种应用。很多存储应用属于专门研究领域，例如分布式并行计算、海量数据管理和检索、数据库管理、数据归档和长期保存等方面。下面仅列出和 I/O 存储技术紧密相关的四种典型存储应用相关技术点的研究情况。

2.4.1 数据精简

当前数据量以指数级规模增加，需要尽量减少实际的资源占用，从而提高存储效率。

重复数据删除技术用于发现和挖掘不同数据块之间的冗余。最常用的方法是划分数据到固定大小数据块 (chunk) 并通过较小的指纹标识冗余。文件内容能够通过合并 Chunk 指纹而被重建，这些指纹顺序存放在文件收据 (recipe) 中。当重删系统规模增加时，相应的文件收据将占据很大的磁盘空间，尤其当冗余率很高的时候。因此，Meister^[56]提出高效可

扩展压缩策略，可将文件收据大小减少 93%。

固定 Chunk 重删系统所面临的严重问题是 Chunk 分散（fragmentation），它导致完全恢复时的性能下降，最近的研究表明备份数据的恢复延迟随着系统运行时间而下降几个数量级。Lillibridge^[57]使用增加 cache 大小、容器限制和使用向前相似区域分析来减轻性能损失。容器限制是一个吸收时操作，以牺牲部分重删率来减少 chunk 离散的代价；而向前相似区域是一个新的恢复时 cacheing 和预取技术。

由于数据呈指数级增长，需要对于主存储进行实时压缩。但是在数据路径上增加压缩操作会消耗大量 CPU 和内存资源，并导致巨大延迟。Harnik^[58]使用两级过滤器快速标识压缩时能够明显减少存储空间的数据，减小压缩代价。第一级过滤器在数据集层（逻辑卷和文件系统），在那里估计剩余数据的全局压缩比。根据这个结果，可以对数据集压缩触发进行选择，或者使用第二级过滤策略。第二级过滤器策略以在线的方式检查写到存储系统的数据，并决定其压缩率。

Meister^[59]首先研究高性能计算环境下的重删问题，量化分析 4 个典型数据中心（BSC, DKRZ, RENCI, RWTH）。使用重删能够明显地减少存储容量。与传统的基于备份的重删不同，这个工作分析 PB 级的在线文件系统数据。结果表明，使用文件内部重删，有 20% ~ 30%（最高可达 70%）的在线数据能够被重删系统精简。但是全文件重删效果是不明显的。

对于已备份数据而言，有时需要删除特定的数据。但是重删系统中数据 chunks 可能属于多个所有者。尤其在分布式系统中，经常会有增加、移除结点或者结点失效的情况出现，这使得在分布式全局重删系统中删除数据面临挑战。Strzelczak^[55]描述了一种针对可扩展、内容寻址的全局重删系统中的删除算法，仅仅建立很小限制从而保证安全数据回收的可能，而删除操作是并行的，且能够和用户前端读写操作并行完成。

2.4.2 图运算

交互式图分析应用（社交网络分析）数据存取过程具有很大的非局域性和数据并行度，访问模式依赖于数据相关性，并且以读操作为主。针对实际社交网络和 Web 应用所产生的大图，其计算需要分布式集群及其相应算法的支持。典型分析操作本地内存访问大约 300Mref/s，远程内存访问大约为 100Mref/s，本地 NVM 存取大约 1Mref/s。

GraphChi^[60]是基于磁盘存储的图处理系统，它能够高效处理几百亿边规模的大图，它把大图分解为小部分，利用并行滑动窗口进行操作。GraphChi 能够利用 PC 集群完成高级数据挖掘、图挖掘和机器学习。

由于图分析具有串行小读的特点，可以充分利用非易失存储提供性能。Pomace^[61]利用 Non-volatile 内存加速 Grappa 图分析，让每个 I/O 请求做更多有用工作。相对于传统的基于单个字的并行方式从而减少存取延迟，Pomace 虽然增加延迟，但是其提高通道并行度，并实现预取字。它开发独立操作芯片，在使用闪存 SSD 时能达到 40Kref/s，而使用 PCM 时处理性能大于 5Mref/s。

在大图中查询中需要做大量的路径处理（Path Processing），由于这类处理具有很强

的因果关系，因此数据几乎不能预取，只能查询到数据后再前进，在这个过程中会产生许多无用的查询数据。固态存储具有延迟时间短、随机 I/O 存取性能高效的特点，如果针对常用的资源描述框架元数据模型（Resource Description Framework（RDF）meta-data model）进行特定的设计和优化^[62]，那么构建固态存储感知的查询系统能够加速查询性能。

2.4.3 存储安全

存储安全是数据安全的最后防线，因此存储系统也需要考虑数据安全问题。

当分布式系统规模不断增长，至少一个结点妥协（compromised）的概率增加。在很多超级计算、云计算、安全内容传输网络系统中，处理存储数据时仅使用未加密或者简单建模策略，主要依赖于物理隔绝，这样无法针对妥协计算机和内部恶意者提供安全保护。现有的加密方案也无法为大数据提供有效率的细粒度加密。Horus^[63]的目标是加密大数据集使用 KHT（Keyed hash tree）为不同区间的数据集产生不同的密钥，提供细粒度安全，一个区域的密钥不能使用在另一个区域。Horus 能够减少密钥管理和分布过程中的开销，通过提供端到端数据加密和减少对于信任操作者和云服务提供商的依赖。

最近出现了在线服务提供商使用或者泄露用户的隐私数据的事件。加密技术能够减少这种风险，但是不能消除。因为数据存取模式仍然会隐藏有用信息从而便于破解。Shroud^[64]是一个通用存储系统，能够隐藏运行服务器上的数据存取模式，从而保护用户隐私。Shroud 的功能是使用隐私保护的虚拟磁盘，用户查询地址无需透露块地址。Shroud 在几百个 TB 数据中隐藏存取动作，达到这个目标靠自适应遗忘（oblivious）RAM 算法，并且具有大规模并行性。

重删系统也存在着巨大的安全隐患，因为逻辑删除的数据有可能在物理上还存在，这对于重删系统提出挑战。Sanitization^[65]是从存储系统中安全擦除敏感数据的过程，保证有效地恢复到敏感数据未存储前的状态。

2.4.4 科学计算

随着计算能力不断超越 I/O 处理能力，I/O 已经成为高性能计算机领域的瓶颈。

当前双精度仿真数据经常采用精度无丢失保存，但是这种类型的科学分析理论上可以接受小错误，这也仅仅导致局部细节丢失。Jenkins^[66]开发出多精度级别的细节库（Precision Level of Detail，APLOD），可按照用户定义的字节界限区分双精度数据集。APLOD 参数化分析精度和 I/O 性能的平衡点，标定最大相对错误边界，维护 I/O 存取模式，降低操作代价。

在超大规模计算中，由于 I/O 性能的限制，使得及时保存大量原始仿真数据到持久存储变得越来越困难，一个可能的方案是改变数据处理的流水线模式，从后处理中心模式转变为基于现场处理和迁移处理并发模式。使用主计算资源的计算是现场处理，而基于异步数据传输的一组二级资源上的计算是离线迁移处理。Bennett^[67]设计和实现了应用于大规模科学仿真中的三种常用分析技术，包括算法开发、协调不同分析流执行的资源

调度策略，以及在两种处理间实现高效数据移动的技术。

层次性、多解析度数据表示确保大规模仿真的交互性分析和虚拟化。这些技术的未来应用将使用层次性 Z 序列 (hierarchical Z ordering) 存储高性能计算仿真输出，这个 HZ 序列将数据从笛卡儿坐标空间转换一维局域于解析度的顺序阵列。但当仿真数据的维度不是 2 的偶数幂时，并行 HZ 序列会产生离散内存和网络存取模式，从而限制 I/O 性能。Kumar^[68]提出的新仿真数据集并行 HZ 序列技术，把仿真数据重构为大数据块 (2 的幂数倍)，提高 I/O 效率。

3 国内研究进展

目前国内对于计算机存储系统的研究、开发和应用蓬勃发展，从国家到企业都越来越重视数据存储技术的研发工作。国家科技部门最近几年支持了一大批存储器件、系统和应用方面的科研项目，提高了我国在计算机存储领域技术自主研发和创新能力，取得一系列突出的进展和成果。

由于每个计算机系统都有存储子系统，所有的数据必须保存在存储系统中，读写数据都要和存储系统打交道，而数据中心又是以数据为核心的，因此国内很多关于计算机系统和应用方面的研究都涉及存储系统。而本文描述的国内研究进展主要涉及针对存储系统的专门研究。

3.1 新型存储器件

在新型固态存储器件研究方面，与国外对 PCM、MRAM、RRAM 的研究相比，国内在这几个方面的研究与国外存在着很大的差距，不仅从事相关研究的机构以及高校很少，而且国内的材料、制造工艺和体系构建等方面的研究也与国外存在差距。

目前国内研究 PCM 和 MRAM 相关技术的机构以及公司主要有中国科学院上海微系统与信息技术研究院、中芯国际和 SST、北京时代全芯公司等。高校中仅有复旦大学、上海交通大学、华中科技大学有 PCM 的研究团队。

中国科学院上海微系统与信息技术研究院对相变材料体系有比较系统的研究，成功开发出中国第一款具有自主知识产权的相变存储器 (PCRAM) 芯片。中芯国际与其联合搭建的 130nm 和 45nm PCM 芯片工艺基本上达到国际领先的水平。SST 也拥有完善的芯片设计系统技术，目前三家正联合研发产品级 PCM。此外，我国依托华中科技大学筹建的武汉国家光电实验室信息存储材料及器件研究团队在高密度、低功耗电阻式相变存储器技术、相变随机存储器功能芯片、相变存储器芯片的关键材料等多个有关相变存储器领域技术的研究，已经取得了重大成果，在 2008 年就研发出了具有简易读写以及擦除功能的相变存储器功能芯片。武汉光电国家实验室还在存储级别对 PCM 的混合存储领域、原型平台、PCM 模拟器等多个方面进行研究。目前已经有了一定的研究成果，研制出可

以使用的 PCM 子卡。与此同时，还开发出了 PCM 仿真器。

复旦大学也开展了 MRAM 的研究，华中科技大学和清华大学正开展忆阻器原理、器件和系统相关的研究，也有成果发表在科学通信上。也有一些其他研究机构在从事新型存储器件的研究。

在新型光存储方面，清华大学一直在从事多维度高密度全息光存储原理、器件和系统研究，而北京理工大学正在研究 TB 级同轴全息光盘的研究，华中科技大学针对未来高密度光盘编码、伺服和大规模光盘库开展研究。

3.2 存储系统

针对存储系统的研究在国内广泛开展，华中科技大学、清华大学、中科院计算所、国防科技大学、北京大学、西北工业大学等都有专业从事存储系统的研究团队，也取得一系列成果。

最近几年国家从器件、设备和系统等多个方面对于存储系统研究给予支持，包括 2012 年国家自然科学基金委对于存储领域进行的重点支持，科技部从 973 项目、863 项目和重大支撑计划等方面也对国内存储研究给予资助和支持。

在磁盘阵列方面，华中科技大学提出一系列磁盘阵列新型编码，包括 H-Code、P-Code、M-Code 等；在磁盘阵列重构优化方面提出了 VDF、PRO、Workout 等技术。清华大学在磁盘阵列扩展性等方面获得一些成果，论文发表在 FAST2011 上；中国科技大学、南开大学和成都电子科技大学等在新型编码方面也提出一些新技术，一些成果也在国际著名期刊和会议上发表，例如，针对 RDP 单盘失效重构的 RDOR 方法的研究成果发表在 SIGMETRICS2011 上。

在固态存储系统方面，华中科技大学在多通道闪存阵列和调度以及存储级 PCM 方面取得一些成果，实现利用芯片和通道的高度并行性进行调度。清华大学在 2013 年 FAST 会议上发表有关减少文件系统写放大的论文（其思想在前面介绍过）。重庆大学和国防科技大学在固态存储系统设计和优化方面也做了许多实际工作。

在超大规模存储方面，国防科技大学开发适应天河超级计算机 I 和 II 的高性能存储系统，其核心是分布式并行文件系统的研究；华中科技大学开展分布式文件系统的研究，包括基于内核级的分布式系统和用户态轻量级的分布式文件系统，也开展云存储相关的研究；北京大学开发出基于 P2P 的云存储系统；中科院计算所一直在开发适合各种应用的分布式存储集群。相对于国内研究机构，以淘宝公司、百度公司和腾讯公司为代表的互联网公司更是构建满足自身庞大业务需要的专用大规模存储系统，取得实际的成果。另外，也涌现出一大批“网盘”应用，它们的后台也是基于各种分布式文件系统或者就是大容量主机系统。在这个方面，国内产业与学术界的合作和国外相比还是相对缺乏。

另外，在存储系统能效优化方面，华中科技大学对高能效存储控制器和分布式存储系统进行了研究，以期降低数据中心的数据存储能耗。另外，也开展了基于光盘的大数据长期归档的研究。国防科技大学等单位在国家支撑项目支持下也开展了大量存储设备

能耗优化的研究。

除国内研究机构和大学之外，国内以华为公司为代表的企业也在从事固态存储器件和系统的研究，华为公司的 SSD 加速卡已经可以和 FusionIO 的性能进行竞争，目前也在开展 PCM 相关的产品化工作。

4 国内外研究进展比较

对比国内外研究进展，虽然近几年我国在存储技术方面取得一系列突破性进展，部分研究工作能够被国际相关著名学术会议和期刊所接受，也被国际同行广泛认可。但是整体而言，尚缺乏源头性创新的研究工作，大部分研究工作处于跟踪水平，只在少量技术点上达到国际水平。

和国内其他计算机系统技术研发一样，由于学术界缺乏产业界支持，包括资金、问题发现和实践验证，很多研究仅仅停留在概念和原理设计及验证上，即使有少量亮点成果也往往是国外原创性工作的细节扩展，因此影响力较低、持续性较差。

除了少数企业在局部或者应用技术研发上予以支持，目前计算机存储技术的研发投资主要来源于国家各种科技项目支持，因此产学研没有有效相互转换和推动。与其他计算机系统研究问题一样，由于没有实践，因此对很多技术细节的掌握和理解还非常缺乏。另外，国家每个项目竞争性和独立性也造成国内研发工作不能成体系，研究深度不足，实际应用的能力较差。

在存储介质方面，国内研发机构或企业到目前为止还不能完全独立设计、生产大容量存储介质和设备，虽然有一些研究所能够生产样品，或者购置生产线生产，但无法支持基础学术研究，也无法支持实际的产业发展。国内研究往往依赖于外国学术报道和论文等公开数据，虽然在一些概念性设计方法方面取得一定进展，但是整体上无法与国外学术和产业界研究实力相比。另外由于国内材料、电子和计算机系统结构属于不同专业，全面整体研究尚有欠缺。在存储级内存方面，由于没有样品仅有少量 PCM 和 MRAM 的研究，因此目前国内主要还仅限于研究闪存。而光盘在过去十年容量增加较慢，国外研究也较少，在新型多维大容量光盘研究领域，国内研究反而具有一定优势。

在存储设备和结点方面，由于其能够在市场上直接销售，因此国内有很多企业开发相关产品。正如处理器一般选用 Intel 和 AMD 一样，国内存储系统也往往直接来源于国外成熟技术或者产品集成。存储系统和其他计算机系统一样需要各种软硬部件技术的支撑，但国内研究难以在整体技术上进行把握，导致大部分研究工作只能从部分部件入手进行局部优化和改进，或者纯粹进行应用开发。在新型 SCM 操作系统和专用服务器方面，国内研究还未深入开展；而磁盘阵列方面的研究，由于不是国际学术界的热点，国内研究反而较为突出。国内研究机构虽在重删、虚拟化、备份等存储服务和应用方面提出部分优化算法，但是整体发展较慢。在产业方面，存储服务器使用商用软硬部件集成，技术含量较少；而国内企业研发出的中低端磁盘阵列已占据一定市场，但从事高端磁盘

阵列研发的仅有少量公司，离国际水平还有一定差距。

在大规模存储系统方面，目前高性能计算平台、企业数据中心和互联网数据中心都需要大规模存储系统，国内研究者实际运行的系统主要来源于国外企业产品或者开源项目，国内研究者基于这些平台能够针对应用需要开发出调度算法、管理监控系统和自动构建技术。由于大数据分析的需要，目前针对 MapReduce 模式研究与应用获得国内学术和企业界的关注，但目前还是处于跟踪和应用阶段。

5 发展趋势与展望

存储器件是计算机存储的物质基础。磁盘仍然是最廉价和普遍的大容量在线存储介质，这种状况在未来 5~10 年内应该不会改变，针对磁头的研究是提升磁盘容量的关键。闪存和硬盘各有特点，具有明显的互补性，因此混合磁盘技术能够发挥各自的优势，也是未来基本存储设备。

相变存储器和磁内存有可能替代现有的动态内存甚至 cache，能够极大地改变整个计算机系统软硬件结构，并有可能深刻地改变现有操作系统架构和功能，以数量级的方式提升存储性能，为系统整体性能和能耗优化提供巨大的设计空间。因此存储级内存研究将受到更多的关注。

忆阻器不仅可作为存储器件，而且可能成为计算存储一体化模式（人脑模式）的载体。在大数据时代，历史数据具有重要的价值，而 TB 级甚至 PB 级多维高密度光盘是长期数据高效能保存的重要技术。玻璃和 DNA 存储等新型存储技术是极具潜力的超长期数据存储技术，但十年内尚不具备商业价值。

分布式计算机存储结构也是未来技术发展的重要方向。由于大数据存储和处理的需要，多级分布计算和存储结构为每个存储部件增加处理单元，实现局部数据的就近处理，避免数据集中处理所产生的迁移。

未来典型应用负载行为能够被进一步精确定义、提取、预测和模型化，利用负载模型能够开发出面向场景的最优能效系统架构。而动态适应存储结构也是未来研究的重要方向。

PB 级甚至 EB 级存储系统会出现并获得极大的发展，就像 CPU 存在能量墙一样，大规模存储系统的资源和能量消耗的问题尤其重要，仅仅靠规模增长是无法解决这一问题，提高大规模系统整体能效是不可避免的研究问题。

6 结束语

30 年前，存储系统更多地以计算机外部设备的角色出现，其服务对象主要为中央处理器。20 年前，网络存储系统的出现和发展，使用户数据存储和管理成为存储系统必须

考虑的问题，而存储系统也从计算机部件成为功能独立的计算机系统。十年前，互联网发展突出海量数据收集和整体的重要性，集群存储把数据处理和存储紧密联系在一起；当前，云存储发展能够降低所有用户保存和使用数据的成本，存储服务成为必然。更多数据被产生，更多数据被消费，数据的重要性与日俱增，未来消费者不再也不可能关心存储系统的细节，而存储技术需要弥补介质到服务需求的巨大鸿沟。

参考文献

- [1] Richard Wesel, Zhang Tong, Adam Williamson, et al. Dynamic Voltage Allocation Based on Mutual Information for NAND Flash Memory [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [2] Myeongwoon Jeon, Jungwoo Lee. Cell to Cell Interference Mitigation Techniques for Flash Memory [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [3] Oliver Hambrey, Thomas Parnell, Oleg Zaboronski. Exploiting Variation within MLC NAND to Improve Post ECC reliability [Z]. The 4th Annual Non- Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [4] Qin Minghai, Eitan Yaakobi, Paul H Siegel. Inter- cell Interference Free Codes for Read/Write in Flash Memories [Z]. 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [5] Jiang Anxiao(Andrew), Li Yue, Jehoshua Bruck. Correcting Errors in MLCs with Bit-fixing Coding [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [6] Ravi Motwani, Chong Ong. Digital Signal Processing for High Endurance Solid State Drives [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [7] Zhang Xinmiao. Low-Energy and Low-Latency Error-Correction for Phase Change Memory [Z]. The 4th Annual Non- Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [8] Eran Sharon, Idan Alrod. Coding Scheme for Optimizing Random I/O Performance Sandisk [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [9] Noha Elarief, Bella Bose, Samir Elmougy. Limited Magnitude Error Detecting Codes for Flash Memories [Z]. The 4th Annual Non- Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [10] Frederic Sala, Ryan Gabrys, Lara Dolecek. Dynamic Threshold Schemes for Multi-Level Nonvolatile Memories [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [11] Toshihiro Sugii, Yoshihisa Iba, Masaki Aoki, et al. High-performance STT-MRAM and Its Integration for Embedded Cache Memory [Z]. The 4th Annual Non- Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [12] Subir Chowdhury, A 12 MHz, Sub Micro Amps Leakage and Low Active Power 64KB FRAM With Novel Design Techniques [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [13] Bipin Rajendran, Liu Yong, Jae-sun Seo, et al. RRAM Devices for Large Neuromorphic Systems [Z]. The 4th Annual Non-Volatile Memories Workshop 2013 , CA , March 3-5 , 2013.
- [14] Mohammed Affan Zidan, Hossam Aly Hassan Fahmy, Muhammad Mustafa Hussain, et al. Memristor-based Memory: The Sneak Paths Problem and Solutions [J]. Microelectronics Journal , October 29 , 2012.

- [15] Alvaro Padilla, Kumar Virwani, Rohit S Shenoy, et al. Access-Devices Based on Mixed-Ionic-Electronic-Conduction (MIEC) Materials for Multi-Layer Crosspoint-Memory: Yield, Speed, and Scaling [Z]. The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [16] H Zhou, A Jiang, J Bruck. Error-Correcting Schemes With Dynamic Thresholds In Non-Volatile Memories [Z], IEEE International Symposium on Information Theory(ISIT2011), Saint-Petersburg, Russia, July, 31-August, 5, 2011.
- [17] Zhao Kai, Zhao Wenzhe, Sun Hongbin, et al. LDPC-in-SSD: Making Advanced Error Correction Codes Work Effectively In Solid State Drives [Z]. The 11st USENIX Conference on File and Storage Technologies (FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [18] Peter Zijlstra, James W M Chon, Min Gu. Five-dimensional Optical Recording Mediated By Surface Plasmons In Gold Nanorods [J]. Nature, 2009, 459(7245): 410-413.
- [19] Li Jianhua, Cao Liangcai, Gu Huarong, et al. Orthogonal-reference-pattern-modulated Shift Multiplexing For Collinear Holographic Data Storage [J]. Optics Letters, 2012, 37: 936-938.
- [20] George M Church, Yuan Gao, et al. Next-Generation Digital Information Storage in DNA [J]. SCIENCE, SEP 28, 2012, Vol(337): 1628.
- [21] Nick Goldman, Paul Bertone, Siyuan Chen, et al. Towards Practical, High-Capacity, Low-Maintenance Information Storage In Synthesized DNA[J]. Nature, vol(494), no: 7435, 20130123: 77-80.
- [22] Gupta A, Kim Y, Urgaonkar B. DFTL: A Flash Translation Layer Employing Demand-based Selective of Page-level Address Mapping [Z]. In Proceeding of ASPLOS' 09, 2009.
- [23] Youyou Lu, Jiwu Shu, Weimin Zheng. Extending the Lifetime of Flash-based Storage through Reducing Write Amplification from File Systems [Z]. The 11st USENIX Conference on File and Storage Technologies (FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [24] Joshua Schiffman, Yiyi Zhang, et al. Getting Real: Lessons in Transitioning Research Simulations into Hardware Systems [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [25] Suzhen Wu, Hong Jiang, Bo Mao. IDO: Intelligent Data Outsourcing with Improved RAID Reconstruction Performance in Large-Scale Data Centers [Z]. The 26th Large Installation system Administration Conference (LISA2012), San Diego, CA, 2012.
- [26] James S Plank, Mario Blaum, James L Hafner. SD Codes: Erasure Codes Designed for How Storage Systems Really Fail [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [27] Ji Yong Shin, Mahesh Balakrishnan, Tudor Marian, et al. Gecko: Contention-Oblivious Disk Arrays for Cloud Storage [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [28] James S Plank, Kevin M Greenan, Ethan LMILLER. Screaming Fast Galois Field Arithmetic Using Intel SIMD Instructions [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [29] Arup De, Maya Gokhale, Rajesh Gupta, et al. Minerva: A Compute Capable SSD Architecture for Next-Generation Non-Volatile Memories [Z], The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [30] Adrian Caulfield, Steven Swanson. QuickSAN: A Storage Area Network for Fast, Distributed, Solid State

- Disks [Z]. The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [31] Matias Bj?rling, Jens Axboe, David Nellans, et al. Linux Block IO: Introducing Multi-Queue SSD Access on Multicore Systems [Z]. The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [32] Mu-Tien Chang, Paul Rosenfeld, Shih-Lien Lu, et al. Technology Comparison for Large Last-Level Caches(L3 Cs): Low-Leakage SRAM, Low Write-Energy STT-RAM, and Refresh-Optimized eDRAM [Z]. 19th IEEE International Symposium on High-Performance Computer Architecture(HPCA-19), Shenzhen, China, 2013.
- [33] Cristian Ungureanu, Biplob Debnath, Stephen Rago, et al. TBF: A Memory-Efficient Replacement Policy for Flash-based Caches [Z]. Proceedings of the 29th IEEE International Conference on Data Engineering (ICDE), Brisbane, Australia, April 8-11, 2013.
- [34] Yiying Zhang, Gokul Soundararajan, Mark W Storer, et al. Warming Up Storage-Level Caches with Bonfire [Z], The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [35] S Venkataraman. Consistent and DurableData Structures for Non-Volatile Byte-Addressable Memory [Z]. In Proceedings of the 9th USENIX Conference on File and Storage Technologies(FAST2011). 2011.
- [36] Iulian Moraru, David Andersen, Michael Kaminsky, et al. From File System Designer to Persistent Data Structure Designer [Z]. The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [37] Eunji Lee and Hyokyung Bahn, Sam H Noh. Unioning of the Buffer Cache and Journaling Layers with Non-volatile Memory [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [38] Hu Jingtong, Chun Jason Xue, Qingfeng Zhuge, et al. Scheduling to Optimize Cache Utilization for Non-Volatile Main Memories [Z]. IEEE TRANSACTIONS ON COMPUTERS, 2012.
- [39] Lanyue Lu, Andrea C Arpaci-Dusseau, et al. A Study of Linux File System Evolution [Z]. The 11st USENIX Conference on File and Storage Technologies (FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [40] Ao Ma, Chris Dragga, Andrea C Arpaci-Dusseau, et al. ffsck: The Fast File System Checker [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [41] Pradeep Shetty, Richard Spillane, Ravikant Malpani, et al. Building Workload-Independent Storage with VT-Trees [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [42] Ricardo Koller, Leonardo Marmol, Raju Rangaswami, et al. Write Policies for Host-side Flash Caches [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [43] Vasily Tarasov, Dean Hildebrand, Geoff Kuenning, et al. Virtual Machine Workloads: The Case for New NAS Benchmarks [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [44] Mai Zheng, Joseph Tucek, Feng Qin, et al. Understanding the Robustness of SSDs under Power Fault [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.

- [45] Thanh Do, Tyler Harter, Yingchao Liu, et al. HARDFS: Hardening HDFS with Selective and Lightweight Versioning [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15 , 2013.
- [46] Madalin Mihailescu, Gokul Soundararajan, Cristiana Amza. MixApart: Decoupled Analytics for Shared Storage Systems [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15 , 2013.
- [47] Edmund B Nightingale, Jeremy Elson, Jinliang Fan, et al. Flat Datacenter Storage[Z]. The 10th USENIX Symposium on Operating System Design and Implementation (OSDI2012), Hollywood, USA, Oct 8-10 , 2012.
- [48] Daniel Pollack. HSS: A Simple File Storage System for Web Applications [Z]. The 26th Large Installation system Administration Conference(LISA2012) , San Diego, CA, 2012.
- [49] Robert Henschel, Stephen Simms, David Hancock, et al. Demonstrating Lustre Over A 100 Gbps Wide area Network Of 3, 500 km [Z]. The International Conference for High Performance Computing, Networking, Storage and Analysis(SC2012) , Salt Lake City, USA, Nov 10-16 , 2012.
- [50] Bing Xie, Jeffrey Chase, David Dill, et al. Characterizing Output Bottlenecks In A Supercomputer [Z]. The International Conference for High Performance Computing, Networking, Storage and Analysis (SC2012) , Salt Lake City, USA, Nov 10-16 , 2012.
- [51] Thanasis G Papaioannou, Nicolas Bonvin, et al. Scalia: An Adaptive Scheme For Efficient Multi- Cloud Storage [Z]. ACM Symposium on Cloud Computing, San Jose, USA, Oct 14-17 , 2012.
- [52] Devesh Tiwari, Simona Boboila, Sudharshan Vazhkudai, et al. Active Flash: Towards Energy-Efficient, In-Situ Data Analytics on Extreme-Scale Machines [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013) , San Jose, CA, USA, February 12-15 , 2013.
- [53] David Shue, Michael J Freedman, Anees Shaikh. Performance Isolation and Fairness for Multi-Tenant Cloud Storage [Z]. The 10th USENIX Symposium on Operating System Design and Implementation (OSDI2012) , Hollywood, USA, Oct 8-10 , 2012.
- [54] Andrew Wang, Shivaram Venkataraman, Sara Alspaugh, et al. Cake: Enabling High- level SLOs on Shared Storage Systems [Z]. ACM Symposium on Cloud Computing, San Jose, USA, Oct 14-17 , 2012.
- [55] Przemyslaw Strzelczak, Elzbieta Adamczyk, Urszula Herman-Izycka, et al. Concurrent Deletion in a Distributed Content-Addressable Storage System with Global Deduplication [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013) , San Jose, CA, USA, February 12-15 , 2013.
- [56] Dirk Meister, André Brinkmann, Tim Sü, et al. File Recipe Compression in Data Deduplication Systems [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013) , San Jose, CA, USA, February 12-15 , 2013.
- [57] Mark Lillibridge, Kave Eshghi, Deepavali Bhagwat. Improving Restore Speed for Backup Systems that Use Inline Chunk-Based Deduplication [Z]. The 11st USENIX Conference on File and Storage Technologies (FAST2013) , San Jose, CA, USA, February 12-15 , 2013.
- [58] Danny Harnik, Ronen Kat, Dmitry Sotnikov, et al. To Zip or Not to Zip: Effective Resource Usage for Real-Time Compression [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013) , San Jose, CA, USA, February 12-15 , 2013.
- [59] Dirk Meister, Jürgen Kaiser, Andre Brinkmann, et al. A Study On Data Deduplication In HPC Storage Systems [Z]. The International Conference for High Performance Computing, Networking, Storage and

- Analysis(SC2012), Salt Lake City, USA, Nov 10-16, 2012.
- [60] Aapo Kyrola, Guy Blelloch, Carlos Guestrin. GraphChi: Large-Scale Graph Computation on Just a PC [Z]. The 10th USENIX Symposium on Operating System Design and Implementation(OSDI2012), Hollywood, USA, Oct 8-10, 2012.
- [61] Jacob Nelson, Brandon Holt, Brandon Myers. Pomace: A Grappa for Non-Volatile Memory [Z]. The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [62] Manos Athanassoulis, Bishwaranjan Bhattacharjee, et al. Querying Persistent Graphs using Solid State Storage [Z]. The 4th Annual Non-Volatile Memories Workshop 2013, CA, March 3-5, 2013.
- [63] Yan Li, Nakul Sanjay Dhotre, Yasuhiro Ohara, et al. Horus: Fine-Grained Encryption-Based Security for Large-Scale Storage [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [64] Jacob R. Lorch, Bryan Parno, James Mickens, et al. Shroud: Ensuring Private access to Large-Scale Data in the Data Center [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [65] Fabiano C. Botelho, Philip Shilane, Nitin Garg, et al. Memory Efficient Sanitization of a Deduplicated Storage System [Z]. The 11st USENIX Conference on File and Storage Technologies(FAST2013), San Jose, CA, USA, February 12-15, 2013.
- [66] John Jenkins, Eric R. Schendel, Sriram Lakshminarasimhan, et al. Byte-precision level of detail processing for variable precision analytics [Z]. The International Conference for High Performance Computing, Networking, Storage and Analysis(SC2012), Salt Lake City, USA, Nov 10-16, 2012.
- [67] Janine C Bennett, Hasan Abbasi, Peer-Timo Bremer, et al. Combining in-situ and in-transit processing to enable extreme-scale scientific analysis [Z]. The International Conference for High Performance Computing, Networking, Storage and Analysis(SC2012), Salt Lake City, USA, Nov 10-16, 2012.
- [68] Sidharth Kumar, Venkatram Vishwanath, Philip Carns, et al. Efficient data restructuring and aggregation for I/O acceleration in PIDX [Z]. The International Conference for High Performance Computing, Networking, Storage and Analysis(SC2012), Salt Lake City, USA, Nov 10-16, 2012.

作者简介

曹强博士，华中科技大学武汉光电国家实验室光电信息存储功能实验室（暨信息存储教育部重点实验室）教授，中国计算机学会高级会员，中国计算机学会体系结构专业委员会常委，信息存储专业委员会委员，IEEE会员，ACM会员。承担、参与或者完成十多项国家研究项目，包括国家自然科学重点基金，国家自然科学基金，863课题和973课题。在国内外学术刊物上发表论文40多篇，其中包括USENIX ATC、DSN等一流学术会议和TC、TPDS、TOS等一流期刊，申请专利20多项，软件著作权8项。撰写出版专著《海量网络存储系统原理与设计》一部，2008年获湖北省技术发明二等奖。



计算系统虚拟化的研究新进展与趋势

CCF 系统软件专委会

摘要

计算系统虚拟化是一种在硬件和软件之间去耦合的技术，可以支持在一台物理机上同时运行多个操作系统，并通过虚拟机之间的动态资源调整以及虚拟机迁移等方法，提升系统资源利用率和可管理性。随着多核技术的发展，以及大规模数据中心和云计算应用的开展，对计算系统虚拟化带来了一些新的挑战。如何根据新的硬件特征以及大规模应用的特点，有针对性地提供系统虚拟化支持，是当前研究的热点。本报告以云计算作为背景，从嵌入式虚拟化技术、GPU 虚拟化技术、云计算平台中虚拟资源性能计量方法、云计算资源管理与调度以及虚拟化安全等方面对近年来国内外的研究进展进行综述和比较，并对未来的发展趋势进行分析。

关键词：虚拟化，云计算，嵌入式虚拟化，GPU 虚拟化，虚拟资源性能计量，虚拟资源调度与管理，虚拟化安全

Abstract

Virtualization technology, a technique to decouple hardware and software in computer system, permits the simultaneous run of multiple operating systems on a single physical machine, and improves the system resource utilization and manageability by such means as virtual machine migration and dynamic resource adjustment among different virtual machines. Along with the development of multi-core technology as well as the application of large-scale data center and cloud computing, there have emerged many challenges for the virtualization technology. Therefore, it is the hotspot of current research how to provide pertinent technical support for system virtualization based on the features of the new hardware and large-scale applications. Under the background of cloud computing, this report conducts a summary and comparative study at home and abroad in terms of the embedded virtualization technology, GPU virtualization technology, the measurement of virtual resource property on the cloud computing platform, the management and scheduling of cloud computing, and the security of virtualization, and further prospects their future and acquires the apocalypse respectively.

Keywords: virtualization, cloud computing, embedded virtualization, GPU virtualization, virtualized resource performance measurement, virtualized resource scheduling and management, virtualization security

1 序言

计算系统虚拟化是一种在硬件和软件之间去耦合的技术，可以支持在一台物理机上

同时运行多个操作系统，并通过虚拟机之间的动态资源调整以及虚拟机迁移等方法，提升系统资源利用率和可管理性。计算系统虚拟化是主流云计算平台的关键支撑技术。随着多核技术的发展，以及大规模数据中心和云计算应用的开展，对计算系统虚拟化带来了一些新的挑战。本报告以云计算作为背景，选取从云端到终端的 5 项计算系统虚拟化关键技术，探索支撑云计算的计算系统虚拟化新特征、新需求、新技术以及新的解决方案。通过对国内外近年来的研究进展，对目前国内研究现状与进展进行总结，并通过分析与对比，对未来可能的研究领域进行展望。

本文涉及的 5 项关键技术，包括：嵌入式虚拟化技术、GPU 虚拟化技术、云计算平台中虚拟资源性能计量方法、云计算资源调度和云计算平台中虚拟化安全。1) 随着移动操作系统的快速发展，嵌入式智能设备在人们的日常生活中发挥着越来越重要的作用，嵌入式虚拟化技术应运而生。嵌入式虚拟化技术经过近些年的发展，呈现出了各式各样的技术特征。由华中科技大学撰写的“嵌入式虚拟化技术的研究进展与趋势”（郑龙、廖小飞、邵志远）针对嵌入式虚拟化研究所面临的挑战，综述了近年来国内外嵌入式虚拟化技术的研究进展，基于分析与对比的结果，进而对其未来研究趋势进行了展望。2) 随着 GPU（Graphics Processing Unit，图形处理器）计算性能以及可编程性的逐步增强，具备大量运算单元、较快内存频率和较大内存位宽的特性使得 GPU 在通用计算领域大放异彩。在此背景下，GPU 虚拟化技术引起产业界及学术界的广泛关注，成为研究热点之一。相比于较为成熟的 CPU（Central Processing Unit，中央处理器）虚拟化领域，GPU 虚拟化面临着诸多困难与挑战。在 GPU 高速地升级换代且缺少相关文档的背景下，GPU 虚拟化技术需要在高复杂度的环境下设计出高性能的解决方案。由 Intel 亚太研发集团的董耀祖撰写的“GPU 虚拟化技术的研究进展与趋势”对国内外 GPU 的虚拟化方法进行了全面的介绍。3) 在云计算的 IaaS 平台上，如何精确的计量计算资源的性能成为关键问题，其目标是向用户提供量化的简明的性能数据，帮助用户全面掌握资源的性能，也能指导资源的合理分配。提供可计量的服务，也已经成为云计算的一个本质的特征。由北京大学、云南大学撰写的“云计算平台中虚拟资源性能计量方法的研究进展与趋势”（张彬彬、汪小林、罗英伟）对目前 IaaS 平台常用的资源性能计量方法进行了调研，讨论其中存在的问题，分析未来的发展趋势。4) 从云计算服务提供商的角度看，一个关键问题是解决负载波动和节能之间的矛盾：一方面，因为不良用户体验导致客户流失，直接关系到服务提供商的生存；他们尽量部署足够数量的服务器以应对负载的动态变化；而另一方面，这样做导致很多时候服务器都处于空闲状态，浪费电能。基于虚拟化技术的服务器整合（Server consolidation）技术提供了一种有效利用数据中心服务器和能源的方法。由北京大学撰写的“云计算资源调度的研究进展与趋势”（宋维佳、陈琪、肖臻）介绍了国内外对虚拟资源调度技术的研究现状和最新研究成果，并进行对比分析。5) 虚拟化技术给计算机系统的安全性带来了机遇与挑战。一方面，由于虚拟化在操作系统下增加额外的软件层，因此可以解决许多传统与新型的安全问题。另一方面，由于虚拟化软件栈的日趋完善与复杂化，其自身的安全性也变得非常重要。由上海交通大学、华中科技大学撰写的“云计算平台中虚拟化安全的研究进展与趋势”（陈海波、夏虞斌、

邹德清、戚正伟) 探索了虚拟化平台的相关安全威胁、需求与相应的解决方案。

2 嵌入式虚拟化技术的研究进展与趋势

2.1 引言

计算机系统虚拟化是一种利用软件实现对系统软件或者硬件的模拟，使得其他软件可以运行于模拟环境之上的技术^[1]。虚拟化技术带来了诸如系统整合、资源复用、计算优化、成本节约等优势，广泛应用于大型机、服务器和桌面机，很大程度上降低了企业界运营成本。虚拟化技术于 2006 年至 2010 年均成功入选国际著名研究暨顾问机构 Gartner 年度十大战略技术。随着小众智能移动生态系统的繁荣^① 及手机硬件条件的成熟^②，虚拟化技术正在步入一个新的发展空间：嵌入式设备（如图 1 所示）。

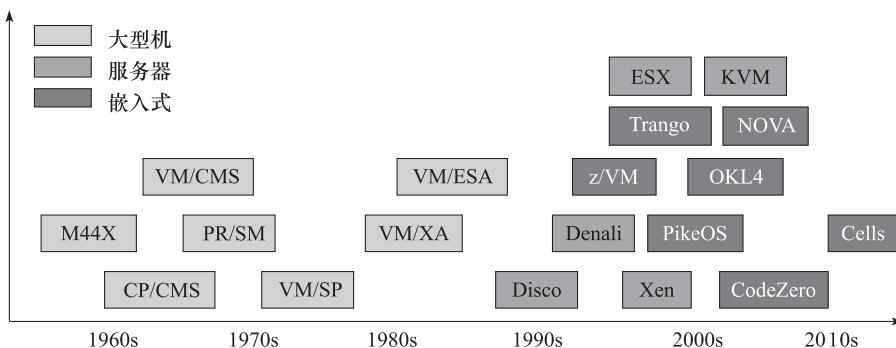


图 1 虚拟化技术发展历程

传统的虚拟化技术在物理机和操作系统之间构建虚拟机监控器（Virtual Machine Monitor, VMM）层，用于调度共享硬件，管理分配给每个操作系统资源^[2]。相较于传统的虚拟化技术，嵌入式虚拟化技术也有自己的 VMM 层，但是不同的是：传统虚拟化技术大都面向服务器和桌面应用，且对实时性要求不高；而嵌入式虚拟化技术则要求针对不同的应用进行特定地专有裁剪与设计，这是嵌入式系统设计的关键所在。嵌入式虚拟化技术是近几年才发展起来的一项远未成熟的新技术。在这一有着很强市场的新型领域，国内外学术界和企业界争先恐后地对其进行了大量的相关研究，但这些技术大都处于尝试阶段。本文将结合国内外嵌入式虚拟化技术的研究现状探索嵌入式虚拟化技术所面临的挑战，展示嵌入式虚拟化技术应如何延伸其触角，找到新的应用领域，进而对嵌入式虚

① Gartner 研究表明：2009 年智能手机出货量为 1.7 亿部，相较于去年增加了 28%；另外，移动设备入围 2013 年十大战略技术之首。此外，市场调研组织 Forward Concepts 预测，2014 年后智能手机的出货量将会以 24% 的速度增长。

② 例如，三星于 2013 年 3 月 15 日正式发布其八核智能手机 Galaxy S4。

拟化技术的未来发展趋势进行展望。

本节旨在让读者对嵌入式虚拟化技术的出现、现状和发展趋势有个全面、详细的了解和认知。

2.2 嵌入式虚拟化

2.2.1 研究背景

通用计算机系统的虚拟化在大型机、服务器和桌面虚拟化中已经取得了令人瞩目的成就。由于移动智能设备的高速发展，虚拟化技术下一个阶段的发展方向逐渐向嵌入式方向靠拢。著名嵌入式系统厂商 Wind River System 认为：“随着移动智能设备逐渐取代个人 PC 端，嵌入式虚拟化将会是未来计算的下一个方向”。究其原因，潜在的市场价值是其流行的根本原因，具体来说，有如下几方面因素：

1) 嵌入式硬件的高速发展。如何让计算能力强大的芯片充分发挥效用？做到这一点能够减少硬件设备的数量，降低设备投入。硬件更新换代或采用多核芯片后，如何减少产品开发成本，快速推向市场？多核平台将对软件开发带来重大冲击。这是因为，多数过去研发的面向单核的软件一方面无法充分发挥多核芯片的计算能力，另一方面可能因并发执行而产生故障。

2) 移动智能系统的多功能需求。以手机为例，除了通话，人们还期望利用手机上网、付款、收发电子邮件、看电视、理财、玩游戏、导航和移动办公等。手机已经成为电脑、钱包、电视机、游戏机和导航仪等工具的组合。为支持更多功能，一种应对办法是在单个设备上集成数个 DSP 处理器和通用处理器。这一方面增加了硬件成本；另一方面增加了功耗和重量；使得手机系统的应用前景受限。

3) 移动智能系统的高可靠性。如智能手机等嵌入式系统具有独立的操作系统，可以由用户自行安装软件。这带来两方面的威胁。一是设备越来越容易受到病毒和恶意代码的攻击。未来几年，手机所面临的安全威胁将超越个人电脑，成为个人信息安全的第一大隐患。二是设备故障的几率增加。安装的软件品质低劣或发生冲突，都会导致故障。

针对上述移动端问题及传统系统虚拟化技术的成熟，将虚拟化技术应用于嵌入式系统能有效解决或缓解这些问题。

2.2.2 概念及特征

嵌入式虚拟化系统是一类在嵌入式系统中部署虚拟机监控程序，并添加灵活性和一些高级功能的专用新型系统。目前，嵌入式虚拟化技术普遍采用半虚拟化技术^②，其主要原因是：硬件仿真或全虚拟化技术所带来的高额开销对嵌入式系统而言是不可接受的。

② 半虚拟化是指在嵌入式硬件平台和操作系统之间加入一层虚拟机管理器，由后者构造出嵌入式操作系统所需的运行环境，即虚拟机^[2]。

图 2 显示了嵌入式虚拟化系统和传统虚拟化系统的设计结构。嵌入式虚拟机监控程序由一个虚拟平台和一个微核组成；这种设计结构允许虚拟机监控程序同时支持多种不同的虚拟机（操作系统和应用程序）和单独的应用程序。

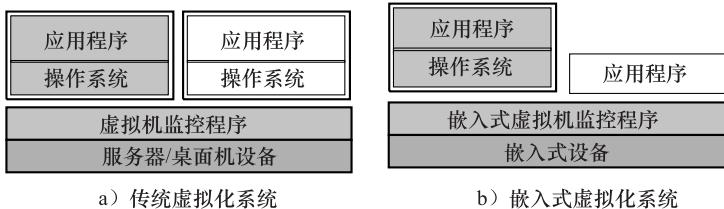


图 2 传统虚拟化系统和嵌入式虚拟化设计构架

如今的物理设备范围发生了转变：从拥有服务器级功能的强大处理器（例如，对虚拟化技术的硬件支撑）到拥有较少资源的功耗敏感性系统。这种多样性变化对于嵌入式虚拟机化技术而言是一个新的发展契机，同时也带来了比传统的系统虚拟系统要求更高、不同的、但受限的运行环境，其特征主要包括如下几方面：

1) 效率。无论是传统虚拟机监控器，还是嵌入式虚拟机监控器，都致力于提高其运行效率。但相对于传统虚拟机监控器而言，嵌入式虚拟机有着更加严格的资源限制，其规模尽量要小，且必须拥有极高的内存使用率。

2) 安全性及隔离性。由于嵌入式虚拟监控系统规模小，因此就相对容易验证和降低系统出错率。实际上，有些嵌入式虚拟机监控程序供应商已经正式验证了他们的虚拟机监控程序，并保证它们没有 bug。虚拟机监控程序越小，平台就越安全，这是因为虚拟机通常是系统唯一以特权模式运行的部分，充当了可信计算基^[3]（Trusted Computing Base, TCB）的角色。

3) 通信。嵌入式虚拟监控程序提供了多个 VM 和应用程序共享一个硬件平台的功能，但通常还扩展了一些通信机制供它们交互。这个通道既要有效又要安全，并允许特权和非特权应用程序并存。

4) 实时功能。嵌入式虚拟机监控系统必须支持实时调度功能，其主要原因之一是嵌入式虚拟化技术重视用户体验，其反应速度必须放在第一位。

2.2.3 技术挑战

嵌入式虚拟化技术带来了功能整合、成本控制、功耗管理、高可靠性等诸多好处，但是也面临着许多挑战，部分问题在传统虚拟化系统中不存在。

1) 如何有效地保证系统的实时性要求？运行于嵌入式系统内的应用大多属于实时应用。在宿主虚拟机无法占用足够 CPU 带宽的情形下，实时应用有可能无法满足实时要求，即实时任务有可能无法在规定时限内完成。为了解决诸如此类问题，修改实时应用或操作系统是无法解决问题的，那么虚拟机监控程序如何调度虚拟机将会是重中之重。

2) 如何很好地适应嵌入式平台的多样性？嵌入式系统种类多样，包括移动终端、医

疗设施、工业控制器、无线电通信设备、网络设备、机器人、车载电脑和数字家电等。嵌入式硬件平台的种数比通用计算机硬件平台的种数多出几个数量级。操作系统和应用软件也是多种多样，甚至没有操作系统。嵌入式虚拟化技术要得以广泛应用，就需能够很好的支撑多种多样的软硬件平台。

3) 如何有效管理电源? 用电池供电的嵌入式系统，电源的管理将会是一个很重要的问题。在没有运用虚拟化技术的嵌入式系统中，操作系统或应用软件对此采取了专门的节电措施。运用虚拟化技术后，客户操作系统依托的是虚拟机监控器所构造的虚拟机，独立的电源管理决策不再适用。例如，操作系统不能因其内部无工作负载就将整个硬件系统置于待机状态，因为其他虚拟机还有可能处于运行状态。虚拟机监控器如何有效地对电源进行全局管理将是一个不得不面临的问题。

4) 如何保证不同虚拟机间安全又有效的通信? 在移动终端中，不同虚拟机的进程之间需要协作的情况随处可见。例如，手机视频处理就可能涉及多个虚拟机进程：运行在处理基带业务的客户操作系统内的进程通过移动通信网络下载视频文件，而运行在处理常规应用的客户操作系统内的播放器完成视频播放；因此两者之间将进行大数据量传输。这种情况下，虚拟机之间可以建立网络连接进行通信，但却非常的低效。一个高效的办法是共享内存，不过这又与资源隔离的原则相冲突。如何找到一种安全又有效的通信机制是其能实用的必要条件之一。

2.3 国外研究现状

嵌入式虚拟化技术在国外率先兴起并展开了深入的研究，不少的开发产品已经投入商用多年。它们不仅来自于开源社区；而且还有一些专有版本。下面主要介绍国外对于嵌入式虚拟化系统的一些研究动态。

法国 TRANGO 公司^[4]主要是开发面向手机的实时监控管理器，并为 ARM 体系结构 (ARMv5 和 ARMv6) 和 MIPS 体系结构提供虚拟化方案，可运行于其上的系统包括 Symbian OS、WinCE、基于 Linux 的移动操作系统。TRANGO 虚拟化方案有着诸多优点：1) 它具有将私有或敏感代码和数据与受 Linux PLX 许可证限制的内容隔离的功能；2) 它具有将一个丰富的操作系统和一个以商业为主的应用和服务并行运行的能力；3) 用户可以通过其 ARMv6 CPU 核性能的技术提升将原先需要两个或更多的处理器才能完成的功能集中到一个处理器中；例如，以往需要一个额外处理器来实现基带或调制解调器的功能等情况，通过 TRANGO 提供的虚拟化方案，智能手机可以很好的将这些功能进行整合。由于 TRANGO 在嵌入式虚拟化技术研制方面的成功，VMware 于 2008 年 10 月将其收购，并入 VMware 手机虚拟化平台产线 (Mobile Virtualization Platform, MVP)，最终命名为 VMware Horizon Mobile。

德国 SYSGO 公司的 PikeOS^[5]实时操作系统 (Real-time Operating System, RTOS) 是一个能够让多个应用程序与 Android 和 Linux 等操作系统安全地在单一硬件上平行运行的

虚拟机监控程序。随着移动支付、跨装置间敏感资料串流、高价值媒体内容处理以及其他消费应用的兴起，微处理器和系统安全的重要性日渐提升。顺应这种趋势，PikeOS 可提供结合 RTOS 和安全虚拟机环境的独特方案。图 3 显示了 PikeOS 的系统构架，它主要实现了一个分离内核的设计思想。PikeOS 引入嵌入虚拟机监控程序允许在同一平台上同时使用较早的应用程序和更新的应用程序。PikeOS 在内核层实现了一套构架和平台支撑包，用于特殊的硬件环境（例如 x86、PowerPC、MIPS 等）以及提供虚拟化平台的分离内核，与虚拟机监控程序类似，分离内核安全地隔离上级客户操作系统。PikeOS 现阶段广泛应用航空航天电子工业，且逐步通过 AUTOSAR^[6]进入汽车产业。

Open Kernel Labs (OKL) 于 2006 年在美国芝加哥成立，主要致力于嵌入式微内核和嵌入式虚拟机监控程序的开发。该实验室在嵌入式领域中提出了 Microvisor，一个具有虚拟化功能的微内核。OKL 实验室最有名的产品莫过于 OKL4 Microvisor。OKL4 基于 L4[⊖]微内核，实现了安全单元的分区（用于对构架中的 VM 进行分区）。OKL4 Microvisor 拥有特权级，其他的程序（包括所有的 VM、应用程序、驱动程序等）放在隔离区中，且使用一种高效的进程间通信机制令各单位进行通信和合作，如图 4 所示。OKL4 内核及相关组件是轻量级的，因为所有的硬件设备驱动程序和单独的应用程序被放置于 Microvisor 之外。现如今，OKL4 Microvisor 已成功部署于 15 亿台设备中，其中大多数为手机平台。世界上第一部利用虚拟化技术在只有单一处理器上并发运行两个操作系统（Linux 和 BREW）的 Motorola Evoke QA4 手机正是基于 OKL4 Microvisor 上实现的。

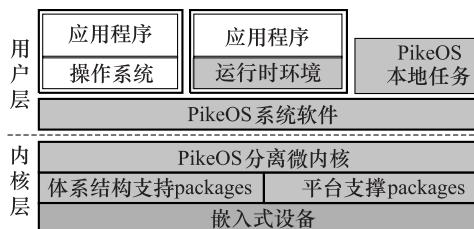


图 3 PikeOS 系统结构

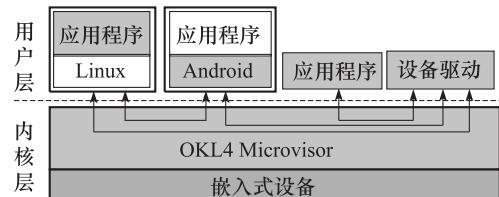


图 4 OKL4 Microvisor 系统结构

NOVA^[8]微虚拟机监控程序于 2010 年问世。相比于以前的微内核技术，NOVA 进一步实现了只有 9000LOC 的轻量级微核，包含一些用于非特权代码的独立层。NOVA 将系统结构分层三层（如图 5 所示）：一层微虚拟机监控程序；一层用户级环境；一层 gues-tos 层。这些核心组件包括：一个根分区管理器，用于资源分配、底层硬件设备的驱动程序和用来管理内存映射的虚拟机监控器。NOVA 基于全虚拟化技术实现，因此其未更改客户操作系统。相比于第一代微内核 Mach、Chorus 以及第二代微内核 L4，NOVA 被视为第三代微内核。

在已有的商用嵌入式虚拟化解决方案中，VirtualLogix 公司的 VirtualLogix VLX^[13]是其中的佼佼者。VirtualLogix VLX 是典型的嵌入式虚拟化技术，它实现了在硬件层和软件层

[⊖] L4 的灵感来自于 Mach^[7]，Mach 是美国卡内基 - 梅隆大学开发的内核，作为传统 UNIX 系统内核的一个简易替代品。

之间的一个抽象层（如图 6 所示）。如此这样，多个不同特点的客户系统产生不同的资源请求（如 CPU、real-time clock）时，可通过虚拟化硬件资源来满足它们的要求。VLX 最大的特点就是引入了分区的概念。共享存储器经过分区后分配给不同的客户操作系统，这样做好处在于各个客户操作系统可以运行自己的内存管理机制而不必干扰其他操作系统。这种机制也使得含有 MMU 的操作系统（如 Linux 和 WinCE）和不含 MMU 的操作系统（如 Nucleus 和 WRTx）可以在一个处理器上共存；同时该机制也增强了嵌入式系统的性能，提升了系统的安全性。

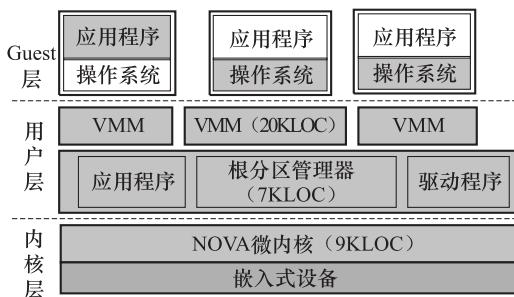


图 5 NOVA 微虚拟机监控器程序构架

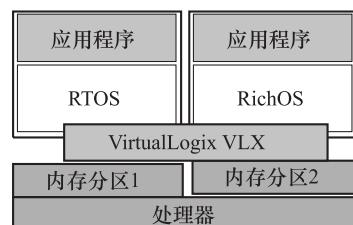


图 6 VirtualLogix VLX 系统构架

Wind River Hypervisor^[10]是 Wind River System 公司的多核软件解决方案的组成部分，它目前可以支持的操作系统有 VxWorks、Windows、Wind River Linux。类似于传统虚拟器，Wind River Hypervisor 基于硬件资源构造出多个虚拟机，其未来计划进一步扩大硬件的支持范围，包括 ARM 处理器系列等。

Secure Xen on ARM^[9]开源项目是由三星电子公司 2007 年 4 月发起，其主要针对 3G 及以上版本的手机网络中安全可信的实际需求，为 3G 及以上版本移动电话建立轻量级虚拟化技术。其主要涉及思想是结合 Xen 虚拟化平台提供的安全功能和 ARM 平台的高运行效率，开发出运行于 ARM 平台上的半虚拟化 Xen 嵌入式系统。现阶段，Xen 中的大部分核心组件如 xend、xenstore 及 xenconsole 均已能成功运行于 ARM 平台上。该项目尚处于研发过程中，其源码可以从 <http://wiki.xensource.com/xenwiki/XenARM> 下载。类似于开源的 Xen-ARM 项目，基于 KVM 的 KVM-ARM^[11]也致力于开源解决方案，但是这些方案虚拟机监控器性能较差，目前皆处于开发测试阶段。

Cell^[12]是美国哥伦比亚大学软件系统实验室的一个开源手机虚拟化项目，该项目的学术论文获得了第 23 届国际著名学术会议 SOSP 的最佳论文。Cells 基于 Android 实现了操作系统的虚拟化，在同一个手机可以同时运行多个 Android 系统，并可以自由切换，且性能非常理想，非常适合 Android 虚拟化需求。是目前最好的开源嵌入式手机虚拟化解决方案。在内核层，Cells 集成了手机多功能的需求，用单一的微核创建不同的 VP（Virtual Phones），并将它们隔离，且每个虚拟的 VP 都有自己的地址空间（如图 7 所示的 Cells 系统结构）。

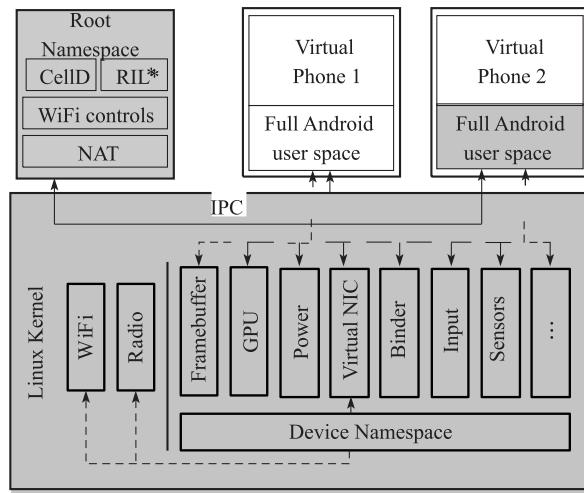


图 7 Cells 系统结构

2.4 国内研究现状

在嵌入式虚拟化领域，国内研究进展较慢，且大多基于 Xen-ARM/KVM-ARM 等开源项目，处于跟踪或探索阶段。针对嵌入式领域中的某些特定属性，例如，安全性^[15]、隔离性^[16]、鲁棒性等，国内也只是仅有一些零星的学术探讨。

浙江大学 SMART^[14] 嵌入式系统提出了子构件设想。SMART 体系结构融合了细粒度构件、弹性设计、任意剪裁等优点，具有代码高质量、高可靠性、高实时性、应用定制等特点。SMART 体系结构的软件模块由构件组成，而构件由子构件组成，通过构件拆分完成通用接口普适性与特定模块专用性的良好结合。用户可以根据自己的应用需求，选择不同的子构件构建出符合需求的嵌入式系统。但是该系统未充分利用虚拟化所带来的诸多好处。

文献[18]针对嵌入式虚拟执行环境中的实时性能特性，深入分析了 KVM 虚拟化执行环境中影响实时性能的原因，在此基础上测试了采用提升优先级和 CPU 专有核策略下的 KVM 实时性能，并提出了一种基于时间的可预测调度算法来改善实时客户虚拟机的实时性能。文献[19]针对当前手机领域面临的安全问题和现有手机安全解决方案的不足，深入剖析了其中的原因并提出了一种基于轻量级执行环境的嵌入式系统虚拟化模型；并在该模型的基础上实现了新的手机安全解决方案。而文献[20~22]则分别对嵌入式虚拟系统中的镜像存储、域间数据共享、及其他一些关键技术（如中断技术、设备驱动、I/O 读写）进行相关研究。总的来说，这些方法技术都只关注到了嵌入式虚拟化系统设计中某一方面或几方面的特征，更重要的一点是它们大多都是基于国外开源项目；因此，其本身离一个完整而自主的嵌入式虚拟化系统有着很长的距离。

为了填补国内在嵌入式虚拟化领域所处的空白，华中科技大学服务计算技术与系统暨集群与网格计算实验室开发了自主的嵌入式虚拟机管理器 VNeo 系统^[17]。VNeo 采用类似于传统虚拟化中的半虚拟化技术，基于三星著名开源项目 Xen-ARM 而开发。VNeo 作为一个嵌入式虚拟化系统，主要在能耗、性能、内存消耗和用户体验上做了重大改进，能够使得一个嵌入式设备能同时运行两个不同的客户操作系统，并支持它们之间的实时切换。目前，最新版的 VNeo 基于 GTK 图形界面能够同时运行 ARM Linux 和 Android 系统。为了满足嵌入式虚拟化的特殊需求，VNeo 做了如下几点调整：1) 将驱动程序转移到虚拟机管理器中，这样 VNeo 系统就能对硬件进行统一管理，最大限度地减少能耗；2) 去除了传统 Domain 0 虚拟机，大大地降低了 VNeo 对内存资源的消耗；3) 在虚拟机监控程序中引入 FAT32 文件系统，主要用于存储小文件和客户操作系统的镜像；4) VNeo 系统减少了客户操作系统启动时的超级调用次数和内存拷贝次数，大大减少了客户操作系统的启动时间。如图 8 所示，VNeo 系统主要包括：两个客户操作系统，一个控制系统，驱动程序和中断管理模块。VNeo 现阶段还只是处于模拟器测试阶段，下一阶段的 VNeo 系统将会支持更多的系统，进一步增强用户体验，并能够在真机上真实运行。

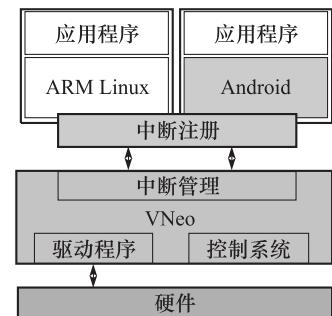


图 8 VNeo 整体构架图

2.5 国内外研究进展比较

从 20 世纪 60 代中期 IBM 实现世界上第一台虚拟化机器 M44/M44 X，历经半个世纪的发展，虚拟化技术逐渐成熟并发生了许多特征变化，为了顺应时代的发展，其中嵌入式虚拟化技术异军突起。为了实现嵌入式虚拟化所带来的诸多挑战，国外率先开展了大量的研究，不少虚拟化领域或嵌入式领域厂商纷纷推出自己的产品。随着市场的嵌入式虚拟化的需求越来越大，不仅有大批公司投身嵌入式虚拟化技术的研究中，而且参与其中的高校也越来越多，例如，日本的富士通公司、早稻田大学和筑波大学、澳大利亚的新南威尔士大学、美国的佛罗里达大学、芝加哥大学等等，都相继加入了这场令人憧憬的研究之中。相对于国外浓烈的研究气氛，国内相对冷淡，除了浙江大学等个别高校展开零星的研究外，几乎没有公司和专门的研究机构。要实现嵌入式虚拟化技术所带来的诸多挑战，国内研究人员或机构可谓任重道远，同时应提早开展研究并加强重视，并努力提升国家未来在嵌入式虚拟化领域中的战略技术储备。

表 1 列出了国内外嵌入式虚拟化技术特征对比图。从表中可以开出，开源项目支持的架构和系统相对较少、且实时性也不高，离商用产品有着很长的一段路；而国内研究基本上大体都基于 Xen-ARM 开源项目，因此距离国外嵌入式虚拟化技术先进水平还有着很长的一段路要走。

表 1 国内外嵌入式虚拟化技术特征对比

	PikeOS	MVP	OKL4	VLX	NOVA	Xen-ARM	Cells	VNeo
支撑平台	PowerPC/ ×86/ ARM/MIPS/ SPARC/ SuperH	ARM/MIPS	ARM/ ×86/ /MIPS	ARM/ ×86/ PowerPC/	×86 等	ARM32/ ARM64	×86	ARM
支撑系统	Linux/ Android/等	Symbian/ WinCE/Linux /Symbian/ Android	Symbian/ Linux/Android	Linux/DSPs/ Windows/ Android/ Symbian/ VxWorks	Symbian/ Linux/Android	Linux	Android	Android Linux
用途	航天航空、 汽车	移动设备	手机、消 费电子	手机、网 络设备	N/A	3G 智能 手机	智能手机	智能手机
虚拟方式	半虚拟化	半虚拟化、 硬件虚拟化	半虚拟化	半虚拟化、 硬件虚拟化	全虚拟化	半虚拟化	半虚拟化	半虚拟化
开源	否	否	部分	否	否	是	是	是
实时性	是	是	是	是	N/A	否	否	否

2.6 发展趋势与展望

正如本文上面描述的大量嵌入式虚拟机监控程序实例，其未来的系统架构和设计模式仍将会朝着微内核的方向发展。这种设计理念通常精简高效，有利于提高性能，减少代码，改善安全性和可靠性。具体来说，其未来发展有如下几方面：

1) 硬件级虚拟化。嵌入式系统对于性能一般要求较高，任何的性能损耗都是不可接受的，未来如果从硬件级支撑虚拟化，嵌入式虚拟化系统的性能将会大大提高。

2) 全虚拟化。尽管半虚拟化一个很普遍又有效的方式，但是其实现非常耗时。同时，调整操作系统来适应虚拟化技术是一个繁琐的任务，特别是在待虚拟化操作系统不可修改的情况下。NOVA^[8]即是属于全虚拟化的第三代嵌入式虚拟化监控器。

3) 易于管理。随着移动支付、跨装置间敏感资料串流、高价值媒体内容处理以及其他消费应用的兴起，未来嵌入式虚拟化技术将会越来越普遍，这同时也要求虚拟机的操作必须简单、直观、易于管理等。

4) 轻量级的虚拟机监控程序。嵌入式设备硬件资源通常非常有限，如何在有限的硬件资源上，高效地虚拟与分配同一硬件资源给不同的虚拟机使用将会是一个非常有挑战性的工作，这其中就包括基本的瘦内核或微内核设计理念，即尽量将虚拟机监控程序瘦身。

2.7 小结

随着移动市场的高速发展，嵌入式虚拟化技术应运而生。本节旨在让读者对嵌入式

虚拟化技术的出现、现状和发展趋势有个全面、详细的了解与认知。本节首先介绍了嵌入式虚拟化的基本概念和基本特征，进而围绕所面临的诸多挑战，综述了近年来国内外关于嵌入式虚拟化技术的研究现状。并在对比分析的基础上，展望了嵌入式虚拟化的未来发展方向及研究趋势。

3 GPU 虚拟化技术的研究进展与趋势

3.1 引言

随着 GPU (Graphics Processing Unit, 图形处理器) 计算性能以及可编程性的逐步增强，其具备大量运算单元、较快内存频率和较大内存位宽的特性使得 GPU 在通用计算领域大放异彩。在此背景下，图形虚拟化技术引起产业界及学术界的广泛关注，成为研究热点之一。图形虚拟化技术通常也被称作 GPU 虚拟化技术。相比于较为成熟的 CPU 虚拟化领域，GPU 虚拟化面临着诸多困难与挑战。在 GPU 高速地升级换代且缺少相关文档的背景下，图形虚拟化技术需要在高复杂度的环境下设计出高性能的解决方案。

对于图形虚拟化的解决方案，需要从多方面评价该方案的综合表现。这些指标包括：性能表现、保真性、安全性以及复用性^[23]。性能表现具体指在图形虚拟化环境下，虚拟机的图形性能表现。主要评价方式是在相同硬件配置下（尤其是 GPU 硬件），虚拟机环境相比非虚拟化的原生环境在图形性能表现上的性能损失。保真性关注虚拟化环境下，物理 GPU 的能力与各项特性是否能被虚拟机所发挥。安全性包含虚拟机之间的隔离性，虚拟机的可移植性等方面。复用性关注多个虚拟机是否能够共享一个物理 GPU 设备，这对于图形虚拟化的实用性非常重要。

根据当前的各类图形虚拟化解决方案，我们可以根据其虚拟化实现手段，将他们分类为软件图形虚拟化，硬件图形虚拟化，以及结合了软件层面与硬件层面，介于前两者之间的介入（mediated）图形虚拟化。

软件虚拟化技术针对图形栈的较高层级进行虚拟化，通过软件模拟的方式实现 GPU 虚拟化。软件虚拟化技术不需要依赖 GPU 架构的底层实现细节。在宿主机端，该技术通过 GPU 供应商提供的 API (Application Programming Interface, 应用程序编程接口) 和原生驱动程序访问物理 GPU，客户机与宿主机则通过软件方式交互。鉴于软件模拟的有限性，物理 GPU 的特性在软件虚拟化解决方案中往往无法全部实现。软件虚拟化的另一个挑战是如何在性能和虚拟环境下的安全性上进行平衡。

硬件虚拟化技术针对客户机的图形栈与物理 GPU 硬件的交互进行虚拟化。硬件虚拟化技术需要涉及图形栈底层，因此这项技术与 GPU 的架构模型相关。得益于与 GPU 硬件直接进行交互，硬件虚拟化技术在理论上将获得更好的性能表现并最大限度地支持 GPU

硬件提供的各类特性。然而，复用性是硬件虚拟化所面临的一项挑战。除此之外，硬件虚拟化依赖于具体 GPU 的相关文档，在实现上相比软件虚拟化更加复杂。

介入虚拟化技术介于软件虚拟化技术和硬件虚拟化之间。介入虚拟化同硬件虚拟化类似，需要涉及 GPU 架构模型底层实现，因此保证了其拥有较好的性能表现和保真度。同时，介入虚拟化通过软件模拟的方式实现 GPU 资源管理等事务，因此能够支持多台虚拟机对同一物理 GPU 硬件的复用性。类似于硬件虚拟化技术，介入虚拟化在实现上较为复杂，同时还需要考虑安全性、引起性能损失的额外开销等问题。

3.2 国外研究现状

国际上，产业界及学术界在图形虚拟化领域已经进行了近 10 年的各类研究，提出了多种基于不同实现技术的图形虚拟化方案。对于软件虚拟化技术，学术界提出了 VMGL^[24]、Blink^[25]、GVIM^[26] 和 VirtGL^[27] 等众多解决方案，实现了对 OpenGL、Direct3D、CUDA 等 API 予以硬件加速的图形虚拟化系统。在产业界中，Oracle（甲骨文）、VMware 和 Citrix（思杰）等公司下的虚拟化产品也提供了基于软件虚拟化的图形加速选项。本节将以学术界中开源的 VMGL 项目和产业界中 VMware 公司基于其虚拟化产品提出的 GPU 虚拟化方案为例进行介绍。

3.2.1 VMGL

VMGL 由多伦多大学的 H. Andrés Lagar-Cavilla 和卡内基 - 梅隆大学的 Niraj Tolia 等人于 2007 年首次提出。VMGL 提出一个跨平台、独立于 VMM（Virtual Machine Monitor，虚拟机监视器）和 GPU 的基于 OpenGL（Open Graphics Library，著名的跨编程语言、跨平台的图形编程接口）的虚拟化解决方案。虚拟机可以使用 GPU 硬件加速其 OpenGL 图形程序以获得更好的图形性能^[24]。

VMGL 实现在用户空间中，其中宿主机可以直接访问 GPU 硬件资源。客户机通过网络传输将 OpenGL 的 API 调用传送至宿主机，宿主机交由 GPU 执行后将结果返回客户机。

VMGL 中主要包括三个用户空间模块，分别是 VMGL library、VMGL stub 与 VMGL X Server Extension。在客户机中，VMGL library 代替原有的 OpenGL 库，在宿主机上创建 VMGL stub 以存储 OpenGL 命令。VMGL stub 与宿主机上的 OpenGL 库交互，代表客户机上的应用程序执行图形处理及渲染操作。在图 9 中，Viewer（观察程序）负责显示客户机的 2D 图形界面输出，捕获用户的输入并将之提交至客户机，在 VMGL 中一般采用 VNC（Virtual Network Computing，一种实现屏幕画面显示和远程操作的软件）实现。VMGL X Server Extension 用以协调 Viewer 的 2D 输出以及 VMGL stub 的 3D GL 输出，监视基于 OpenGL 的窗体的大小、位置等变化并将之提交至 VMGL stub，使得基于 OpenGL 的图形程序得以正常显示。总体来说，VMGL 对 OpenGL 的 API 进行虚拟化，通过 API 远程执行的方式实现了基于 OpenGL 程序的图形虚拟化。

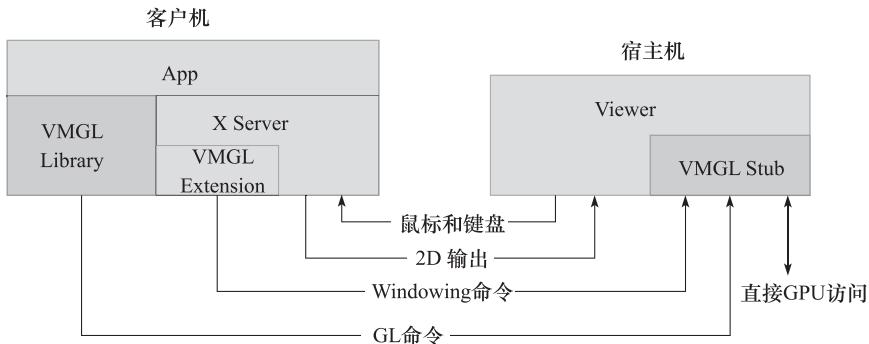


图 9 VMGL 的架构示意图

3.2.2 基于 VMware 虚拟化产品 GPU 虚拟化方案

VMware 公司的 Micah Dowty、Jeremy Sugerman 等人基于 VMware 的虚拟化产品提出一套提供 OpenGL 和 Direct3D（微软公司在 Windows 系列操作系统上提供的 3D 图形编程接口）加速的图形虚拟化解决方案。该方案主要通过在客户机端模拟 GPU 设备实现图形虚拟化，其结构如图 10 所示^[23]。

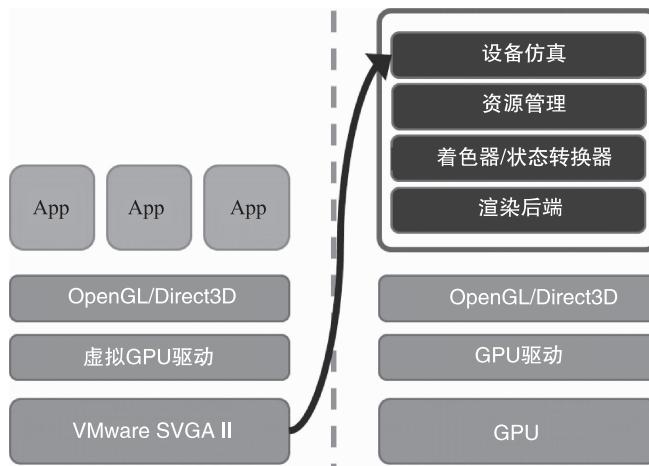


图 10 基于 VMware 系列产品的 GPU 虚拟化方案示意图

在此方案下，客户机拥有一块虚拟显卡，即模拟的 VMware SVGA II 显卡，客户虚拟机通过虚拟 GPU 驱动与该模拟显卡交互。虚拟显卡设备为客户机提供了三种虚拟硬件资源，包括寄存器、GMR（Guest Memory Regions，客户内存区域）和一个先入先出（FIFO，First In First Out）的命令队列。虚拟显卡上的寄存器用以模拟真实显卡的相关操作，并维护虚拟机的状态和图形渲染相关的数据等信息。GMR 抽象出供虚拟显卡使用的内存区域，包括一块虚拟显卡内存（VRAM），该内存实际由宿主机在物理内存上予以分配。命令队列的主要作用是作为传输层负责处理结构无关的 SVGA3D 渲染协议^[23]。

SVGA3D 协议是一套自定义的、针对 Direct3D API 的简化封装。客户机使用 SVGA3D

命令而非底层 GPU 命令，因而提升了整个系统的可移植性。客户机的虚拟驱动发送命令后，宿主机的后端程序周期性地读取这些命令，将之分类并交由宿主机的驱动程序进行图形渲染等处理。VMware 公司在宿主机上已经实现支持 Direct3D 和 OpenGL 的 3D 加速后端。整体来说，这套解决方案通过软件图形虚拟化方式，对虚拟机中运行的基于 Direct3D 和 OpenGL 的图形程序提供硬件加速。

3.2.3 基于 IOMMU 的相关技术

由于 GPU 硬件相关资料的公开程度较低，硬件虚拟化技术主要以产业界的研究为主。

IOMMU (Input/Output Memory Management Unit，输入输出内存管理单元) 可以将系统 I/O 设备可见的虚拟地址映射成物理地址，从而使 I/O 设备可以在虚拟环境中正常工作。SUN 公司、Intel 公司和 AMD 公司等都推出了支持 IOMMU 技术的硬件平台。对于 $\times 86$ 微处理器体系结构，AMD 公司提出的 AMD-Vi 技术以及 Intel 公司提出的 Intel VT-d 技术可以将 GPU 硬件设备分配给指定虚拟机使用，实现 I/O 操作的直通 (pass-through)，使得虚拟机能够完全利用物理 GPU 的性能进行图形渲染等操作，这项技术又被称为直通虚拟化技术。目前，众多主流的虚拟化平台已经支持 AMD-Vi 或 Intel VT-d 技术实现基于硬件虚拟化的图形虚拟化技术。然而上述直通虚拟化在实现 GPU 虚拟化的同时舍弃了复用性，因此物理 GPU 无法为多台虚拟机共享使用，导致这项技术的实用前景较为有限。

SR-IOV (Single-root I/O virtualization) 虚拟化是通过 I/O 设备硬件的改进，实现直通虚拟化和 I/O 共享。SR-IOV 虚拟化在硬件上将单一的物理设备表现为多个具有独立功能的 PCI 设备 (Virtual Function, VF)，而每一个 VF 都具有独立的与性能相关的 I/O 实例，通过硬件内部的调度将单一的物理设备表现为多个设备。但是由于图形设备硬件实现的复杂性，具有 SR-IOV 功能的图形设备目前并没有成为方向。

对于介入图形虚拟化方向，NVIDIA (英伟达) 公司于 2012 年宣布将提供支持硬件虚拟化的 GRID (旧称：VGX) 系列 GPU。下面将对 GRID 系列 GPU 在硬件层面支持的虚拟化技术进行简要介绍^[28]。

3.2.4 NVIDIA GRID 系列

NVIDIA 公司自 2012 年发表了支持硬件虚拟化的 GRID 系列 GPU 产品。针对不同的用户应用场景，基于 GRID 系列 GPU 的平台提供了 GPU 直通，基于软件实现的 GPU 共享，以及基于硬件实现的 GPU 虚拟化三种技术^[28]。

对于 GRID 系列提供的硬件 GPU 虚拟化技术，其结构如图 11 所示。

在 Hypervisor (虚拟机管理程序) 层，GRID VGX technology 模块在软件层面提供了对 GRID GPU 硬件虚拟化的支持。通过 Hypervisor 设置，用户可以根据需求为虚拟机配置 NVIDIA 公司提供的不同种类的虚拟 GPU，虚拟机可以通过 NVIDIA 驱动与虚拟 GPU 进行交互。虚拟机中的 NVIDIA 驱动通过 GRID VGX technology 和 Hypervisor 进行 GPU 内存映射、GPU 资源分配等配置。相关配置完成后，虚拟机中的 NVIDIA 驱动可以直接与物理的 GPU 硬件进行交互，处理虚拟机中的图形渲染。多台虚拟机可以通过 GRID VGX technology 共享一个物理 GPU 资源。目前关于 GRID VGX technology 的更多技术细节尚未公开，然而从图

11 中可以推测 GRID 系列的硬件 GPU 虚拟化技术可以被归类到介入图形虚拟化中。

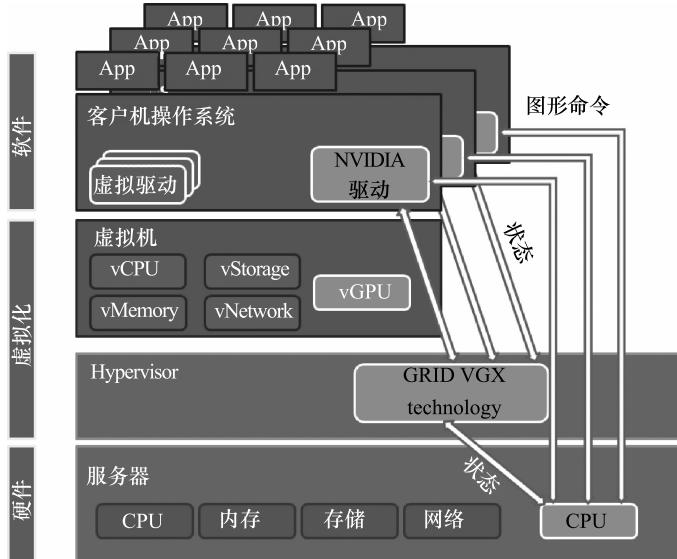


图 11 GRID 系列硬件 GPU 虚拟化技术示意图

NVIDIA 公司目前正与 Citrix 公司进行合作, Citrix XenServer 虚拟化平台将首先实现 NVIDIA GRID 系列的硬件虚拟化功能, 并预计于 2013 年中期发布相应产品。

3.3 国内研究现状

近年来, 国内学术界对于图形虚拟化领域也进行了相关研究。vCUDA 项目作为国内相关研究的典型代表, 关注图形虚拟化在高性能计算 (High Performance Computing) 领域中的发展前景。下面将对 vCUDA 项目进行介绍^[29]。

vCUDA 框架由湖南大学的 Lin Shi、Hao Chen 等人最早于 2009 年提出并于 2012 年发表相关改进, 针对虚拟化环境中的高性能计算领域提出了基于 NVIDIA 公司提供的并行计算架构 CUDA (Compute Unified Device Architecture) 的图形虚拟化解决方案。vCUDA 将客户机内的 CUDA API 拦截并转发给可以直接访问硬件 GPU 资源的宿主机, 交由宿主机的原生 CUDA 库及 GPU 驱动进行计算处理。vCUDA 框架针对客户机与宿主机之间的数据传输, 设计了专用的远程过程调用 (Remote Procedure Call, RPC) 基础结构优化传输性能^[29]。

如图 12 所示, vCUDA 基于客户端/服务器

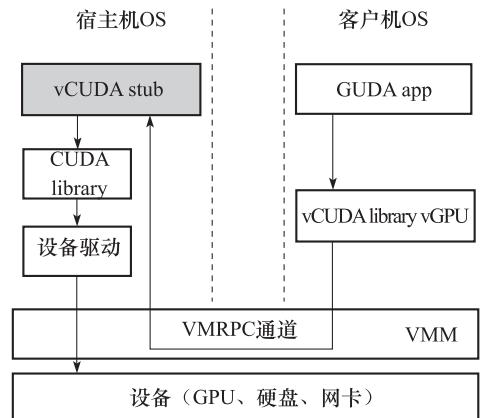


图 12 vCUDA 架构示意图

模型，包括三个用户空间下的模块，分别是：vCUDA library，vGPU 以及 vCUDA stub。vCUDA library 替代客户机操作系统中的原生 CUDA 运行时库，它负责将该客户机应用程序发出的 API 调用截获并转发至 vCUDA stub。当客户机上的应用程序执行 API 调用后，其相关信息及参数将被封装并置入一个全局的 API 队列，该队列维护 API 的相关状态，并依据不同的策略周期性地通过远程过程调用进行传输。vGPU 是位于内存中的一个数据结构，由 vCUDA library 创建。vGPU 对真实 GPU 的部分特性予以抽象，为客户机中的应用程序提供了一个虚拟的底层设备。vCUDA library 为每一个应用程序创建一个虚拟 GPU 的上下文以记录虚拟设备的各项属性和状态。vGPU 负责管理客户机的设备内存分配，并通知 vCUDA stub 相应地分配物理的设备内存，同时 vGPU 需要维护客户机本地内存地址与宿主机的远程内存地址之间的映射关系。vGPU 也负责记录当前的 CUDA 上下文环境以支持挂起和恢复操作。vCUDA stub 负责接受远程发送的请求并创建对应的执行上下文环境，最后将执行结果返回给客户机。vCUDA stub 同时管理物理的硬件资源，通过周期性地与 vGPU 同步以维护宿主机与客户机之间的状态一致性。

3.4 国内外研究进展比较

根据上文的相关介绍可以看出，国内对于图形虚拟化的相关研究虽然在数量有限，却处于国际主流研究领域前沿。如湖南大学 Lin Shi 等人提出的 vCUDA 项目较早地关注到图形虚拟化结合高性能计算领域，在 CUDA API 的图形虚拟化方面提出了一个较为完善的解决方案，为国际上其他关注图形虚拟化在高性能计算领域应用的研究提供了重要参考。然而，国内对于介入虚拟化和硬件虚拟化方面鲜有研究。其主要原因在于这两种实现方式依赖于 GPU 硬件模型，在学术界难以找到相关研究，而在国外，基于这两种实现方式的相关研究主要以产业界为主。

3.5 发展趋势与展望

对于软件虚拟化，支持 OpenGL、Direct3D 等图形 API 加速的图形虚拟化在众多主流虚拟化产品上已经基本实现。而在学术界方面，对于 OpenGL、Direct3D 加速的研究相对而言也比较成熟。随着人们越来越关注 GPU 在通用计算领域的发挥，图形虚拟化如何为 CUDA、OpenCL（Open Computing Language，开放计算语言，一个针对异构平台编程的框架）等涉及 GPU 通用计算的框架或标准提供更完善的支持成为当前软件虚拟化的研究热点之一。

在 x86 微处理器体系结构下，随着近年来各大 CPU 厂商提供 CPU 在硬件层面对虚拟化的支持，系统级虚拟化相关技术得到了长足的发展。尽管 GPU 在硬件层面提供对虚拟化的支持刚刚开始起步，对于硬件虚拟化，可以预见支持硬件虚拟化的 GPU 及其相关技术将成为未来的主流发展方向之一。

3.6 小结

相对于 CPU 虚拟化、内存虚拟化等较为成熟的虚拟化研究领域，图形虚拟化技术起步较晚。然而近年来，得益于虚拟化产品的潜在市场和广阔前景，加之 GPU 在图形渲染和通用计算领域的重要角色，图形虚拟化技术已成为虚拟化领域的热门技术话题。尽管当前 GPU 硬件的标准化程度远远不及 CPU，相关的接口或技术规范也鲜有公开，导致图形虚拟化的实现面临诸多困难。随着人们对 GPU 在虚拟化环境下的使用越发重视，图形虚拟化技术将更为产业界、学术界和虚拟化产品终端用户所关注。可以预见，图形虚拟化技术将拥有更为光明的发展前景。

4 云计算平台中虚拟资源性能计量方法的研究进展与趋势

4.1 引言

云计算的 IaaS (Infrastructure as a Service, 基础设施即服务) 模式成为商业应用以后，用户可以向 IaaS 平台按需租用具有一定级别计算能力的计算资源，资源提供方根据资源的性能和租用的时间计价。亚马逊弹性计算云 (Elastic Compute Cloud, EC2)：微软 Windows Azure、谷歌计算引擎 (Google Compute Engine)、Rackspace 云服务器等都是这样的业务模式。在这种业务模式中，资源提供方划分云计算平台的计算资源，并计量其性能，向用户呈现，供用户根据自身的计算需求按需选择租用一份性能适当的计算资源；用户认为所租用的计算资源实际表现出资源提供方宣称的性能。这样的业务模式是效用计算 (utility computing) 思想的一种实现，计算资源成为商品，计算资源的性能成为衡量商品质量的依据。资源的分配和计价方式、用户对资源的选择，都以资源的性能作为依据。

因此，在 IaaS 平台上，如何精确的计量计算资源的性能成为关键问题，其目标是向用户提供量化的简明的性能数据，帮助用户全面掌握资源的性能，也能指导资源的合理分配。提供可计量的服务，也已经成为云计算的一个本质的特征^[30]。本节对目前 IaaS 平台常用的资源性能计量方法进行了调研，讨论其中存在的问题，分析未来的发展趋势。

4.2 常用的云计算资源性能计量方法

现有的云计算 IaaS 的商业业务模式中，计算资源提供方一般将资源划分为不同配置的实例 (instance)，每一种实例具有一定的计算能力和存储容量，并具备一定的 I/O 性能，用户根据自己的计算需要租用其中一种或多种实例来部署自己的应用。从目前的实

现来看，一个实例就是一个独立运行的虚拟机，用户租用一个云计算实例，从技术上看就是租用了一个虚拟机的处理能力。云计算资源提供方需要对其所提供的虚拟机的性能进行计量，也就是在虚拟化环境中计量虚拟资源的性能，这与非虚拟化的物理平台上的性能计量存在不同。下面就虚拟机的 CPU、存储、网络等方面分别说明现有的性能计量方法：

(1) CPU 计算能力的计量

在物理计算环境中，CPU 的计算能力由多个方面的指标共同体现，例如主频、核心数量、核心架构、L1/L2/L3 Cache 容量、总线类型和频率等数据。IaaS 云计算平台基于虚拟化系统、CPU 资源的虚拟化技术，实际上是采用分时复用的方式使多个虚拟机共享物理 CPU 资源，现有的虚拟化方式，很难明确各虚拟 CPU 的各项性能评价参数。此外，IaaS 云计算平台用于商业应用时，为了向用户提供一个简单明了的计量结果，降低用户选择计算资源的复杂性，也需要简化 CPU 的性能评价指标。因此，在现有的 IaaS 平台上，CPU 的计算能力采用的是模糊的计量方式：通常人为定义一个计算单位 (computing unit)，以其计算能力作为标尺，计量虚拟机的 CPU 计算能力（虚拟 CPU 的性能）。目前业界还没有通用的虚拟 CPU 标准计算单位，各 IaaS 云计算平台对虚拟机的计量方法几乎都延续自亚马逊 EC2 服务所采用的计量模式，再制定自己的计量标准。亚马逊的 EC2 采用 EC2 计算单位 (EC2 Compute Unit, ECU) 计量云计算中虚拟 CPU 的计算能力，一个 ECU 被定义为相当于 2007 年生产的主频为 1.0 ~ 1.2GHz 的 Intel Xeon 处理器或 AMD Opteron 处理器（或 2006 年生产的主频为 1.7GHz 的 Intel Xeon 处理器）所具有的计算能力^[31]。类似地，微软的 Windows Azure 服务用单个“虚拟核”的计算能力（相当于 1.6GHz 的 Intel Xeon 处理器）作为计量单位，由虚拟机使用的虚拟核的数量来衡量不同的 CPU 计算能力，并提供了特小型（与其他虚拟机共享虚拟核）、小型（1 个虚拟核）、中型（2 个虚拟核）、大型（4 个虚拟核）、特大型（8 个虚拟核）等几种标准的虚拟机模型作为计费的单位^[32]。谷歌云计算引擎、Rackspace 云服务器、阿里云云服务器等也都使用“虚拟核”的计量单位计量虚拟机的 CPU 计算能力^[32~34]。

(2) 存储能力的计量

对存储能力的计量相对容易。在虚拟化中，内存和外存属于可被划分使用的资源，每一个虚拟机可独占使用一定数量的物理内存和外存空间。因此，在 IaaS 云计算平台中，可以为每一个虚拟机分配定量的内存和外存空间，在常见的 IaaS 商业平台中，与物理计算机的存储能力计量方式类似，仍然常用 MB 或 GB 计量内存容量，用 GB 或 TB 计量外存容量。

(3) 网络传输能力的计量

Rackspace 和阿里云都支持为每一个虚拟机分配定量的网络带宽，以 Mbps 计量^[34,47]。但在虚拟化环境中，由于多个虚拟机通常共享网络设备，存在竞争关系，如何保证虚拟机使用的网络带宽并充分利用网络资源是一个研究难点，Windows Azure 和谷歌云计算引擎仅统计网络数据流量作为费用结算依据，并未就网络带宽给出性能承诺，亚马逊 EC2 也未对虚拟资源的 I/O 带宽做出量化保证，仅粗略的给出高、中、低三个可选

的级别，只对同一个集群配置组内部的虚拟机之间的网络通信给出了 10Gbps 的带宽说明^[31]。

4.3 存在的问题

目前的云计算虚拟机性能计量方法主要存在以下两方面的问题：

(1) 粗糙的计量方式

目前除了存储能力采用定量的方式计量外，在 CPU、I/O 性能等方面的计量方法，还难以全面呈现虚拟机的性能。

虚拟 CPU 的计量方法以亚马逊 EC2 服务为代表，该典型的性能计量方式设定一个 ECU 或虚拟核的计算能力作为计量单位，将虚拟机的 CPU 计算能力规范化到该计量单位，以虚拟机提供多少个 ECU 或虚拟核来呈现虚拟机的 CPU 计算能力。但是，在确定作为计量标准的 ECU 或虚拟核时，却只给出了模糊的说明，例如，一个 ECU 被定义为相当于 2007 年生产的主频为 1.0 ~ 1.2GHz 的 Intel Xeon 处理器或 AMD Opteron 处理器，或 2006 年生产的主频为 1.7GHz 的 Intel Xeon 处理器所具有的计算能力，这个定义体现了主频、CPU 的生产厂商、所属系列和出厂时间这几个要点，并简单的将这三类 CPU 的计算能力进行了等价，作为虚拟 CPU 性能的计量单位。若使用该 ECU 的定义去衡量虚拟 CPU 的性能，将无法全面呈现虚拟 CPU 的计算能力，比如 CPU 的计算能力受到物理 CPU 架构方式的影响，例如，在 AMD Opteron 处理器中，每个核心具有自己的 L2 Cache，采用了集成的内存控制器，直接访问自己的内存；而在 Intel Xeon 处理器中，多核心共享 L2 Cache，通过共享的前端总线访问集中内存，两者的 Cache 容量和访问延迟等方面也存在差别，那么即使两个 CPU 的主频相同，它们在访存延迟和处理器间通信等方面都会表现出很大的性能差异^[35]。即使简单比较两者的定点和浮点计算性能，也存在差异。亚马逊 EC2 简单的将这两种处理器等价处理，无法准确呈现不同平台上虚拟机的实际性能和适合的应用类型。

虚拟机的 I/O 性能的计量则更加粗糙，部分 IaaS 平台仅支持为每一个虚拟机分配定量的网络带宽，EC2 不支持定量设定虚拟机的 I/O 带宽，而只提供了“高、中、低”三个可选的性能级别。Windows Azure 和谷歌云计算引擎则不支持设定虚拟机的 I/O 方面的性能。

目前的这种粗糙的性能计量方式，难以全面呈现云计算平台上虚拟机的性能，但是这样的计量方式，却提供了一种简化的性能计量结果，在同一个云计算平台上，用户易于把握不同虚拟机之间的相对性能差异，能快速对性能差异的程度做出一个粗略的判断。作为商业应用，如何能做到细致全面的计量并得到简单易用的结果，是一个研究的难点。

(2) 误差较大的计量结果

一方面，粗糙的计量方式决定了难以得到精确的计量结果，另一方面，现有的云计算平台几乎都采用虚拟化技术作为其底层架构，虚拟化的原理（运行在同一台物理主机上的虚拟机采取分时复用或划分使用的方式共享该主机上的 CPU、内存、外设等资源），

决定了虚拟机之间存在难以避免的对底层物理资源的竞争^[36]，特别是以分时复用方式共享的资源（CPU、网络带宽等），很难保证共享这些资源的虚拟机之间的性能隔离，如果在计量时忽略资源竞争造成的性能相互影响，计量结果将会出现较大的误差。在现有的对云计算平台虚拟资源的性能计量中，这种误差广泛存在。例如，在对亚马逊 EC2 的 ECU 性能评估中，当运行较长时间的真实的科学计算时，发现其实际性能达不到其标称的性能，其实际的浮点计算能力比理论值下降了 6~8 倍（选择 1.1GHz 的 AMD Opteron 处理器的浮点计算性能作为理论参考值）^[37~39]；而采用更适合部署在云计算平台上的面向服务的应用作为基准程序评估 EC2 的性能时，只有在单个虚拟机独自运行时，能够保持较稳定的性能，而当多个相同的云计算虚拟机同时运行相同的应用时，它们之间则呈现了很大的性能差别^[40]。文献[42]中在 75 个相同的 EC2 虚拟机上运行同样的 I/O 测试，结果出现了较大的差异，其标准方差达到了平均值的 16%。而文献[41]中采用了一组微观基准程序对 EC2 的 CPU、内存、I/O 的性能都进行了持续一个月的长期测试，测试结果同样表明，与物理环境相比，EC2 提供的虚拟机在 CPU、内存、I/O 各方面的性能都很不稳定。

计量误差影响了云计算平台资源的合理分配和使用：资源提供方无法得到准确的性能计量结果，那么要保证其承诺的计算能力，需要提供冗余的资源以保证突发负载时的计算性能；而用户发现资源难以达到标称的性能，为了满足待部署的应用的突发负载，一般会选择一个能够应付突发负载的虚拟机。这种使用方式实际上又回到了传统计算模式，在大部分时间，资源的利用率都不高。

虽然 EC2 等云服务平台支持用户根据负载情况调整所租用的云计算资源的计算能力，但是目前这种调整并不是在线进行，而需要停下原来运行的虚拟机，重新配置一个性能更好的虚拟机继续工作，也就是用户改租另一个虚拟机，这个过程造成的服务停止时间在几分钟的数量级。微软的 Windows Azure 启动一个虚拟机的时间长达 10 分钟^[46]。精确的性能计量能一定程度上避免运行时的资源调整带来的服务不可用时间。

4.4 国内外研究现状

国外典型的 IaaS 平台，包括亚马逊的 EC2、微软的 Windows Azure、谷歌云计算引擎、Rackspace 云服务器等采用类似的虚拟资源性能计量方法，其计量的细节、存在的问题、对计量效果的评测在上文 4.2 节、4.3 节中进行了说明。这些工作说明，IaaS 平台上的虚拟资源性能计量方法有待进一步研究。

目前已经有相关的工作深入探讨云计算中资源性能评测的方法：文献[43]梳理了已发表的研究工作中关于云计算的性能评测中采用的指标，并对这些指标进行了分类，给出了测量每一种指标通常所采用的基准测试程序。文献[44]进一步将分类好的指标组织成词典的形式，供后续研究查询。SPEC 组织了一个关于云计算评估的研究组，着重于云计算平台伸缩性、性能隔离、可靠性、可扩展性等相关指标的测量，但目前 SPEC 仍没有发布相关的基准程序^[45]。这些工作关注性能的评测方法和性能评价指标，虽然可以按照这

些工作中给出的评测方法给出虚拟资源在各性能评价指标上的评测数据，但是，全面罗列虚拟资源的性能参数并不适用于商业平台，大量细节的数据反而使得用户难以判断各虚拟机之间的性能特征。因此，如何简明地呈现量化的性能计量结果，也就是各性能参数如何以简单准确的形式呈现给用户也是一个待研究的重点。

国内 IaaS 云计算服务还不成熟，阿里云云服务器平台提供了可供用户租用的虚拟机，但其对虚拟机性能的计量方式仍很粗糙^[34]，腾讯公司正在开发腾讯云平台，拟提供 IaaS 服务^[48]，其他很多云服务平台仅提供存储服务。要提供高质量的 IaaS 服务，虚拟机的性能计量方法是一个关键，还存在很大的研究空间。

4.5 发展趋势与展望

如何准确全面的计量云计算平台上虚拟机的性能并简洁的呈现是亟待研究的方向，要达到这个目标，需要从以下方面开展研究：

- 1) 建立完善的虚拟资源性能评价指标，在文献[43~45]等工作的基础上，继续研究除了目前各云计算平台所采用的 CPU 主频、存储容量、I/O 带宽等性能指标之外，实际影响云计算平台上部署的常见服务运行性能的因素。
- 2) 建立一个（或一组）标准化性能计量模型，也即计量资源性能的标准单位，在全面评测虚拟资源性能的基础上，将评测结果规范化到标准计量模型，用标准单位量化性能计量结果。
- 3) 加强虚拟机之间的性能隔离，或者正确评估虚拟机之间的资源竞争对虚拟机性能的影响，避免计量结果出现较大误差，这是提高云计算平台服务质量的关键。

这三个方面的研究可以促进对资源性能的全面准确的计量。此外，为了更好的呈现云计算资源的性能特征，发挥云计算环境的优势，还可以从计量的角度呈现云计算的虚拟化平台不同于传统计算平台的特征：

1) 虚拟资源的弹性特征。云计算平台采用的虚拟化环境中，虚拟机和底层物理资源的松耦合表现为虚拟机和底层物理资源的映射关系可以被动态的调整，可以根据实际的计算需求动态增减虚拟机所占用的资源数量，这是主动的性能调整。此外，一个虚拟机虽然在很多方面表现为一个独立运行的计算环境，但是其底层的虚拟化实现使得一个虚拟机往往由于资源竞争等其他虚拟机的影响表现出很大的性能不确定性，这是被动的性能变化。如果在性能的计量结果中能够体现性能变化的程度和范围（弹性特征），可以帮助资源提供方合理分配资源，也使用户对虚拟资源的性能波动有正确的预期。

2) 多虚拟机的部署特征。在虚拟化环境中，多个虚拟机在底层物理资源上的部署方式可能极大地影响虚拟机的性能：CPU 物理架构会影响多个虚拟机的性能，例如，多核平台上部署多个虚拟机时，虚拟机与物理 CPU 核心的不同的对应关系会造成不同的性能表现，针对特定的缓存架构，用户可以利用核心之间的共享 L2 Cache 在虚拟机间共享常用数据，避免频繁的 Cache 失效；也可以利用核心和内存的直联结构避免多核对前端总线的竞争。网络的实现方式（包括同一台物理主机上虚拟机间 VLAN 的配置、物理主机

之间路由器和交换机的使用情况、物理主机的地理位置分布等) 都会影响虚拟机的网络性能, 现有的计量方法并没有体现出这方面的特征, 虽然 EC2 支持用户确定每一个虚拟机部署在哪一个物理位置上的哪一个 Zone 中^[31], 但用户仍然无法知道各虚拟机之间的网络拓扑关系。如果在计量结果中呈现多虚拟机的部署特征, 将有利于用户合理使用资源部署服务。

为了保持透明性, 现有的云计算服务向用户完全屏蔽了虚拟化特征, 包括虚拟机的性能的弹性、虚拟机在底层物理资源上的部署细节等, 使用户难以准确全面的评估所租用资源的性能。在性能计量时, 是否应该适度的呈现虚拟机的虚拟化特征, 还有待讨论和验证。

4.6 小结

本节总结了现有的 IaaS 云计算平台常用的虚拟机性能计量方法: 采用标准单位 (ECU、GCEU、虚拟核等) 计量虚拟机的 CPU 计算能力, 为虚拟机分配定量的内存、外存容量体现虚拟机的存储能力, 通过设定虚拟机使用的网络带宽衡量虚拟机的网络 I/O 性能。本节分析了这种性能计量方式存在的问题: 标准单位的定义粗糙, 难以准确呈现虚拟机的实际性能; 虚拟机之间存在难以避免的资源竞争, 使得度量结果存在较大的误差。这些问题, 影响了云计算平台的资源分配和用户的使用感受。如何准确的计量虚拟机的性能, 并呈现为简明的计量结果, 是提高 IaaS 云计算平台服务质量的关键。本节指出, 建立完善的虚拟资源性能评价指标和标准化性能计量模型是优化性能计量方法的基础, 提高虚拟机之间的性能隔离是改善计量结果准确性的关键, 此外, 还要进一步探索能否在度量结果中适度呈现虚拟机的弹性特征和多虚拟机的部署特征。

5 云计算资源调度的研究进展与趋势

5.1 引言

云计算^[49]模型把计算看成是等同于水电的资源。计算用户只需要购买在线的计算服务而不必建设机房购买服务器, 因此节省了投资、降低了成本。其发展前景得到了工业界的广泛认同。从云计算服务提供商的角度看, 这涉及建立和维护大型数据中心, 其中有很多问题需要解决。一个关键问题是负载波动和节能之间的矛盾。一方面, 因为不良用户体验导致客户流失, 直接关系到服务提供商的生存, 他们尽量部署足够数量的服务器以应对负载的动态变化, 而这样做导致很多时候服务器都处于空闲状态, 浪费电能; 有调查表明, 数据中心的大多数服务器利用率只有 5% ~ 15%^[50]。另一方面, 大型数据中心的服务器数量庞大, 电能成本惊人: 根据 Amazon EC2 在加州和弗州两个地区提供的

虚拟机价格和这两个州的电价之间的差异可粗略推算，即使在电价较低的弗州，电力成本占总成本 15% 以上^[51]；据统计 2006 年，全美国的数据中心能耗已达 610 亿千瓦，是 2000 年能耗的 2 倍^[52]，占全美总电能消耗的 1.5%。

除物理改造现有数据中心和在新建时采取节能方法外，由虚拟化催生的服务器整合（Server consolidation）技术提供了另一种有效利用数据中心服务器和能源的方法。多个服务（一般是高峰期互相错开的）被包装在各自的虚拟机中，共享一台物理服务器以提高资源的利用率。这涉及一系列问题需要解决。例如，把哪些服务聚合在一起才能更有效地发挥服务器的能力？在物理机上，各服务竞争资源时，应该用何种手段，以什么样的策略进行分配？虚拟机热迁移^[53]提供了动态调度虚拟机（服务）的手段，那么应该采用什么迁移策略折中保证服务质量和服务的平衡？我们统称解决上述问题的种种方法为虚拟资源调度技术。本文将介绍国内外对虚拟资源调度技术的研究现状和最新研究成果，并进行对比分析。

5.2 国外研究现状

国际上针对云计算资源调度的研究可归类到如下两个层次：同一物理机上的虚拟机之间（本机资源调度）和不同物理机之间（全局资源调度）。本机资源调度涉及 CPU、内存和 I/O 在虚拟机之间的优化分配；全局资源调度则涉及负载如何在物理机之间优化组合。

当多个虚拟机在一台服务器上运行时，它们竞争着服务器的资源。虚拟机管理器（Virtual Machine Monitor, VMM）或者 hypervisor 负责这些资源的调度。一般来说，虚拟机管理器提供调度机制，而由上层软件（如 VMware Distributed Resource Scheduling 软件）来确定调度的策略。根据资源类型的不同，我们将本地调度机制分为 CPU 调度、内存调度和 I/O 调度。

5.2.1 本地资源调度

本质上，VMM 在各个虚拟机之间调度 CPU 与操作系统在进程之间调度 CPU 类似。不同的是，虚拟机管理器调度的单元是虚拟 CPU（VCPU）。每个虚拟机有一个或多个 VCPU，由虚拟机管理器决定何时把 VCPU 调度到真实的 CPU（我们不用“物理 CPU”这个词，因为它可能是多核 CPU 上的某个核，或者是某个 hyper-threading 逻辑 CPU）上运行。其中有诸多复杂的因素需要考虑。

根据应用的权重为虚拟机按比例分配 CPU 资源是常用的 VCPU 调度方法。Xen2.0 采用 Borrowed Virtual Time 调度策略、Simple Earliest Deadline First^[54] 和 Xen3.0 之后缺省的 Credit^[54] 调度算法都属于这一类。它们都是基于 round-robin 扩展而来，只是每个虚拟机具有不同的权重，获得的时间片数量不同。

同一虚拟机的多个 VCPU 之间有协同调度（co-scheduling 或 gang-scheduling）的问题。假设两个运行在不同 VCPU 上的线程 A 和 B 需要进行同步，例如可能出现线程 A 等

待 B 的执行结果。如果线程 B 所在的 VCPU 没有获得执行，线程 A 所在的 VCPU 也将得不到执行。虚拟机管理器不知道这种依赖关系，无法确定应该首先调度哪个 VCPU。为了缓解这个问题，ESX Server 3.0 以前采用严格协同的调度策略（strict co-scheduling），即要求虚拟机的多个 VCPU 必须同时运行；但这样做带来两方面的问题：VCPU 较多的虚拟机很难得到运行的机会，且虚拟机的 VCPU 数量不能超过系统中实际虚拟机的数量。后来的 ESX Server 跟踪每个 VCPU 未得到运行的相对时间，在调度中尽量保持虚拟机的每个 VCPU 步调一致，从而放宽了该限制。理想的协同调度允许虚拟机管理器获得操作系统进程/线程的同步依赖关系，从而正确决策应该让哪个 VCPU 得到执行^[54]。这种做法的缺点是需要对操作系统进行改造。

文献[55]针对多层次（multi-tier）的 Internet 应用提出了一种本地 CPU 调度策略。其核心思想是利用应用性能（如吞吐量、响应时间、损失率）和调控目标的差距来调整虚拟机的 VCPU 的调度权重。这个负反馈回路起到稳定应用性能的作用。实验表明，与现有主流调度算法相比，每个虚拟机的 CPU 利用率都得到了较好的控制，尤其是在系统饱和的时候，不同应用的 QoS 得到公平的保证。

文献[56]在此基础上进行了两方面改进。首先，通过引入磁盘 I/O 调度^[57]，实现了对多维资源的控制。另外，与文献[55]中单纯使用负反馈进行资源调整的方法不同，文献[56]建立了应用性能的二阶模型：认为应用下一个调度周期的性能表现，与前两个调度周期的性能表现，以及本调度周期的资源分配和上一个调度周期的资源分配这四个因素呈正比关系；在每一轮调度中，历史性能数据被用于回归分析以更新模型参数，用以预测下一个调度周期中应用的性能。因为这个模型没有考虑各资源之间的互相制约关系，所以确定资源需求的时候，各资源调整的幅度以平方和的形式计入代价，防止资源分配的突变。当资源紧张出现竞争时，文献[56]以所有应用性能下降的平方和最小为目标寻找合理的本地资源分配方案。文献[56]与[55]有类似的不足：考虑的资源维度依然有限，内存和网络难以考虑进去；除 Internet 应用之外的其他应用，例如 Map Reduce 和高性能计算或者虚拟个人桌面机等不容易用文中所述的应用性能模型进行建模；另外，该模型对负载的突变处理能力有限。

不同于其他资源，虚拟机对内存资源的占有是持续的，而且内存在虚拟机之间的共享代价较高，不仅因为页表的改写导致 TLB 失效；在一个页面交给另一个虚拟机之前，出于安全的考虑，需要将整页数据重刷。Ballooning^[58]通过虚拟机内部的一个“气球”进程，给 VMM 提供一种从虚拟机中搜集空闲内存的手段。当 VMM 需要从虚拟机中挤出内存时，便通知气球进程（balloon）执行 inflate 操作：气球进程向虚拟机操作系统申请内存页面并将它们 pin 住以防被回收；接着，气球进程将这些页面交给 VMM 使用。VMM 将内存还给虚拟机时，它通知气球进程执行 deflate 操作：气球进程便把指定数量的内存页面释放给虚拟机操作系统。

Transcendent Memory^[59]在 hypervisor 中维护了一个公共的内存池，提供了两个操作 tmem_get 和 tmem_put，分别用于整页数据的写入和读出；并修改了操作系统的 cache eviction 和 swap 操作，分别称为 pre-cache 和 pre-swap。公共内存池可提供两种逻辑内存

池：持续性内存池保证写入的数据不会丢失，但可能因为内存不足写入失败；非持续性的内存池则不能保证内存不会丢失，但能保证内存总可以写入。所谓 pre-cache，是指虚拟机操作系统在系统内存紧张需要回收 page cache 的时候，先调用 tmem_put 操作将内存写入非持续性的内存池，再回收该内存页面；下次再需要此页数据时，操作系统首先从非持续性内存池中用 tmem_get 查询，是否它是否还存在，若存在，就读取之以节省读磁盘的开销，否则读取磁盘。pre-swap 则是在系统内存不够，需要换出一些页面时，操作系统首先将尝试将页面写入到持续性的内存池，如果写入成功，就不必把页面再写入磁盘以节省写磁盘开销，否则直接写入磁盘。这种将虚拟机操作系统的 page cache 集中起来统一管理的方式，大大提高了内存的使用效率，使每个虚拟机的内存分配可以被压缩到最小。当虚拟机的内存不够使用，pre-swap 操作则利用公共内存池存放换出的页面，避免高昂的 swap 操作。Transcendent memory 在提高内存使用效率的同时，还规避了预测工作集的困难。Transcendent Memory 目前的缺点是不支持全虚拟化。

I/O 调度和 I/O 资源的种类、I/O 虚拟化技术密切相关。在云计算环境中最主要的 I/O 资源类型是磁盘和网络 I/O。而 I/O 虚拟化的方式有纯软件虚拟化、IOMMU 辅助的软件虚拟化和硬件虚拟化。

Xen hypervisor^[60]采用称为“Split Driver Model”的纯软件 I/O 虚拟化方式为客户机 OS 虚拟了通用的网络、磁盘等设备。但这种调度方式本身难以保证各个虚拟机之间的公平性；也不能隔离各个虚拟机之间的 I/O 操作；各个虚拟机的磁盘 I/O 聚合起来处理，能减少磁头的移动时间，First-Come First-Served 调度无法挖掘这种优化潜能。文献[61]提出一种基于 Round Robin 的磁盘调度算法，能确保 I/O 时间（而不是 I/O 请求数量）在各个虚拟机之间的公平分配，并且保证一个虚拟机的 I/O 负载不会影响另一个虚拟机。这个调度算法虽然达到了公平和隔离性的目的，但是对 burst I/O 的情况会丢失性能^[62]。文献[62]指出，虚拟机上客户机 OS 内部的磁盘调度都是基于这样的假设：磁盘是磁头 - 转盘的机电系统或者新型的 SSD 磁盘，但实际上虚拟磁盘表现出来的特性与这个假设完全不同；因此，改善磁盘 I/O 调度的性能需要修改客户机 OS 内部的磁盘调度算法，同时需要 VMM 的磁盘 I/O 调度算法相配合。

5.2.2 全局资源调度

当本地资源的绝对数量不能满足负载需求时，必须通过其他手段来进行调度。一种办法是根据虚拟机中应用的负载事先周密规划，确定虚拟机在物理机上的搭配，确保在较长时间内不会出现过载的情况；这种方法叫做静态服务器整合。另一种办法是，在出现本地资源不够时利用虚拟机热迁移技术把虚拟机迁移到负载较低的物理机上面去以均衡负载。对于常见的 Web 应用类型，还可以将负载重定向到较空闲的物理机上去。

根据服务各自对资源需求的特点，结合物理服务器的性能，寻找相对固定的搭配组合方式将服务汇集到物理服务器上运行。这种整合方式为静态整合。静态整合的好处是实施简单，但寻找优化的固定组合方案很困难。一方面，负载的预测困难；另一方面，即使可以确定服务的负载，在此基础上找到最优方案也是个 NP 完全问题^[63]。实用可行

的方案需要为服务预留足够资源以应对剧烈波动的情况；在寻找优化的搭配组合时也往往采取近似算法。另外，在实际的部署中，服务可能是多层次的（Multi-tier），每个层次（tier）可能被独立封装在各自的虚拟机中，成为静态整合的一个处理单元。文献[63]对静态整合进行了详细的描述。该方案用 Yield 概念^[64]来描述每个服务对资源的需求，规避了对预测负载的需求；然后实验对比了精确算法、装箱算法和遗传算法。Stillwell 的试验结果表明，在实用算法中，矢量装箱算法的成功率、最优近似度和计算时间是最优的；在众多矢量装箱算法中，文献[65]提出的算法结果最接近最优上限。在计算量可接受的范围内，基因算法比最简单装箱算法（文献[63]将这类算法提出来归为贪心算法）的成功率和优化度更好，但是其优化度不及矢量装箱算法。

虚拟机热迁移^[53,66]能够在虚拟机运行过程中将它透明地迁移到另一台服务器上运行。这种特性方便于虚拟机的资源配置的动态调整：当某个物理服务器上太繁忙时，可以将其上部分虚拟机迁移到别处，使其不致影响应用的服务质量；当一些物理服务器空闲时，又可以把虚拟机都聚集起来，释放空闲服务器以节省能源。

文献[67]设计了一个利用虚拟机热迁移技术实现负载均衡的虚拟机调度器 Sandpiper。Sandpiper 由一个集中控制器和在各个物理机上的监控器组成。监控器定期将各虚拟机的 CPU、内存和网络 I/O 使用统计数据发送给集中控制器（若有可能，从虚拟机内部看到的内存、网络 I/O 甚至应用的性能统计也发送给集中控制器；这种方法涉及虚拟机的内部修改，因而称为灰盒策略）。集中控制器根据各个物理机和虚拟机的统计数据判断哪里发生了资源短缺，然后用启发式算法计算出调节方案：一组虚拟机迁移计划，以此控制各物理机上的监控器实施调度。

Sandpiper 考虑了三种资源类型：CPU、内存和网络，并用一个取值为三类资源空闲率乘积倒数的 volume 指标来表示物理机或虚拟机的繁忙程度。这种做法的好处是将多个维度的资源合一，使调度算法简洁易操作。Sandpiper 的核心算法针对每一个热点（发生资源短缺的物理机），为 volume 与其内存镜像大小比值最大的虚拟机寻找一个 volume 最小的物理机作为迁移目标；直到所有热点被消除为止。但是 volume 指标的缺点是不能表示到底是哪个资源紧张，因此无法有效地把嗜不同资源的虚拟机搭配起来有效地发挥服务器性能。

VirtualPower^[68]的目标是绿色计算。它在虚拟机管理器层为虚拟机提供了软件 ACPI 和 DVFS 的接口，一方面可以支持虚拟机透明地进行节能操作，另一方面利用虚拟机的节能操作，通过 CPU 调度、真实的硬件降频和基于虚拟机迁移的动态服务器整合三种手段来降低功耗。从该研究的结果来看利用硬件 ACPI 和 DVFS 最多只能节省 8% 的电能，而结合动态服务器整合腾出空闲物理主机节能的方式能够节省约 34% 的电能。文献[68]是我们知道的第一个用虚拟机迁移技术进行绿色计算的工作，但总的说来，文献[68]中的方法不是一个真正的调度算法，它没有考虑如何提高系统性能的问题。

虚拟机和运行它们的物理服务器可以用矢量装箱算法进行描述^[63,67,69]。但是用标准装箱算法解决虚拟机动态调整的问题，需要解决如下几个困难：与静态整合的装箱模型不同，物品（虚拟机）已经在箱子（物理服务器）中，不能使用离线的装箱算法来处理

而只能使用在线（或者半在线）的装箱算法；另外，装箱问题不考虑放入、取出物品的代价，而迁移的代价不能忽略；由于负载的变化，虚拟机的需求也是变化的，也即物品的大小是动态变化的。

将负载（如 HTTP 请求）通过应用层交换机（L7 switch）重定向到对应某个应用服务器的方法可以导向负载以控制资源在应用之间的分配。这种模式在虚拟化云计算兴起之前就已经存在，其相关研究成果兼具借鉴和实用价值。

MSRP 调度算法^[70]适用于无状态的 Internet 应用。它允许每个应用提供自己的估价函数（utility function）来描述其期待的服务标准，如为每分钟 1000 个以内的 HTTP 请求付 1 分钱/请求，而为每分钟 1000 ~ 2000 之间的 HTTP 请求付 0.5 分钱/请求；同时，为处理这些请求而运行每个物理服务器是有开销的，譬如 0.1 元/分钟；它再根据每个应用不同的出价和运行它们的成本来计算资源的配比方案以最大化自己的收益。

MSRP^[70]的优点是可以描述动态的复杂的 SLA，而且同时考虑了负载均衡和绿色计算。但是文献[70]假设所有应用可以在物理服务器上同时运行，这在 2001 年的数据中心中可能是现实的，但是在应用越来越庞大的今天，就不适用了。另外，MSRP 仅考虑了 CPU 资源的调度。

当今的 Web 应用较大，无法在同一物理机上启动较多的应用；而且庞大复杂的应用启停时间较长。文献[69]针对这种场景建立了调度模型。他们认为最要紧的资源是服务器的 CPU 和内存：CPU 资源随着负载的变化而变化，是负载相关资源；而应用占用内存的大小几乎恒定不随负载而变化，是负载无关资源。把应用的负载看做流，把 CPU 的资源看成边的容量，可以把当前应用布局（placement）转化为一张流图，并用最大流算法求出其够承受的最大负载。文献[71]提出通过定期地检查流图，确定应用负载是否被消化；如果没有消化，则尝试根据当前负载和布局重新运行最大流算法看是否能在不改变布局的情况下通过调整负载来满足应用需求；如果尝试失败，再考虑用调整应用布局的方法来解决。

文献[69]和文献[71]针对当今的 Web 应用复杂庞大和对内存的需求较大的特点，提出了合理的资源调度算法，但是缺乏对绿色计算的考虑。并且，与其他大多数基于负载调度的应用一样，只能处理应用中无状态的层次，不能解决后端存储的性能瓶颈。最后，文献[69]和文献[71]不支持应用层的 SLA，SLA 在进入算法之前已经转换成了 CPU 指令数。

即时消息应用、多人在线网络游戏、流媒体服务等 Internet 应用需要客户端和服务器之间保持一个较长时间的 TCP 连接。这一类型应用的结构与 Web 应用的结构类似，不同的是 Web 服务器的角色由一组与用户保持长期 TCP 连接的连接服务器（Connection Server, CS）代替。在这类应用场景下，资源调度过程中需要考虑两个限制：一方面，CS 处理登录请求的速率是有限的，因此 CS 启动后负载上升需要时间。另一方面，因为无法预知用户在何时断开连接，调度算法也不能做到迅速释放 CS，否则将引起服务器主动断开连接（Server Initiated Disconnection, SID），服务质量下降。

文献[72]在 Windows Live Messenger 的基础上设计了连接请求的调度策略，辅以开机

关机以保持活动 CS 的数量与负载相匹配，实现了负载均衡和绿色计算。为了保证服务质量，它定义了 CS 支持的最大连接数 N_{\max} 和最大登录处理速率 L_{\max} ；在这两个限制条件下，新进入的登录请求按照公平的策略重定向到各 CS，使负载被满足的同时在活动 CS 上均匀集中。CS 被尽可能分别地向两种状态集中：全速运行或空闲。关闭近似空闲的服务器不会引起大量的 SID，打开新的服务器也考虑了负载上升的时间。

文献[72]的工作基于 Windows Live Messenger，由于负载的周期性和稳定性非常好，它采用的以最近和前几周相同时间点的实测负载为参数的线性预测模型就有很好的预测效果。在游戏服务器由于新推出一个人物、副本会引起用户的大量增加；视频流媒体网站的用户量可能也会随突发事件的影响而起伏。这些因素导致该工作在适用性方面受到限制。

5.3 国内研究现状

国内对云计算资源调度问题也有深入的研究。针对基于虚拟机迁移的云计算资源调度问题，文献[73]提出了一种半在线的装箱算法解决了传统装箱算法无法支持物品变化而且迁移数量过多的问题。文献[73]将物品按大小分成 T、S、L、B 四个类型（为了简单起见，后续表述用字母代表该类型的物品），其大小范围分别处于 $(0, 1/3]$ 、 $(1/3, 1/2]$ 、 $(1/2, 2/3]$ 、 $(2/3, 1]$ 之中。为了有效利用空间，S 和 S 放一起，B 单独放，L 和 T 放一起，T 和 T 或 L 放一起；并且在处理新物品时确保最多存在一个只装了 S 的箱子；最多存在一个剩余空间大于 $1/3$ 的、只装着 T 物品的箱子。但是，原算法不涉及物品大小变化。为了解决物品变化时的处理，依次考察大小发生变化的所有物品，及时调整所在箱子的物品，以保证原算法的限制条件得以满足，保证理论的最优近似度。举例说明：某物品从 L 变成了 B，那么判断该物品所在箱子是否有 T 物品，如果有，把这些 T 物品取出作为新物品放入；再如，某物品从 B 变成了 S，把 S 从这个箱子中取出作为新物品放入。

实验表明，算法 [73] 性能相对于其他算法更好，并能满足实际负载的稳定性需求。以通常的虚拟机/物理机 10: 1 的配比，在使用负载预测的条件下，平均每个物理机每调度周期将发生约 0.4 次迁移（包括迁入和迁出）；再考虑到系统中出现负载波动时产生的调整远远大于负载较稳定的时候，物理机出现的瞬时峰值远远高于这个数量，引起较大的瞬时网络负载。

装箱算法兼具负载均衡和绿色计算的双重效果。它监测负载发生变化的虚拟机，检查所在的箱子的性质是否发生变化。若性质发生变化，有两种情况：一是箱子过满，即物理机发生了资源紧张；二是箱子太空，即资源没有充分利用。这两种情况都将被算法消除。注意到装箱算法的每次调度都最大化物理机资源利用率，虽然能取得较好的绿色计算效果，但负载的小波动易引起物理机的资源紧张。

装箱算法要求每个箱子的大小相同，因而物理机必须同构。而实际数据中心的设备可能有新有旧，难以满足这个条件。当然也可以将同构的物理机划分到同一个调度域中，

限制在域内进行调度来缓解这个问题。但总的说来，装箱算法在绿色计算时趋于激进，系统的稳定不佳。偏度算法更加偏向负载均衡，牺牲了部分资源利用率换取了系统的性能^[74]。偏度算法通过在调度中保持每个物理机各维度资源利用率的平衡以减少物理机资源的浪费。

偏度（Skewness）是服务器各个维度资源利用率的不平衡程度的度量，取值为各资源利用率的相对标准差。偏度指引着系统调度过程中的迁移。算法分两步进行：第一步根据预期负载找出系统中过于繁忙的服务器（热点），然后寻找一种迁移方案使得消除热点后，系统总偏度下降最多（或上升最少），如此往复直到系统中所有热点消除；第二步则是找出系统中所有负载过低的服务器（冷点），然后寻找一种迁移方案使得冷点释放后，系统总偏度下降最多（或上升最少），如此往复直到冷点无法被释放或者释放的冷点达到一定数量为止（为了维持系统稳定性的阈值）。偏度算法中有多个可调阈值：热点阈值、冷点阈值、最佳利用率阈值、绿色计算阈值，最大迁移阈值。调整这些阈值可以方便地控制系统在性能和节电之间的权衡。偏度算法使用的负载预测方法和文献[73]中相同。

在系统出现整体资源紧张的时候，热点消除步骤会找不到解：即找不到任何迁移方案能够使得热点被解决。在这种情况下，偏度算法将物理机的容量按照1.05倍的比例放大，再重新计算热点、消除热点；若依然无解，则继续放大1.05倍重新计算。这种方式实际上将负担均匀地分摊到各个物理机上，即认为：在繁忙的时候，多数应用平均地降低性能比让少数应用性能不可用要公平。与文献[55, 56]相比而言，偏度算法没有支持应用的优先级。当然，优先级在虚拟机动态组合的环境下要比本地调度困难得多。

偏度算法与装箱算法^[73]相比的一个优势是，偏度算法不要求物理服务器同构。与装箱算法相同的是，偏度算法不了解虚拟机内部的应用，不能提供应用级的SLA。

针对基于虚拟机复制的云计算资源动态伸缩调度问题，文献[75]提出了一种新的非均分负载的解决方案。通常，人们都默认云计算系统中的资源是无限的，能够满足任意的资源需求，大量的相关研究也是基于这个假设进行的，而实际上云计算仅仅是在有限的计算机集群中实现多用户间的资源共享。例如，亚马逊EC2就是利用它们的剩余资源对外提供云计算服务的，目前已有许多小型的电子商务网站在使用亚马逊EC2对外提供服务，而当系统中的多个用户同时遇到负载高峰时，例如，在Christmas Week，大量的客户涌入各种网购平台进行购物时，亚马逊自己也会遭遇负载高峰，在这种情况下，云计算系统中就有可能会出现资源紧张甚至不足的状况。文献[75]研究的问题就是在这种系统资源紧张的情况下，云计算服务提供者应该如何对各种应用的虚拟资源进行调度，使得其能够尽可能地满足更多应用对资源的需求满足率，同时在系统负载较低时，尽可能地最小化使用的服务器数量，然后通过关闭空闲的服务器来实现绿色节电计算。

文献[27]选择CPU和内存作为其主要考虑的调度资源，其中CPU作为需要分配的瓶颈资源，内存则决定了每台服务器能够同时运行的应用实例的数量。它将资源动态伸缩调度问题转化为一个CCBP模型（Class Constrained Bin-Packing Problem），它将每台服

务器看成是一个箱子，每种应用看成是一个颜色种类，应用对资源的需求看成是需要装到箱子里的物品，服务器的资源总量看成是箱子的容量，而由内存决定的每台服务器至多可同时运行的应用实例种类数看成是种类约束 c 。它创新性地提出了一个加强的半在线装箱的颜色集算法，限定除了最后一个颜色类，其他颜色类均包含 c 种颜色，不同颜色类中各自独立地采用贪心的算法对其所包含的物品进行装箱。它支持应用的资源需求随着时间的变化而变化（也即同类物品的数量可发生变化），同时也支持应用的动态加入和离开（也即颜色类的加入和离开）。当应用的资源需求发生变化时，它创新性地采用查找链的方式移动负载的分配，尽可能地避免新的应用实例的启动，有效地减少了应用实例布局变化这类代价较大的操作。同时，它采用虚拟机的 suspend 和 resume 操作来进一步优化应用实例的启动和初始化时间，通过将一台应用已经初始化完毕的虚拟机的内存镜像 suspend 到硬盘，之后在启动时，利用 resume 操作，就能快速恢复虚拟机的运行，resume 的时间与应用启动的时间是独立的，它只能取决于服务器从硬盘上读取数据的速率，现代硬盘技术能够使得这个过程只需要花费几秒的时间（而以往的常规启动需要几分钟），实验证明使用 resume 恢复一台内存为 1 GB 的虚拟机大约可节省 70% 的时间。另外，它通过将空闲的服务器设置到待机模式来节省电能的消耗，之后在需要时快速地使用在线唤醒技术唤醒服务器。TPCW 基准测试实验数据表明将空闲服务器设置到待机模式可以节省大约 85% 的电能，而使用在线唤醒技术能够使得服务器在 5 秒之内恢复工作，从而验证了其绿色节能计算的能力。因此，该策略能够较好地实现资源动态伸缩调度的目标：最大化应用需求满足率，最小化应用实例布局，以及最小化使用的服务器数量（也即最小化电能消耗）。经过严格的理论推导，其使用的服务器数量与最优结果的比值 $R = \min(2, 1 + (1 - 1/v)/(c * t))$ ，其中 v 为服务器的资源容量， c 为种类约束， t 为应用平均负载与服务器资源容量的比值。通过模拟实验以及真实环境测试，CCBP 调度算法可达到较高的需求满足率，较少的应用实例布局变化，较少的服务器使用量，非常好的规模扩展性，较好的应用请求响应时间，以及较好的 QoS 质量。

国内在本机资源调度领域的研究也在不断开展。虚拟机对内存的占有是持续的。相对于其他资源而言，调度器需要通过某种方式了解虚拟机使用内存的情况。MEB^[76] 在虚拟机管理器中记录每个虚拟机访问内存的情况，用基于 LRU 的算法计算出虚拟机的工作集大小。其内存调度策略是，基于工作集预测，如果物理机的内存能够满足虚拟机的需求，那么，剩余的内存将按工作集大小比例分配给所有的虚拟机；若出现竞争，内存调度策略是最小化总缺页数，以确保应用受到的影响最小。REMOCA^[77] 指出，现今网络的速度比磁盘读写的速度要快。REMOCA 利用数据中心其他主机的空闲内存作为本机内存和磁盘 SWAP 空间中的 cache 层，在内存出现暂时短缺的时候减少内存垃圾，保证应用性能。

5.4 国内外研究进展比较

总的来说，国内关于云计算资源调度研究都集中在三个问题：1) 如何在变化的负载

下保证应用的性能；2) 如何在保证性能的前提下，最小化系统的资源、能源开销；3) 如何在资源竞争出现时，保证资源分配的公平性。国内云计算资源调度研究方面的论文在高水平期刊和国际会议上均有不少成果发表。从研究的问题和研究成果来看，国内外的研究水平是大致相同的。另一方面，国内的研究主体主要是高校和科研院所，而国外的研究主体除此之外，工业界（如微软、谷歌等）也积极参与并有公开学术成果发表。国内工业界在云计算资源调度方面具有很强的基础和实力，例如，淘宝购物网站能承受单天上百亿的交易量，腾讯公司的微信用户数量已经超过 3 亿，冲击传统的电信业务。但是我们很少发现他们在云计算资源调度方面公开发表的学术成果。

5.5 发展趋势与展望

服务级别目标 Service Level Objectives-SLO 对云计算服务商将提供什么性能服务进行了定量描述。实现这个目标是资源调度器的任务。一般 IaaS 服务商如 Amazon、Rack-Space、盛大云等，都是按照虚拟机的性能来描述其服务级别的，例如，Amazon EC2 可以让用户指定其虚拟机有几个 CPU，多大内存，多少磁盘空间，是否有 GPU 支持等。但云计算服务使用者大都是互联网应用提供者，他们最关心的并不是云计算数据中心是否为其应用保留了多少的计算资源——应用需要的资源往往是随负载变化而变化的。他们关心应用性能：对 Web 应用来说，高吞吐量和低延迟是重要的；对在线游戏来说，低断线率和低延迟是重要的。如何通过资源调度支持服务级别目标 SLO 是云计算资源调度需要研究的关键问题之一，涉及如何建立应用性能和资源分配模型和如何隔离虚拟机性能两个方面的具体问题。以 SLO 为目标的云计算资源调度将是研究的主流。

硬件技术的持续发展给云计算资源调度不断带来了新的挑战和机遇。影响 CPU 调度的新因素是多核、众核技术。一个 Intel Xeon 7000 CPU 中有 8 个核心，配合 hyper-threading 技术，一颗 CPU 可支持 16 个逻辑 CPU。CPU 中包含越来越多的处理单元可以运行越来越多的进程，不同进程竞争总线、Cache、TLB 等硬件资源，导致性能瓶颈。新的操作系统架构^[78~81]倾向于将进程尽可能固定在某个核上面运行，并尽可能减少全局锁的使用以使得 cache missing 和流水线刷新这类费时的操作减少。纯软件 I/O 虚拟化需要在虚拟机和 VMM 之间频繁传递 I/O 请求和响应，因此效率不高。文献[62]中用 Xen 进行的实验表明，相对于非虚拟化环境，虚拟磁盘 I/O 的读写性能只有 77% 和 51%；网络收发的性能只有 45% 和 71%；该作者建议使用虚拟机客户机 OS 直接访问硬件来提高性能。这需要 IOMMU 硬件的支持。IOMMU 类似于 MMU，实现设备总线地址到内存地址的映射和保护^[82]。在这种模式下，I/O 调度由虚拟机各自管理，VMM 不再负责 I/O 调度的问题。以 PCI-SIG 规范 I/O 虚拟化为代表的硬件虚拟化技术更是将 VMM 从 I/O 虚拟化调度中解放出来。该规范允许一个硬件被虚拟成多个逻辑设备展现给虚拟机，每个虚拟机直接访问各自的逻辑设备就可以了，为 VMM 省去了设备在各个虚拟机之间进行切换的负担。

5.6 小结

云计算服务模型已经得到了工业界的广泛认同。云计算资源调度涉及本地资源调度和全局资源调度两个层次。本地资源调度主要研究的问题是资源出现紧张时，如何保证资源分配的公平性。全局资源调度的手段主要包括静态虚拟机整合、虚拟机热迁移，和应用复制。全局调度策略的两个目标是保证应用的性能和最小化能源的消耗。国内外在 this 方面的研究基本同步。展望未来，以 SLO 为目标的云计算资源调度将是研究的主流。新型硬件技术的不断发展也带来了新的挑战和机遇。

6 云计算平台中虚拟化安全的研究进展与趋势

6.1 引言

云计算具有与传统完全不同的计算模型。它将计算任务分布在第三方云计算平台中，以此获得按需的计算能力、存储空间和相应的服务。同时，由于需要获得服务器整合、高可靠性、跨平台特性等能力，虚拟化成为云计算的重要手段之一。在新的云环境中，传统的安全手段已经不能够满足新兴的安全需求了。下面列举了云安全的特性：

1) 云计算的虚拟化特性。目前的云服务提供商往往使用虚拟化技术以达到计算资源的最优化利用。虚拟化技术引入了虚拟机监控器以及特权虚拟机，这些部分的加入将增大用户的可信基 (Trusted Computing Base, TCB)。

2) 云计算的服务外包特性。在云计算模型中，租户数据的计算和存储都交由云服务提供商处理。云服务提供商在云计算模型中扮演着最终掌控数据的角色，而用户丧失了对数据、执行流的物理安全保护。

3) 云角色的三方特性。不同于传统互联网，云计算服务角色一般分为终端用户、云用户和云计算服务提供商。云用户使用云计算服务提供商的计算和存储资源为终端用户提供服务。

4) 多租户资源共享特征。由于云端的资源分配是由云计算服务提供商来决定的，所以用户的数据很有可能和其他用户甚至竞争对手共享同一台物理主机或者存储设备。

5) 云端的服务整合特性。云平台为用户提供了透明的服务，然而在云内部的服务整合中数据备份、数据和应用迁移安全、服务接口安全等都需要得到保证。

与传统的安全相比，由于丧失了对计算、存储资源的最终控制权，用户对于云平台提出了更多的安全需求，例如，对于数据、执行和服务的安全需求。与需求更新相对的，云计算的新特性带来了新的威胁，这些威胁对于现有的传统安全手段提出了极大的挑战。为了满足这些挑战云平台需要使用更全面的安全手段和技术增强整个平台的安全性。

图 13 表示了云计算安全的研究体系。

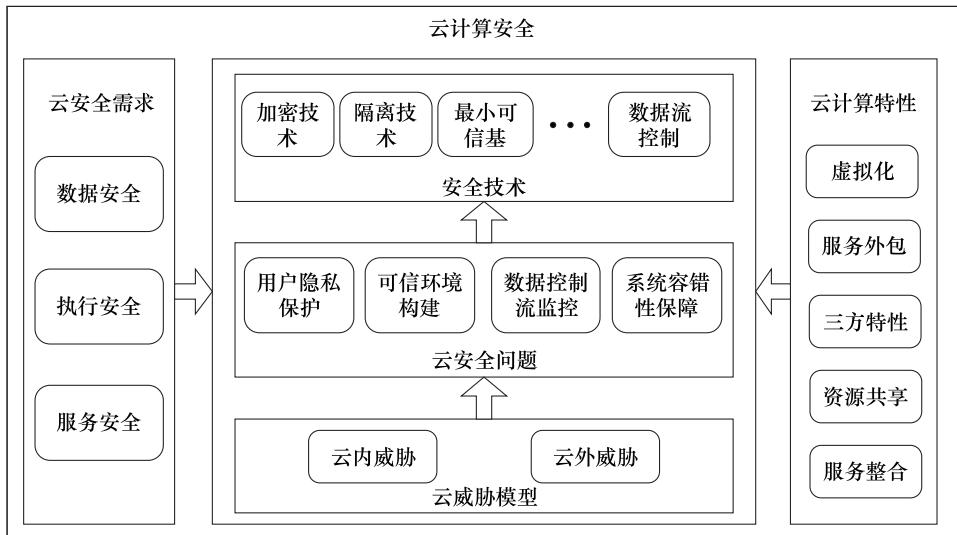


图 13 云计算安全研究体系

目前诸多主流云平台（如亚马逊、微软 Azure 等）均基于虚拟化技术。虚拟化技术为云平台安全带来了新的机遇与挑战。一方面，虚拟化技术为云平台安全提供了新的解决思路。例如，虚拟化层作为硬件与操作系统层中新的一层，可以将传统的安全增强技术应用于其中。另一方面，虚拟化软件的规模与复杂度的不断扩展也使得其本身也成为了安全攻击的对象。例如，主流虚拟化平台（如 Xen、KVM 与 VMware 等）均发现了数以十计的安全漏洞。

本节拟根据图 13 所示的研究体系，主要从基于虚拟化的云平台安全问题、挑战及其解决方案 3 个方面对云平台的系统安全进行分析与讨论。6.2 节主要介绍云安全问题的背景，其中包括云安全的特征、云安全的需求和云计算环境的威胁模型；6.3 节重点讨论了目前虚拟云平台遇到的安全挑战，以及目前相关的支撑技术；6.4 节从软件栈的层次结构角度总结了目前虚拟平台安全的解决方案、相关研究进展以及其优缺点；6.5 节对比分析了国内外研究现状。6.6 节对全文进行小结，并对未来虚拟化安全进行展望。

6.2 虚拟云平台安全威胁

上海交通大学的陈海波等对云平台的安全威胁模型进行了较深入的分析^[83]，并将安全威胁分为来自云平台外部的安全威胁与内部的安全威胁。来自云内部的威胁可以是云计算平台的管理员甚至是云计算服务提供商本身。来自于云外部的威胁包括现有的常规攻击以及由云计算特征产生的旁路攻击和对虚拟机监控器的攻击。此外，云平台资源的内部物理处理资源共享也引起了基于虚拟平台旁路攻击的可能性。

(1) 虚拟平台中的内部安全威胁

根据安全等级的不同，云平台的内部安全威胁可以归结为：1) 无特权的内部威胁，如通过云系统的内部漏洞与社会工程学等手段获取对云资源的非法访问权限；2) 恶意的云平台管理员与操作员通过对虚拟云平台的管理软件进行控制以破坏用户代码与数据的完整性与隐私性；3) 通过控制硬件以获取对云平台的控制权，这是因为由于用户失去了对主机的物理安全保护，硬件的攻击也将成为获取用户隐私的手段之一；恶意的云服务提供商可以通过物理接触、侦听硬件或者使用经过修改的恶意硬件来获得计算中的动态信息。

(2) 虚拟平台中的外部安全威胁

在虚拟云平台的外部威胁方面，我们主要将其归结为来自互联网的威胁、虚拟化层威胁和共享资源的威胁。在来自互联网的威胁中，由于云计算的规模特性使得传统的攻击手段可能造成更加严重的后果。例如在云环境中，一旦虚拟机监控器被攻破，其上所有的虚拟机都将丧失保护。另外，由于云计算平台的开放性，运行在其上的所有软件栈都可能成为攻击的目标，使得攻击面也随之增大。

6.3 问题、挑战与支撑技术

虚拟化技术的广泛应用给云平台安全带来了新的安全机遇与挑战。由于多个不同用户的虚拟机可能共享同一台物理主机，传统的防火墙很难以如此细粒度地插入虚拟机间进行访问控制。同时，在云环境中随时有虚拟机被创建和销毁，并且恶意的用户也可以租用虚拟机，网络隔离也成为一个巨大的挑战。另外，目前安全产品是基于传统软件栈设计和构建的，虚拟化层的加入有可能造成新的隐患产生，如虚拟机间的通信可能不能够被监控软件察觉^[97]。

除此之外，互联网中现有的安全问题将同样威胁着云计算平台，甚至通过虚拟化层的传播可能使得一个漏洞产生更广泛的影响。如果一台服务器遭受到了攻击，虚拟机监控器被攻击者劫持，那么在其上运行的所有虚拟机中的数据和应用都将毫无安全性可言。除了传统的安全问题，云计算的新特性也带来了新的问题和挑战，例如，利用共享资源的旁路攻击、对于虚拟化层的攻击、数据应用迁移中的安全性的保护、动态复杂云环境中的可信问题等。本节将对基于虚拟化的云平台安全中的问题、挑战以及相关支撑技术进行进一步的分析。

6.3.1 用户数据隐私保护

用户的隐私数据是云计算平台中的一个重要问题。如何保护隐私数据不受到云外、云内攻击的威胁，如何保护用户数据能够不被非法篡改和窃取是研究的重点。用户的隐私数据可以细化为静态数据和动态数据两种，静态数据是指用户的文档、报表、资料等不参与计算的隐私信息；而动态数据则是指需要动态验证或参与计算的数据。静态数据通常可以通过加密机制解决其安全性问题。然而，对于动态数据而言，由于其必须以明

文形式存在于 CPU 与内存中，其安全性往往需要通过系统的方法进行解决。

目前，对于隐私数据的保护存在以下几大主要问题：云计算中敏感数据和非敏感数据混合存储，从而难以进行权限控制；云服务提供商容易监守自盗从而窃取用户的隐私数据；用户通常外包整个数据的控制权，隐私数据的完整性难以保证。

国外研究：Craig Gentry 于 2009 从数学上提出了“全同态加密”^[123] 的可行方法^[86]，使这项技术取得了决定性的突破。由于目前软件算法还无法做到使用密文进行所有复杂计算，所以为了加密内存中的数据，许多研究使用硬件来保证动态数据的隐私性。例如，XOM^[124] 通过修改硬件保证了隐私数据在内存中仍然是以密文保存的，数据在进入 CPU 进行计算时才通过特殊硬件解密。麻省理工学院的 CryptDB^[160] 系统使用基于洋葱式的安全等级思想，对不同级别的数据保证不同的隐私性。同时 CryptDB 利用同态加密的手段使密文能够直接进行部分简单的运算，如查询、比较等，保证了高隐私级别的数据在内存中以密文存放。全同态加密是指可以在不解密的条件下对加密数据进行任何可以在明文上进行的运算。

国内研究：迄今为止，对于用户动态数据隐私保护还没有一种彻底的解决方案。同态加密可能可以提供计算过程中的保护。然而，目前其性能开销巨大从而很难被应用到日常应用的计算中。与使用密文保护动态隐私不同，使用可信计算和隔离技术也可以保护用户动态数据的安全性。复旦大学的 CloudVisor^[100] 系统通过使用嵌套虚拟化的技术对虚拟机监控器进行监控，可以防止来自云内部的攻击。CloudVisor 使用 TPM 可信启动技术，整个虚拟机将从用户加密过的镜像启动，所有敏感的 I/O 操作均经过加密。在内存保护上，CloudVisor 使用 EPT 表保护用户虚拟机不受到虚拟机监控器的非法访问。在这种情况下，用户的数据可以和虚拟机监控器很好的隔离开，隐私性因此得到保证。

6.3.2 执行环境可信性

云计算通过服务托管和服务整合可以有效提高资源利用率，从而减少成本并提供资源的灵活性和可伸缩性。然而对于公有云而言，基础设施公有化带来租户对于自身任务应用和数据安全的不可控，引发了租户对于云平台的信任难题。此外，随着功能的增多、代码的膨胀，云平台自身的可信基不断增大，从而可能暴露出更多的安全漏洞。因此，如何保障云平台执行环境的可信性是目前的一个巨大挑战。

执行环境的可信性是指云端计算环境可以忠实执行用户设定的功能，并且不执行任何没有设定的功能的性质。在云计算中如何保障与证明云端的软件的可信性是云安全面临的重大挑战。在可信计算中，用户可以使用 TPM 保证启动环境的可信性。TPM 是可信计算组织（Trusted Computing Group, TCG）为了构建可信的运行环境提出的一种具有安全存储和加密功能的可信平台模块。TPM 使用信任链传递的方式将启动代码的哈希值存放在芯片中，用户可以通过远程验证比较哈希值确定远端代码是否是可信的。然而，与 TPM 类似的基于哈希的认证很难验证依赖于输入或外部数据的程序，如 Web 服务器等，所以相关技术只能保证“启动时安全”而不能保证“运行时安全”。

国外研究：之前的经验显示，越小的程序规模往往具有越强的可信性以及越少的安

全漏洞。因此，可以通过减小可信基来防止系统中的漏洞的出现，从而增强执行环境的可信性，保证程序运行时的安全。Flicker^[108]是卡内基 - 梅隆大学提出的一个最小化可信计算基的执行架构。Flicker 通过 TPM 构建一个安全可信的运行环境供用户的应用程序使用，并且具有极小的可信基。在此架构中，操作系统和虚拟机监控器都被排除在了可信基之外，只有 CPU 和应用程序是受到信任的。

国内研究：复旦大学、华中科技大学、武汉大学、清华大学以及 EMC 中国研究院联合研发的 Daoli 项目试图为云平台的执行提供可信执行环境。其中，复旦大学的 CHAOS 系统^[89,90]通过使用虚拟机监控器，可以在一个不可信操作系统中对云服务进行保护。CloudVisor 系统^[87]则通过嵌套虚拟化技术在虚拟化软件栈不可信的情况下提供对客户虚拟机的透明保护。

6.3.3 数据/控制流保护

在传统的互联网模式中，需要监控的环境相对简单，企业只需要架设防火墙就可以监控外网的不安全流量保证内网的安全性。然而使用云计算后，用户的虚拟机需要面临动态复杂的外部环境，不仅外部的网络环境不可信任，而且运行在相同物理机上的虚拟机也有可能是恶意的，甚至支持虚拟机的监控器也可能偷窥用户隐私数据。除了虚拟机外的环境不可信，云平台内部构造的变化也给传统监控体系提出了巨大挑战。由于虚拟化技术的引入，导致用户的控制流与传统的安全模型不完全契合，攻击者可能通过虚拟化层逃避安全监控^[109]。如何在复杂动态的内外部环境中建立全面、透明、可信的安全监控机制是云安全研究的重点。

国外研究：污染跟踪技术^[142]（taint tracking）是通过标记不信任的内存地址，并追踪感染标记的方法监控程序的运行是否处于安全的状态的一种技术。污染跟踪分析能够有效地阻止大多数篡改内存并转移程序控制流的攻击。Minemu^[143]是目前速度最快的污染动态分析软件，它使用影子内存（Shadow Memory）和额外的寄存器来跟踪污染以及监控程序的运行，应用程序使用的内存和影子内存具有一个固定偏移，以此来简化计算降低性能开销。然而，尽管具有目前最好的性能，Minemu 仍然带来了 1.5 ~ 3 倍的性能损失。

国内研究：控制流完整性保护是另一种保障用户执行代码不被篡改的手段。上海交通大学的 CFIMon^[88]系统通过使用 CPU 已有的 Performance Counter 硬件来检查用户代码中的跳转语句是否跳往可信的地址，以此来防范诸如 return-to-libc 等篡改用户控制流的相关攻击。不同于之前的工作，CFIMon 可以在不修改软件源代码、编译器和运行时的情况下保证程序的控制流完整性。CFIMon 对于 Web 服务器等 I/O 密集型的程序有着良好的性能，然而对于 CPU 密集型的程序则有着比较大的性能开销。

6.3.4 云平台的容错性

云计算平台使用虚拟化技术达到了极高的资源利用率，使得云平台聚集了海量的计算和存储资源。在这种大规模高并发的情况下，如何保证系统的高可用性，如果发生了软件或者硬件的错误，如何快速侦测错误并从错误中恢复是云计算健壮性的核心需求。

除此之外，由于云平台的开放性特征，导致服务面临安全威胁以及遭受攻击的可能性增大，云平台需要防止威胁蔓延，防范单点攻击破坏整个系统；同时，云环境中多实体间有着复杂的依赖关系，云平台需要确保实体之间的隔离性，保证单个组件的崩溃不会导致所有用户的服务不可用，这些都是云计算安全的重要课题。

在云计算这个大规模分布式系统中时时刻刻都有可能出现软件或硬件的错误，许多现有系统通过冗余技术来提高服务和数据的可用性。模块化和隔离技术能够防止威胁的扩散，是增强云平台健壮性的有力手段。

国外研究：加拿大 UBC 设计与实现的 Xoar^[131] 使用微内核架构对目前已经广泛应用的虚拟机监控器 Xen^[60] 进行模块化重构并且在模块之间采取了相应的隔离措施，将原来的特权虚拟机分割为多个单一用途的服务虚拟机，增强了整个系统的健壮性。这种方法不会带来功能的损失，并且只有极小的性能开销，很适合应用在现在的云平台中。

除了主动防御，检测和恢复也是保障云平台健壮性的重要手段，在云平台中快速侦测错误的发生是系统健壮性的重要组成部分。目前侦测错误所用的方法是设置一个超时段，如果在这段时间内被侦测结点完全没有回馈则认定错误发生。使用类似的被动侦测方法可能会延迟错误发现时间，另一种方法是系统主动报告错误的发生。德州大学奥斯丁分校的 FALCON^[132] 系统提出了一种在软件栈中不同层级（硬件网络层、虚拟机监控器层、操作系统层以及应用程序层）之中插入监视器来确认上层是否仍然工作的方法。该方法能够快速检测云环境中不可用错误的发生，并通过使用更低层的监控器重启不可用的部分，保证了整个系统的高可用性和健壮性。在系统恢复方面，如果虚拟机监控器遭遇了错误，ReHype^[133] 提出一种能够从错误中快速恢复并同时保留用户虚拟机现有工作的方法。当遭遇了虚拟机监控器错误的时候，ReHype 通过使用微重启（Microreboot）技术，在不影响用户虚拟机的情况下重新启动一个虚拟机监控器的新实例并将原有状态合并到新的虚拟机监控器中，从而完成系统的恢复。

国内研究：华中科技大学金海教授团队通过研究虚拟机实时迁移算法^[102]，并采用异步拷贝和对磁盘读写的记录与回放技术来解决了广域网中对用户透明的网络连接保持以及带宽受限的问题，从而保证虚拟机服务的高效容错性；提出了一种基于虚拟机架构的软件故障容忍机制^[103]，采用带权值的营救点和错误虚拟化技术来容忍软件故障，并针对云计算环境下存在多软件实例的情况，采用在多软件实例之间共享软件故障容忍信息来快速恢复故障；提出了一种被动式的软件栈缓冲区溢出故障容忍机制^[104]，其能够精确诊断云平台服务中软件的内存故障，寻找出引发故障的栈缓冲区，并采用访存虚拟化将一个静态栈对象重分配到受保护的内存区域来进行保护，从而容忍软件栈溢出故障。

北京大学的基于虚拟机的零宕机服务 Taiji^[172,173] 是通过虚拟机快速同步实现的容错服务，可以保证系统出现故障时，关键性应用仍能正常运行。Taiji 包含两台物理服务器，分别称为主服务器和备份虚拟机，两台服务器都部署了虚拟机管理器，分别称为主虚拟机管理器和备份虚拟机管理器。关键性应用程序运行在主虚拟机中，虚拟机管理器通过快速复制技术把主虚拟机的状态实时备份到备份服务器上，当服务器出现故障时，关键性应用可以在备份服务器上继续运行。此外，北京大学的 Twinkle 系统^[174] 是针对应用逻

辑层的高可用性问题设计的快速启动技术，当互联网服务出现故障或者资源瓶颈时，Twinkle 通过快速启动新的服务副本，可以较好地保证服务质量。

6.3.5 其他云安全问题

(1) 资源共享问题

云计算的一个重要特征是云中的多租客共享同一物理设备。资源共享能够有效地利用闲置的硬件计算、存储能力，同时为用户提供更强的可拓展性和灵活性。然而，资源共享可能会给用户带来隐私泄露的问题。

国外研究：由于资源共享可能导致旁路攻击（Side Channel Attack）的产生。旁路攻击是指使用来自（如物理实现等）非直接传播途径的信息进行的隐私窃取攻击，而不是基于系统本身逻辑漏洞。由于旁路攻击能够绕开系统本身的防御逻辑，所以往往很难防御。目前旁路攻击主要集中于共享 CPU 缓存的环境中^[161]。HomeAlone^[164]通过逆向使用 L2 缓存的旁路攻击技术作为防御检测的工具来确认是否有其他虚拟主机运行在同一台物理服务器上。

除共享缓存旁路攻击外，内存的去冗余机制（Memory Deduplication）也可能使攻击者窥探到用户运行程序^[162]，其他旁路通道由于信号辨析难度大目前还没有形成有效的攻击。

资源共享的另一个问题是不同用户存在着共享声誉。在一些情况下，如果在云平台中有一个恶意用户的存在，那么有可能会导致许多无辜用户的服务不可用，从而变相的攻击了这些用户的服务可用性。例如曾经有人使用 EC2 的服务制造了大量的垃圾邮件，导致国际反垃圾邮件组织将大段的 EC2 的 IP 地址列入黑名单。这给其他使用 EC2 进行邮件工作的用户带来的巨大的困扰，需要使用邮件服务的用户不得不单独向 Amazon 提交申请^[144]。

国内研究：复旦大学的施继成等人^[94]提出一种缓存染色的算法来保证安全相关的操作不会与其他虚拟机共享，从而避免了共享缓存旁路攻击的产生。

(2) 基于虚拟化的云计算服务整合

云计算能够提供给用户透明的服务体验，将用户从繁琐的硬件维护、负载均衡、性能调优中解脱出来。然而由云平台包办的服务整合过程中可能会暴露一些安全隐患。首先，用户数据的地理位置可能威胁用户隐私。例如，谷歌和亚马逊等大型云服务提供商往往在全球范围内都拥有数据中心，当用户上传了数据就很难确认数据究竟会被存放在什么地点，不同国家地区对于数据的审查可能会暴露用户隐私。其次，数据中心间和数据中心内部都会进行数据迁移和应用的迁移以保证数据可靠性和集群中的负载均衡。然而，在迁移过程中用户的信息需要经过互联网，这有可能遭遇攻击者的侦听和嗅探。最后，云计算往往需要整合各级服务，如将 SaaS 或 PaaS 搭建于 IaaS 之上，如何确认和保证各层间通信接口的完整性和安全性也是一个重要问题。

(3) 用户数据销毁问题

在云计算中，用户隐私数据的生命周期难以保证其完整性。如果云服务提供商私自拷

贝或者对用户虚拟机进行快照 (Snapshot)，用户很难确认数据是否真正销毁以及是否泄漏。

国外研究：华盛顿大学的 Vanish^[163] 系统提出利用 P2P 网络的分布式哈希表对密钥进行存储，利用其自然的衰减来实现密钥的销毁。但是，Vanish 提出的密钥衰减的方法只能确认用户静态数据的销毁。

国内研究：复旦大学的 Dissolver^[140] 系统则提出了一种动态数据销毁的解决方案。它保证用户的数据仅存在于私密的运行空间中，在用户指定的时间点，内存中的动态数据将会被强制销毁。

6.4 多层次虚拟化平台安全解决方案

在上述章节中主要分析了基于虚拟化技术的云平台的安全需求，威胁模型以及面临的问题、挑战与支撑技术。本节将结合虚拟化相关技术，详细地给出云安全的具体解决方案与目前相关的研究系统。根据虚拟化平台现有的架构，软件栈可大致划分为四层：应用程序层、操作系统层、虚拟化层以及硬件层。在层次结构中，高层的软件将向其下一层请求服务，并在安全性上依赖于下层的安全。系统安全增强技术就是通过在下层的安全策略来保证上层软件的安全可靠。在本节中，我们主要讨论云计算安全相关的技术和方法，对于常规安全技术手段（例如通过操作系统保护应用程序等）则不进行详细的介绍。基于虚拟化的云计算平台安全技术的层次划分如图 14 所示。

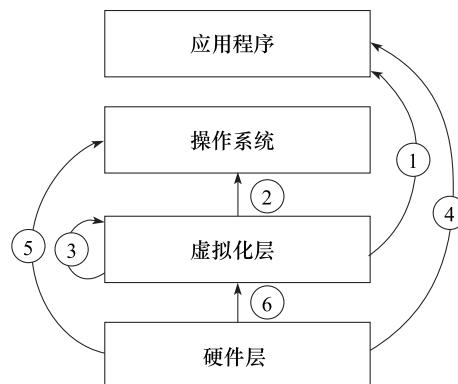


图 14 云计算解决方案层次结构

表 2 列举了目前在系统安全方面的相关研究工作。

表 2 当前系统安全方面的相关研究工作

序号	所处层	保护层	国际相关工作	国内相关工作
1	虚拟化层	应用程序	Overshadow ^[145]	复旦大学 CHAOS ^[89,90] 武汉大学 ^[100] 南京大学 ^[98] 国防科技大学 ^[16] 华中科技大学 ^[105]

(续)

序号	所处层	保护层	国际相关工作	国内相关工作
2	虚拟化层	操作系统	Terra ^[146] , SecVisor ^[147]	复旦大学 CloudVisor ^[87] 华中科技大学 ^[99,106,107] 北京航空航天大学 ^[100] 清华大学 ^[101]
3	虚拟化层	虚拟化层	HyperSentry ^[148] , HyperSafe ^[149] , Nova ^[8]	
4	硬件层	应用程序	Flicker ^[108] , XOM ^[124,150] , AEGIS ^[151] , SP ^[152] , Bastion ^[153]	
5	硬件层	操作系统	HyperWall ^[154]	HyperCoffer ^[97]
6	硬件层	虚拟化层	NoHype ^[155]	

6.4.1 虚拟化层对应用程序的保护

由于目前的商用操作系统大都使用宏内核设计，使得代码庞大、功能繁多、设计复杂，很难保证没有漏洞产生。在云环境中，可以使用虚拟化技术将操作系统排除在可信计算基（TCB）之外。这样即使是操作系统遭到入侵者劫持，云端用户的应用程序仍然是安全的。

国外研究：VMware 公司的 Overshadow^[145] 系统提供了两层隔离，第一层是客户操作系统与 VMM 的隔离，另一层是加密的应用程序与客户操作系统间的隔离。客户操作系统中可运行多个不同的应用程序，包括保护和未保护的。Overshadow 引入了“垫片”(Shim) 的概念，它可以插入客户操作系统与受保护的应用程序之间。垫片处在受保护应用程序的地址空间内，绕开了操作系统与 VMM 直接进行协作。操作系统与受保护的应用程序对于同一块内存的视图是不同的：从应用程序角度，其内存以明文形式存在；而从操作系统角度，应用程序的内存以加密形式存在。VMM 通过截获运行权限的切换，完成对内存视图的转换（即加解密）。此时，操作系统仅负责计算资源的分配，而无法访问应用程序的数据，从而保证即使操作系统被攻击者入侵，用户数据也能得到保护。Overshadow 需要对同一份内存保存两份拷贝，引入了一定的内存开销。同时，在虚拟机监控器不可信的情况下，用户隐私仍然可能泄露。

国内研究：复旦大学的 CHAOS^[91] 系统提供了两层隔离，第一层是客户操作系统与 VMM 的隔离，另一层是加密的应用程序与客户操作系统间的隔离。CHAOS 为操作系统和需要保护的进程提供两份页表。用户的隐私数据将会从操作系统的页表中清除，在保证恶意操作系统无法访问用户数据的同时，仅引入了较小的内存开销；另一方面，CHAOS 会记录内存页的所属者，并拒绝操作系统映射含有数据隐私的页。

为了防止虚拟机之间的信息泄露，武汉大学王丽娜教授团队对虚拟机间通过虚拟机层因外部映射（Foreign Mapping）引起的隐私数据泄露进行了分析^[95]，并且提出通过应用指导的内存隐私保护策略。应用可以指定特定内存不被外部映射，从而虚拟机监控器

保证并且强制这一点。南京大学李小庆等人^[98]提出了一种基于硬件虚拟化技术的单向隔离执行模型，通过将应用程序按照安全需要划分为宿主进程与安全敏感模块，并且在虚拟化层提供一个安全隔离器以支持安全敏感模块的隔离以及与其他不安全进程的交互。由于安全隔离器执行于虚拟化层，因而具有较小的可信基。

针对客户虚拟机中客户操作系统与不可信代码之间的隔离问题，国防科技大学王怀民教授团队提出了一种安全的虚拟机执行环境^[16]，对操作系统与不可信代码之间通过本地虚拟化技术进行隔离，同时提供了为隔离的执行环境提供尽可能原生的操作系统服务。

华中科技大学金海教授团队等提出了一种基于虚拟化技术的安全签名系统 Assured Digital Signing (ADS)^[105]，通过虚拟化技术和可信计算的优势将安全敏感的签名密钥和签名服务移植到可信虚拟机中，在不可信客户虚拟机中仅仅执行对安全不敏感部分且对应用透明，利用虚拟机管理器可信基较小的特点实现了内存映射保护机制以保护通信通道、应用程序和输入输出，实现了在不影响签名效率的前提下为用户程序提供安全可信的数字签名服务。

上海交通大学管海兵教授团队提出了一系列应用程序编程接口 (API)^[170]，使得运行在云环境下的应用程序可以访问隔离的数据和代码。该系统修改了虚拟机监控器，使得其能够直接控制用户应用程序的安全，避免了不可信操作系统和其他不可信进程对敏感数据的窃取。

6.4.2 虚拟化层对于操作系统的保护

在公有云环境中，用户的虚拟机运行在远端服务器中。由于云平台中的虚拟机监控器具有对客户虚拟机完全的控制权，因此用户不得不选择信任云服务提供商及其雇员。然而，云平台的安全威胁可能来自于云服务提供商内部对于租户隐私数据的非法访问。如何保证远端客户虚拟机的隐私性与完整性成为了云安全的重要问题。

国外研究：在传统的软件架构中，驱动程序与特权系统服务是除内核外运行在内核态的组件。这些组件的代码量庞大，并且具有较大的受攻击面，因此很难保证其完整性和安全性。斯坦福大学的 Terra^[146] 系统提出了一个基于虚拟机的可信计算架构。它将不同的虚拟机定制成不同的安全级别，不同安全级别的应用程序运行在不同的虚拟机中。无特殊安全要求的应用程序运行于未经特殊保护的虚拟机中；具有安全保密要求的应用程序则被单独隔离在一个磁盘镜像经过加密处理的虚拟机中。Terra 通过在虚拟机启动之前对其进行度量来防止对环境的篡改，并且负责向远程第三方证明安全的虚拟机的可信性。然而 Terra 仍然需要相信虚拟机监控器，并且即使在安全的虚拟机中仍然无法完全避免恶意的攻击者。卡内基 - 梅隆大学的 SecVisor^[147] 则是使用一个微型的虚拟机监控器保证上层操作系统内核的代码完整性。通过在虚拟化层对内存的保护，SecVisor 可以保证在整个操作系统的生命周期里，只有经过用户确认的代码才能够运行在内核态，从而防御了对内核的代码注入等攻击，增强了操作系统的安全性。但是，知晓 SecVisor 存在的攻击者可以通过对内核数据结构进行破坏或者劫持控制流来进行攻击。因此，SecVisor 只能增加攻击的难度而不能防御来自内核漏洞的攻击。

国内研究：为了防止虚拟机监控器的恶意行为，复旦大学的 CloudVisor^[100] 系统使用嵌套虚拟化技术对虚拟机监控器进行监控。在 CloudVisor 的架构中，虚拟机监控器将被置于非特权指令集中，并使用扩展页表保护用户虚拟机不受到未授权的访问。虚拟机监控器对用户虚拟机进行的所有敏感操作都将下陷到 CloudVisor 中接受安全检查。该架构很好地将安全性与功能性解耦合，在保证功能完备的同时尽量减小了可信基的大小，有利于对可信基进行进一步的正确性验证。在 CloudVisor 中可信基仅有 5500 行代码，相比于庞大的虚拟机监控器具有更加简单的逻辑及更高的可信性。

虚拟化的引入也对安全监控技术提供了新的思路，华中科技大学金海教授团队对通过虚拟机进行安全监控进行了较深入的研究^[99]，对虚拟机安全监控在入侵检测、蜜罐、文件完整性监控、安全监控架构和恶意代码检测与分析等方面的应用进行了较深入的分析，并根据相关技术将安全监控分为内部监控与外部监控。此外，为了适应不同类型虚拟机，从虚拟机底层着手提出了一种基于驱动的通用监控机制——VMDriver^[106]，通过两种设计策略实现通用监控：1) 将客户操作系统中的事件截获和语义恢复相分离，事件截获在虚拟机管理器层，而语义恢复在管理域中；2) 语义恢复通过管理域中的监控驱动来实现，采用传统操作系统中的设备驱动的方式，语义恢复以内核模块方式加载到管理域中，从而实现动态地屏蔽虚拟机中客户操作系统的差异性，并向应用提供统一的接口。在虚拟机通用监控的基础上，提出了面向云环境的第三方可信监控域，提供对管理员和用户操作行为的准确监控、可信度量和报告^[107]。

由于操作系统可能会被恶意软件劫持从而将相关病毒进行隐藏，为此，北京航空航天大学的怀进鹏院士团队提出了一种基于虚拟机监控器的隐藏对象检测方法^[100]。主要的思路是通过显式与隐式的方式对操作系统的对象建立多个视图并进行关联，从而可以检测包括进程、网络与文件等隐藏操作系统对象。目前，基于虚拟化进行安全监控已经具有较多的研究。清华大学的诸葛建伟团队对当前的相关研究进行了综述^[101]，并且比较了不同方法与引用。

在虚拟机在线系统状态监控和电子取证方面，需要解决机器一致状态的准确获取问题，上海交通大学的 Vis 系统^[171] 采用虚拟快照和延迟复制技术，通过硬件虚拟化提供的嵌套页表，监控对目标系统的实时修改，从而获得一致的系统快照。

6.4.3 虚拟化层对自身的安全增强

相比于操作系统，虚拟机监控器的复杂程度大大降低，安全性也更容易验证。但近年来随着功能的不断增加，虚拟机监控器的代码规模也急剧增大。随之而来的是对虚拟机监控器及其控制软件的漏洞攻击。这就需要安全增强技术保证虚拟机监控器本身的安全。

国外研究：虚拟化技术使得云计算具有巨大的服务整合能力，在一个虚拟机监控器上很有可能运行着许多虚拟机和服务。一旦虚拟机监控器被攻破，所有的组件、虚拟机和服务都将受到威胁。为了防御这种缺陷，NOVA^[8] 提出构建一个微内核的虚拟机监控器，把传统的单个整体的虚拟机监控器分装成一个基于组件的系统，通过引入不同组件

之间基于组件功能的访问控制，到达提高虚拟机监控器安全性的目的。

NOVA 将目前的虚拟机监控器划分为更加细粒度的功能组件，包括多个虚拟机监控器、根分区管理、设备驱动、微型监控器以及其他系统服务。在这些组件之间的接口上保证最小特权原则，这样就可以防止攻击者攻破局部就可以获得整个系统的控制权。在权限控制上，系统中只有微型监控器运行在宿主机的特权级上。相比于目前的虚拟机监控器，微型监控器功能明确、代码精简所以更容易保护其安全性。NOVA 通过缩小特权集、模块化组件以及进行组件间隔离等方法，增强了虚拟化层的安全性。但是，NOVA 与现有的虚拟机监控器一样仍然需要其他机制来保护用户隐私不受来自云内、云外攻击的威胁。

另外，HyperSentry^[148] 和 HyperSafe^[149] 试图通过动态检测虚拟机监控器的完整性或强制要求其控制流的完整性来实现增强虚拟机监控器的安全性的目的。其中，HyperSafe 使用了两个技术在 hypervisor 层内部实现自我保护：首先是内存锁，即通过 x86 处理器 CR0 寄存器的 WP（Write Protection）位、内存页表的只读位，以及内存页表所在页的只读位，来保护只读代码不受更改；其次是间接跳转表，即所有返回地址均根据跳转表间接跳转，跳转的目标地址通过静态分析得到，跳转表通过内存锁保护。

国内研究：目前，国内对这方面的研究较少，有必要进行进一步的研究。

6.4.4 硬件对应用程序的保护

在云计算环境中，用户需要应用执行在一个可信的环境中。然而，传统的宏内核操作系统可能含有安全漏洞。除此之外，随着虚拟化层的代码量越来越大，虚拟机监控器本身被攻击面也逐渐扩大。一些研究者希望能够将操作系统和虚拟化层从可信基中排除，仅仅剩下可信的硬件和应用程序。

国外研究：Flicker^[108] 是卡内基 - 梅隆大学提出的一个最小化可信计算基的执行架构。在 Flicker 中，可信基只有 CPU 和需要保护的应用，其他软件均是不受信任的。在运行安全应用之前，Flicker 将暂停整个系统的运行，并使用 TPM 构建一个干净可信的运行环境。在运行过程中，该程序独占硬件资源，保证其安全性。运行结束后，Flicker 将恢复原始系统的运行状态。Flicker 的基础架构让高安全性要求的代码在一个完全隔离的环境中执行。Flicker 还能够对所执行的代码（包括其输入和输出）进行有意义、细粒度的远程证明。即使 BIOS、操作系统和支持 DMA 的设备全部都是恶意的，Flicker 也能够保证以上特性。但是，Flicker 在创建一个安全执行环境时，会暂停整个系统的工作，并且需要一个排他的运行环境。这就破坏了云计算高并发和资源共享的特性，因此 Flicker 并不适用于云计算环境。

另外的研究者使用硬件加密体系结构来保护用户的隐私不会泄露。XOM^[124] 与 AEGIS^[151] 在 CPU 中加入了安全增强的模块和加解密逻辑。用户程序代码和关键性数据在内存中以加密状态存在，运行时 CPU 加载数据和指令后在 CPU 内部解密，在交换出 CPU 之前恢复加密保护，并用哈希值来保证数据的完整性。XOM 系统上运行的受保护程序需要进行修改，通过调用特殊的访存指令来读入或写出受保护的数据。XOMOS^[150] 则

是在 XOM 处理器上实现的操作系统，利用了 XOM 提供的硬件特性，在可信的硬件平台实现不可信的操作系统。虽然 XOM 和 XOMOS 能够为用户隐私带来良好的保护，但是 XOM 架构到目前为止还未有真正的实现，并且 XOMOS 需要重构软件栈，不具有良好的兼容性。

与 XOM 比较接近的是 SP 构架^[152]和 Bastion 系统^[153]。SP 构架针对的是关键性数据的保护，而没有对加密代码的支持。由于 SP 的目标更单一，它对 CPU 的改动比 XOM 等要小很多，但是在 SP 上运行的用户程序仍需要进行修改。

国内研究：目前，国内对这方面的研究较少，有必要进行进一步的研究。

6.4.5 硬件对于操作系统的保护

在 IaaS 模型中，用户提供自己的操作系统运行在云计算平台上。云服务提供商的恶意雇员可能通过虚拟机监控器的管理员权限破坏用户的操作系统或者是窃取用户隐私数据。在这种云环境下，需要通过硬件来保护用户的操作系统安全不受侵害。

国外研究：HyperWall^[154]提出使用隐私保护表（Confidentiality and Integrity Protection Table, CIP）保护用户虚拟机的隐私性、完整性和隔离性。该表存储在物理内存尾端，并且只能被硬件访问。隐私保护表中保存了虚拟机监控器和 DMA 对于用户虚拟机内存的访问控制范围和权限，并会随着用户虚拟机的创建和内存变动而更新，从而保证隐私保护表的实时性。当虚拟机监控器要访问用户虚拟机相应内存时会发生 TLB 不中的情况，在此时 HyperWall 会检查虚拟机监控器访问的地址是否经过授权，防止虚拟机监控器偷窥用户隐私数据。在用户虚拟机产生中断时，所有的通用寄存器会被加密并且他们的哈希值会被储存在安全寄存器中，以防止数据的泄露和篡改。

国内研究：目前，国内对这方面的研究较少，可进行进一步的研究。

6.4.6 硬件对于虚拟化层的保护

主要是为了防止虚拟化层被攻陷，加强虚拟机之间的隔离，并且防止恶意虚拟机对用户代码和数据的窥探。

国外研究：普林斯顿大学的 NoHype^[155]系统提出使用硬件隔离技术静态分配虚拟机资源，去除整个动态虚拟化层，将虚拟化工作完全使用硬件完成。在 NoHype 体系结构中，每个虚拟机使用一个核，从而避免旁道攻击的可能性。从硬件上分隔物理内存，保证不同虚拟机之间的隔离性。修改 I/O 设备保证虚拟机能够直接接入相应的物理设备，从而消除了不同虚拟机的数据混杂，移除动态的虚拟化层能够带来显著的安全性提升。

国内研究：HyperCoffer^[97]提出了通过对处理器进行扩展以在虚拟机监控器与片外硬件不可信的情况下对虚拟机的完整性与隐私性的保护。HyperCoffer 通过在处理器中增加加密与完整性保护单元，从而对片外与片内的数据交换进行加解密与完整性验证，因而片外数据的隐私性与完整性可以得到有效保障。为了保障虚拟机监控器与被 HyperCoffer 所保护的虚拟机的正常通信，HyperCoffer 提出了通过运行于客户虚拟机空间的 VM-Shim

以支持对必要的交互数据的安全保护，并且同时兼容现有虚拟机的功能。

6.5 国内外研究进展比较

云计算等新型计算模式的出现对安全提出了新的需求，而虚拟化技术作为主流云计算平台的支撑技术，为云平台安全性提供了机遇与挑战。从以上分析中可以看出，国外在虚拟化安全方面开展了大量的研究，在用户数据隐私保护、执行环境可信性、数据与控制流保护、容错与资源安全共享方面取得了一系列的研究成果。包括麻省理工学院、斯坦福大学、卡内基-梅隆大学、普林斯顿大学等在内的国外顶级名校均在虚拟化软件栈的各层取得了一系列研究成果。并且，国外相关产业界巨头如 VMware、亚马逊等也纷纷开展了相应的研究。

国内方面，得益于虚拟化开源产品（如 Xen、KVM 等）的开放性以及国内云计算概念的深入与推广，虚拟化安全相关的研究也如火如荼地开展着。目前，国内一批高校如上海交通大学、复旦大学、华中科技大学、北京大学、国防科技大学、北京航空航天大学、清华大学、浙江大学与武汉大学等均在这方面开展了较广泛的研究。从安全属性而言，国内高校在用户数据隐私保护与执行环境可信性和容错等方面取得了国际认可的研究成果。然而，在数据与控制流保护与资源共享方面的研究尚未广泛开展。从虚拟化软件栈的保护层次来看，在虚拟化层对应用程序与操作系统等方面开展了深入的研究并取得了国际一流的研究成果。然而，在虚拟化层对虚拟化层的保护以及基于硬件的安全保护方面研究较少。这些缺口值得成为国内研究者在虚拟化安全方面开展研究的方向。

6.6 小结

虚拟化作为主流云平台的关键技术，如何利用虚拟化增强云平台安全性与如何提高虚拟化自身的安全性变得非常重要。本节旨在介绍云计算环境下虚拟平台安全的背景，包括云安全的新特性、云安全的需求以及云安全的威胁模型。其次，本节分析了云安全中用户隐私数据保护、执行环境可信性、数据/控制流保护以及云计算健壮性等问题以及现有的安全支撑技术。最后，本节按照目前云计算软件栈的层次结构，对现有的云安全系统进行了全面的研究与分析。

目前虚拟云平台中仍有许多安全问题需要解决，集中体现在以下几个方面：

1) 用户隐私数据保护。随着软件向云进行迁移的过程中，越来越多的用户隐私数据可能向云进行迁移。如何有效保护用户数据在传输、使用与共享过程中的安全性，仍然是虚拟云平台的一个主要研究话题。

2) 特权用户访问。云平台操作员等特权用户的非法访问将导致用户隐私的大规模、长时间泄露^[85,112]。尽管当前已经有了相关研究来减小云平台可信基以不再过多地依赖云平台操作员，然而，目前的解决方案仍然具有不可忽视的性能损失以及与现有虚拟化平

台的兼容性问题。如何提供一个低成本且后向兼容的解决方案变得非常重要。

3) 云服务的高可用性。云平台资源的高度集中化也使得云平台高可用性变得非常重要。然而，虚拟化平台的复杂性、云管理软件规模的庞大性使得其保障云平台的高可用性变得非常困难。例如，微软、谷歌与亚马逊等主流云平台提供者今年均发生了多次服务中断。如何通过虚拟化提供的迁移、快照等技术提高云服务的可用性并提供快速恢复仍然值得深入研究，尤其是在提供解决方案的可扩展性与降低性能开销方面。

7 总结

本报告以云计算为背景，选取从云端到终端的 5 项计算系统虚拟化关键技术——嵌入式虚拟化技术、GPU 虚拟化技术、云计算平台中虚拟资源性能计量方法、云计算资源管理与调度和虚拟化安全，对近年来国内外的研究进展进行综述和比较，并对未来的发展趋势进行分析，探索支撑云计算的计算系统虚拟化新特征、新需求、新技术和新的解决方案，希望能够对国内相关的研究人员有所启发。

参考文献

- [1] 金海 . 计算机系统虚拟化：原理与应用 [M] . 北京：清华大学出版社，2008.
- [2] 石磊，邹德清，金海 . Xen 虚拟化技术 [M] . 武汉：华中科技大学出版社，2009.
- [3] John Rushby. Design and Verification of Secure Systems, In Proceedings of the 8th ACM Symposium on Operating Systems Principles (SOSP'81). ACM, NY, USA, 12-21.
- [4] Trango Virtual Processors, http://en.wikipedia.org/wiki/Trango_Virtual_Processors.
- [5] PikeOS, <http://www.sysgo.com/products/pikeos-rtos-and-virtualization-concept/rtos-technology/>.
- [6] AUTOSAR, <http://en.wikipedia.org/wiki/AUTOSAR>.
- [7] Mach, <http://www.cs.cmu.edu/afs/cs/project/mach/public/www/mach.html>.
- [8] Steinberg U, Kauer B. NOVA: A Microhypervisor-Based Secure Virtualization Architecture. Proceedings of the 5th European conference on Computer systems, ACM, 2010, 209-222.
- [9] Xen-ARM, http://www.xen.org/products/xen_arm.html.
- [10] Wind River Hypervisor, <http://www.windriver.com/products/hypervisor/>.
- [11] KVM-ARM, <http://systems.cs.columbia.edu/projects/kvm-arm/>.
- [12] Cells, <http://systems.cs.columbia.edu/projects/cells/>.
- [13] VirtualLogix, <http://www.virtuallogix.com/>.
- [14] SMART 系统, <http://embedded.zju.edu.cn/smart.htm>.
- [15] 王吉辉. 嵌入式实时系统的安全策略研究 [J]. 软件导刊, 2012, 11(7).
- [16] 温研, 王怀民. 基于本地虚拟化技术的隔离性执行模型研究 [J]. 计算机学报, 2008, 31(10).
- [17] VNeo, <https://code.google.com/p/vneo/>.

- [18] 左保京. 嵌入式虚拟化中时间可预测调度算法研究(学位论文), 上海: 上海交通大学, 2011.
- [19] 李茜, 崔丽. 基于嵌入式虚拟化技术的手机安全方案, 微计算机信息, 2012(09).
- [20] 罗晶. 嵌入式虚拟机系统镜像存储的研究[D]. 武汉: 华中科技大学, 2012.
- [21] 李海峰. 基于嵌入式虚拟机平台的域间数据共享系统(学位论文), 武汉: 华中科技大学, 2012.
- [22] 伍箴水. 嵌入式系统中设备虚拟化的关键技术研究(学位论文), 武汉: 华中科技大学, 2012.
- [23] Dowty M, Sugerman J. GPU Virtualization on VMware's Hosted I/O Architecture[J]. ACM SIGOPS Operating Systems Review, 2009, 43(3): 73-82.
- [24] Lagar-Cavilla H A, Tolia N, Satyanarayanan M, et al. VMM-Independent Graphics Acceleration, Proceedings of the 3rd International Conference on Virtual Execution Environments. ACM, 2007: 33-43.
- [25] Hansen J G. Blink: Advanced Display Multiplexing for Virtualized Applications, Proceedings of NOSSDAV. 2007.
- [26] Gupta V, Gavrilovska A, Schwan K, et al. GVIM: GPU-Accelerated Virtual Machines, Proceedings of the 3rd ACM Workshop on System-level Virtualization for High Performance Computing. ACM, 2009: 17-24.
- [27] Weggerle A, Schmitt T, Löw C, et al. VirtGL-a Lean Approach to Accelerated 3D Graphics Virtualization, Cloud Computing and Virtualization 2010.
- [28] How GRID Brings Amazing Graphics to the Enterprise Cloud, (NVIDIA) GPU Technology Conference 2013, <http://on-demand.gputechconf.com/gtc/2013/presentations/S3501-NVIDIA-GRID-Virtualization.pdf>.
- [29] Shi L, Chen H, Sun J, et al. vCUDA: GPU-Accelerated High-Performance Computing in Virtual Machines, Computers, IEEE Transactions on, 2012, 61(6): 804-816.
- [30] P Mell, T Grance. The NIST Definition of Cloud Computing (Final Version), [Online] <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>, 2011.
- [31] Amazon EC2 Instance Types, [Online] <http://aws.amazon.com/cn/ec2/instance-types/>.
- [32] Babcock, Charles. Why Cloud Pricing Comparisons are So Hard, InformationWeek, June 6, 2012, [Online] <http://www.informationweek.com/cloud-computing/infrastructure/why-cloud-pricing-comparisons-are-so-hard/240001491>.
- [33] Google Compute Engine, [Online] <https://cloud.google.com/pricing/compute-engine>.
- [34] 阿里云云服务器技术白皮书, [Online] <http://bbs.aliyun.com/job.php?spm=0.0.0.0.E6LJ0r&action=download&aid=18823&mt=4081066da4b12cf61c90d2a187d71b89>.
- [35] Chona S Guiang, Kent F Milfeld, Avijit Purkayastha, John R. Boisseau, Memory Performance of Dual-Processor Nodes: Comparison of Intel Xeon and AMD Opteron Memory Subsystem Architectures, In Proceedings for ClusterWorld Conference and Expo 2003, 2003.
- [36] Y Koh, R Knauerhase, P Brett, M Bowman, Z Wen, C Pu. An Analysis of Performance Interference Effects in Virtual Environments, In Proceedings of IEEE International Symposium on Performance Analysis of Systems & Software, ISPASS 2007.
- [37] A Iosup, S Ostermann, N Yigitbasi, R Prodan, T Fahringer, D H J Epema. Performance Analysis of Cloud Computing Services for Many-Tasks Scientific Computing, IEEE Transactions on Parallel and Distributed Systems. vol. 22, no. 6, pp. 931-945, 2011.
- [38] A Iosup, R Prodan, D Epema. IaaS Cloud Benchmarking: Approaches, Challenges, and Experience, In the 5th Workshop on Many-Task Computing on Grids and Supercomputers (MTAGS), with the ACM/IEEE

- Conference on High Performance Networking and Computing, 2012.
- [39] Ostermann S, Iosup A, Yigitbasi N, Prodan R, Fahringer T, Epema D. An Early Performance Analysis of Cloud Computing Services for Scientific Computing, Technical Report PDS-2008-006, Delft University of Technology, 2008.
- [40] Jiang Dejun, Guillaume Pierre, Chi-Hung Chi. EC2 Performance Analysis for Resource Provisioning of Service-Oriented Applications, Proceedings of the 2009 international conference on Service-oriented computing, 2009.
- [41] J Schad, J Dittrich, J Quiane-Ruiz. Runtime Measurements in the Cloud: Observing, Analyzing, and Reducing Variance, Proceedings of the VLDB Endowment, vol. 3, no. 1-2, pp. 460-471, 2010.
- [42] Michael Armbrust, Armando Fox, Rean Griffith, Anthony D Joseph, Randy Katz, Andy Konwinski, Gunho Lee, David Patterson, Ariel Rabkin, Ion Stoica, Matei Zaharia. A View of Cloud Computing, Communications of the ACM, vol. 53, p50-58, 2010.
- [43] Zheng Li, O'Brien L, He Zhang, Cai R. On a Catalogue of Metrics for Evaluating Commercial Cloud Services, In Proceedings of 2012 ACM/IEEE 13th International Conference on Grid Computing, 2012.
- [44] Zheng Li, O'Brien L, He Zhang, Cai R. A Factor Framework for Experimental Design for Performance Evaluation of Commercial Cloud Services, In Proceedings of 2012 IEEE 4th International Conference on Cloud Computing Technology and Science, 2012.
- [45] RG Cloud Working Group: SPEC Research Group, [Online] <http://research.spec.org/working-groups/rg-cloud-working-group.html>.
- [46] Z Hill, J Li, M Mao, A Ruiz-Alvarez, M Humphrey. Early Observations on the Performance of Windows Azure, Proceedings of the 19th ACM Int. Symp. High Performance Distributed Computing (HPDC 2010), ACM Press, 2010.
- [47] Cloud Servers Pricing by Rackspace Cloud Computing & Hosting, [Online] <http://www.rackspace.com/cloud/c/servers/pricing/>.
- [48] 腾讯云, [Online] <http://www.qcloud.com/>.
- [49] Armbrust, Michael Fox, Armando Griffith, Rean Joseph, Anthony D. Katz, Randy H. Konwinski, Andrew Lee, Gunho Patterson, David A Rabkin, Ariel Stoica, Ion Zaharia, Matei. Above the Clouds: A Berkeley View of Cloud Computing, Technical Reports of UC Berkeley, February 10, 2009.
- [50] Server Consolidation, <http://www.vmware.com/solutions/consolidation/>.
- [51] Factors Affecting Electricity Prices, http://www.eia.doe.gov/energyexplained/index.cfm?page=electricity_factors_affecting_prices.
- [52] Report to Congress on Server and Data Center Energy Efficiency, http://repositories.cdlib.org/lbnl/LBNL-363_E/, August 2007.
- [53] Christopher Clark, Keir Fraser, Steven Hand, Jacob Gorm Hansen, Eric Jul, Christian Limpach, Ian Pratt, Andrew Warfield. Live Migration of Virtual Machines, In Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation - Volume 2 (NSDI'05), Vol. 2. USENIX Association, Berkeley, CA, USA, 273-286.
- [54] David Chisnall. The Definitive Guide to Xen Hypervisor, Prentice, 2004.
- [55] Pradeep Padala, Kang G Shin, Xiaoyun Zhu, Mustafa Uysal, Zhikui Wang, Sharad Singhal, Arif Merchant, Kenneth Salem. Adaptive Control of Virtualized Resources in Utility Computing Environments, Pro-

- ceedings of the 2nd ACM SIGOPS/EuroSys European Conference on Computer Systems 2007, March 21-23, 2007, Lisbon, Portugal.
- [56] Pradeep Padala, Kai-Yuan Hou, Kang G Shin, Xiaoyun Zhu, Mustafa Uysal, Zhikui Wang, Sharad Singhal, Arif Merchant. Automated Control of Multiple Virtualized Resources, Proceedings of the 4th ACM European conference on Computer systems, April 01-03, 2009, Nuremberg, Germany.
- [57] Ajay Gulati, Arif Merchant, Mustafa Uysal, Peter J Varman. Efficient and Adaptive Proportional Share I/O Scheduling, Technical Reports of HP Labs, 2007.
- [58] Carl A. Waldspurger, Memory Resource Management in VMware ESX Server, Proceedings of the 5th symposium on Operating systems design and implementation, 2002.
- [59] Dan Magenheimer, Chris Mason, Dave McCracken, Kurt Hackel. Transcendent Memory and Linux, Proceedings of the 2009 Linux Symposium, 2009.
- [60] Paul Barham, Boris Dragovic, Keir Fraser, Steven Hand, Tim Harris, Alex Ho, Rolf Neugebauer, Ian Pratt, Andrew Warfield. Xen and the Art of Virtualization, Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles, October 19-22, 2003, Bolton Landing, NY, USA.
- [61] Seelam, Seetharami R, Patricia J Teller. Virtual I/O Scheduler: a Scheduler of Schedulers for Performance Virtualization, Proceedings of the 3rd International Conference on Virtual Execution Environments. ACM, 2007.
- [62] Kesavan, Mukil, Ada Gavrilovska, Karsten Schwan. Differential Virtual Time (DVT) : Rethinking I/O Service Differentiation for Virtual Machines , Proceedings of the 1st ACM symposium on Cloud computing, ACM, 2010.
- [63] Mark Stillwell, David Schanzenbach, Frederic Vivien, Henri Casanova, Resource Allocation Algorithms for Virtualized Service Hosting Platforms, Journal of Parallel and Distributed Computing, Volume 70, Issue 9, September 2010, Pages 962-974, ISSN 743-7315.
- [64] Kai Shen, Hong Tang, Tao Yang, Lingkun Chu. Integrated Resource Management for Cluster-Based Internet Services, ACM SIGOPS Operating Systems Review, v. 36 n. SI, Winter 2002.
- [65] William Leinberger, George Karypis, Vipin Kumar. Multi- Capacity Bin Packing Algorithms with Applications to Job Scheduling under Multiple Constraints, Proceedings of the 1999 International Conference on Parallel Processing, p. 404, September 21-24, 1999.
- [66] VMotion, <http://www.vmware.com/products/vmotion/overview.html>.
- [67] Wood T, Shenoy P, Arun. Black-Box and Gray-Box Strategies for Virtual Machine Migration, In Proceedings of the ACM Symposium on Networked Systems Design and Implementation (NSDI'07).
- [68] Ripal Nathuji, Karsten Schwan. VirtualPower: Coordinated Power Management in Virtualized Enterprise Systems, Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles, October 14-17, 2007, Stevenson, Washington, USA.
- [69] A Karve, T Kimbrel, G Pacifici, M Spreitzer, M Steinder, M Sviridenko, A Tantawi. Dynamic Placement for Clustered Web Applications, Proceedings of the 15th international conference on World Wide Web, May 23-26, 2006, Edinburgh, Scotland.
- [70] Jeffrey S Chase, Darrell C Anderson, Prachi N Thakar, Amin M Vahdat, Ronald P Doyle. Managing Energy and Server Resources in Hosting Centers, Proceedings of the eighteenth ACM symposium on operating systems principles 2001.

- [71] Chunqiang Tang, Małgorzata Steinder, Michael Spreitzer, Giovanni Pacifici. A Scalable Application Placement Controller for EnterPrise Data Centers, In Proceedings of the 16th international conference on World Wide Web (WWW '07). ACM, New York, NY, USA, 331-340.
- [72] Gong Chen, Wenbo He, Jie Liu, Suman Nath, Leonidas Rigas, Lin Xiao, Feng Zhao. Energy-Aware Server Provisioning and Load Dispatching for Connection-Intensive Internet Services, Proceedings of the 5th USENIX Symposium on Networked Systems Design and Implementation , 2008.
- [73] 一种虚拟机调度方法, 发明人: 肖臻、宋维佳、陈琪、罗海鹏、陈曙威, 专利号: 200910238364.2.
- [74] Zhen Xiao, Weijia Song, Qi Chen. Dynamic Resource Allocation Using Virtual Machines for Cloud Computing Environment, Parallel and Distributed Systems, IEEE Transactions on, vol. 24, no. 6, pp. 1107-1117, June 2013.
- [75] Zhen Xiao, Qi Chen, Haipeng Luo. Automatic Scaling of Internet Applications for Cloud Computing Services, To appear in IEEE Transactions on Computers.
- [76] Weiming Zhao, Zhenlin Wang, Yingwei Luo. Dynamic Memory Balancing for Virtual Machines, ACM SIGOPS Operating Systems Review 43. 3 (2009) : 37-47.
- [77] Haogang Chen, Xiaolin Wang, Zhenlin Wang, Xiang Wen, Xinxin Jin, Yingwei Luo, Xiaoming Li. REMOCA: Hypervisor Remote Disk Cache, Parallel and Distributed Processing with Applications, 2009 IEEE International Symposium on, pp. 161-169, 10-12 Aug. 2009.
- [78] Livio Soares, Michael Stumm. FlexSC: Flexible System Call Scheduling with Exception-Less System Calls, In Proceedings of the 9th USENIX conference on Operating systems design and implementation (OSDI '10). USENIX Association, Berkeley, CA, USA.
- [79] David Wentzlaff, Anant Agarwal. Factored Operating Systems (fos) : the Case for a Scalable Operating System for Multicores, SIGOPS Oper. Syst. Rev. 43, 2 (April 2009) , 76-85.
- [80] Silas Boyd-Wickizer, Haibo Chen, Rong Chen, Yandong Mao, Frans Kaashoek, Robert Morris, Aleksey Pesterev, Lex Stein, Ming Wu, Yuehua Dai, Yang Zhang, Zheng Zhang. Corey: an Operating System for Many Cores, Proceedings of the 8th USENIX Conference on Operating Systems Design and Implementation , p. 43-57 , December 08-10 , 2008, San Diego, California.
- [81] Silas Boyd-Wickizer, Austin T Clements, Yandong Mao, Aleksey Pesterev, M Frans Kaashoek, Robert Morris, Nickolai Zeldovich. An Analysis of Linux Scalability to Many Cores, In Proceedings of the 9th USENIX conference on Operating systems design and implementation (OSDI '10) . USENIX Association, Berkeley, CA, USA.
- [82] IOMMU, <http://en.wikipedia.org/wiki/IOMMU>.
- [83] Gartner DataQuest Forecast on Public Cloud Services DocID G200833 , June 2 , 2010.
- [84] Novell Inc, Survey on Cloud Computing, <http://www.novell.com/nes/press/novell-survey-reveals-wide-spread-and-accelerating-enterprise-adoption-of-private-clouds>.
- [85] Jay H, Mark N. Assessing the Security Risks of Cloud Computing, <http://www.gartner.com/DisplayDocument?id=685308>.
- [86] Craig Gentry. Fully Homomorphic Encryption Using Ideal Lattices, Proceedings of the 41st annual ACM symposium on theory of computing, 2009.
- [87] F Zhang, J Chen, H Chen, B Zhang. CloudVisor: Retrofitting Protection of Virtual Machines in Multi-tenant Cloud with Nested Virtualization, Proceedings of the Twenty-Third ACM Symposium on Operating Sys-

- tems Principles, 2011.
- [88] Yubin Xia, Yutao Liu, Haibo Chen, Binyu Zang. CFIMon: Detecting Violation of Control Flow Integrity using Performance Counters, The Dependable Computing and Communications Symposium (DCCS), 2012.
- [89] Chen H, Chen J, Mao W, Yan F. Daonity- Grid Security from Two Levels of Virtualization. Inf. Secur. Tech. Rep. , 2007 123-128.
- [90] Chen F, Zhang F, Chen C, Yang Z, Chen R, Zang B, chung Yew P, Mao W. Tamper-Resistant Execution in an Untrusted Operating System Using a Virtual Machine monitor, Technical Report FDUPPITR-2007-0801, parallel Processing Institute, Fudan University, 2007.
- [91] 陈海波. 云计算平台可信性增强技术的研究(学位论文), 复旦大学, 2008.
- [92] 张逢皓, 陈进, 陈海波, 瞿斌宇. 云计算中的数据隐私性保护欲自我销毁[J]. 计算机研究与发展, 2010. 9.
- [93] Haibo Chen, Xi Wu, Liwei Yuan, Binyu Zang, Pen-chung Yew, Frederic T Chong. From Speculation to Security: Practical and Efficient Information Flow Tracking Using Speculative Hardware, In the proceedings of 35th International symposium on Computer Architecture, 2008.
- [94] Jichen Shi, Xiang Song, Haibo Chen, Binyu Zang, Limiting Cache-based Side-Channel in Multi-tenant Cloud using Dynamic Page Coloring, the 7th Workshop on Hot Topics in System Dependability 2011 (HotDep'11).
- [95] Hanjun Gao, Lina Wang, Wei Liu, Yang Peng, Hao Zhang. Preventing Secret Data Leakage from Foreign Mappings in Virtual Machines, SecureComm 2011 : 436-445.
- [96] Rongwei Yu, Fan Yin, Jin Ke, Lina Wang. Behavior-based Attestation of Policy Enforcement among Trusted Virtual Domains, JNW 5(6) : 642-649 (2010).
- [97] Yubin Xia, Yutao Liu, Haibo Chen. Architecture Support for Guest- Transparent VM Protection from Untrusted Hypervisor and Physical Attacks, In Proceedings of 2013 Internal Symposium on High Performance Computer Architecture (HPCA 2013), February, 2013.
- [98] 李小庆, 赵晓东, 曾庆凯. 基于硬件虚拟化的单向隔离执行模型[J]. 软件学报, 23 (8) : 2207-2222.
- [99] 项国富, 金海, 邹德清, 陈学广. 基于虚拟化的安全监控[J]. 软件学报, 2012, 23 (8) : 2173-2187.
- [100] 李博, 沃天宇, 胡春明, 李建欣, 王颖, 怀进鹏. 基于 VMM 的操作系统隐藏对象关联检测技术 [J]. 软件学报, 2013, 24(2) : 405-420.
- [101] 杨峰, 姜辉, 诸葛建伟, 段海新. 虚拟机环境检测方法研究综述[J]. 小型微型计算机系统, 2012 年 8 期.
- [102] Haikun Liu, Hai Jin, Xiaofei Liao, Chen Yu, Cheng-Zhong Xu. Live Virtual Machine Migration via Asynchronous Replication and State Synchronization, IEEE Transactions on Parallel and Distributed Systems, 2011, 22(12) : 1986-1999.
- [103] Gang Chen, Hai Jin, Deqin Zou, Bing Bing Zhou, Weizhong Qiang, Gang Hu. SHelp: Automatic Self-healing for Multiple Application Instances in a Virtual Machine Environment, Proceedings of the 2010 IEEE International Conference on Cluster Computing, 2010, 97-106.
- [104] Gang Chen, Hai Jin, Deqing Zou, Bing Bing Zhou, Zhenkai Liang, Weide Zheng, Xuanhua Shi. SafeStack: Automatically Patching Stack-based Buffer Overflow Vulnerabilities, To be appeared in

- IEEE Transaction on Dependable and Secure Computing, 2013.
- [105] Weiqi Dai, T Paul Parker, Hai Jin, Shouhuai Xu. Enhancing Data Trustworthiness via Assured Digital Signing, IEEE Transactions on Dependable and Secure Computing (TDSC), Vol. 9, No. 6, pp. 838-851, 2012.
- [106] Guofu Xiang, Hai Jin, Deqing Zou, Xinwen Zhang, Sha Wen, Feng Zhao. VMDriver: A Driver-based Monitoring Mechanism for Virtualization, Proceedings of the 29th International Symposium on Reliable Distributed Systems (SRDS 2010), New Delhi, India, November 1-4 2010, pp. 72-81.
- [107] Deqing Zou, Wenrong Zhang, Weizhong Qiang, Hai Jin, Kan Hu. Design and Implementation of a Trusted Monitoring Framework for Cloud Platforms, Future Generation Computer Systems (FGCS), doi: 10.1016/j.future.2012.12.020.
- [108] McCune J, Parno B, Perrig A, Reiter M, Isozaki H. Flicker: An Execution Infrastructure for TCB Minimization, Proceedings of the 3rd ACM SIGOPS/EuroSys European Conference on Computer Systems 2008. ACM, 2008 315-328.
- [109] Rutkowska. Introducing Blue Pill, The official blog of the invisible things, 2006.
- [110] Novell Inc, Survey on cloud computing, <http://www.novell.com/nes/press/novell-survey-reveals-wide-spread-and-accelerating-enterprise-adoption-of-private-clouds>.
- [111] http://voices.washingtonpost.com/securityfix/2007/11/salesforcecom_acknowledges_dat.html.
- [112] <http://www.techspot.com/news/40280-google-fired-employees-for-breaching-user-privacy.html>.
- [113] IDC, It cloud services user survey, pt. 2: Top benefits and challenges, <http://blogs.idc.com/ie/?p=210>, 2008.
- [114] Cifuentes C, Gough K, Decomilation of Binary Programs, Software Practice and Experience, 1995, 811-829.
- [115] Collberg C, Thomborson C, Low D. A Taxonomy of Obfuscating Transformations, University of Auckland Technical Report, 1997.
- [116] Collberg C, Thomborson C, Low D. Manufacturing Cheap, Resilient, and Stealthy Opaque Constructs, Proceedings of POPL, 1998, 184-196.
- [117] Myles G, Collberg C. Software Watermarking via Opaque Predicates: Implementation, Analysis, and Attacks, Electronic Commerce Research, 2006, 6(2) : 155-171.
- [118] Udupa S, Debray S, Madou M. Deobfuscation: Reverse Engineering Obfuscated Code, Proceedings of Working Conference on Reverse Engineering, 2005, 45-54.
- [119] Popov I, Debray S, Andrews G. Binary Obfuscation Using Signals, Proceedings of Usenix Security Symposium, 2007, 275-290.
- [120] Wang C, Davidson J, Hill J, Knight J. Protection of Software-Based Survivability Mechanisms, Proceedings of DSN, 2001.
- [121] Cloud Security Alliance, www.cloudsecurityalliance.org.
- [122] Lee G. The Challenges of Securing the Virtualized Environment, IEEE Computer Society, 2012.
- [123] IBM Homomorphic Encryption research page, http://domino.research.ibm.com/comm/research_projects.nsf/pages/security.homoenc.html.
- [124] David L, Chandramohan T, Mark M, Patrick L, Dan B, John M, Mark H. Architectural Support for Copy and Tamper Resistant Software, ACM SIGPLAN Notices Homepage archive Volume 35 Issue 11,

- Nov. 2000.
- [125] Trusted Computing Group, Trusted Platform module, <http://www.trustedcomputinggroup.org/>, 2010.
 - [126] G Klein, K Elphinstone, G Heiser, J Andronick, D Cock, P Derrin, D Elkaduwe, K Engelhardt, R Kolarski, M Norrish, T Sewell, H Tuch, S Winwood. seL4: Formal Verification of an OS Kernel, ACM Symposium on Operating System Principle, 2009.
 - [127] R Sailer, X Zhang, T Jaeger, L v Doorn. Design and Implementation of a TCG-based Integrity Measurement Architecture, In Proc. Of the USENIX security symposium, August 2004.
 - [128] E G Sirer, W Bruijin, P Reynolds, A Shieh, K Walsh, D Willianms, F B Schneider. Logical Attestation: An Authorization Architecutre for Trustworthy Computing, ACM Symposium on Operating System Principle, 2011.
 - [129] F B Schneider, K Walsh, E G Sirer. Nexus Authorization Logic: Design Rationale and Applications, ACM Transactions on Information and System Security, 14(1), May 2011.
 - [130] S Ghemawat, H Gobioff, S Leung. The Google File System, Proceedings of the nineteenth ACM symposium on Operating systems principles, 2003.
 - [131] P Colp, M Nanavati, J Zhu, W Aiello, G Coker, T Deegan, P Loscocco, A Warfield. Breaking Up is Hard to Do: Security and Functionality in a Commodity Hypervisor, ACM Symposium on Operating System Principle, 2011.
 - [132] J B Leners, H Wu, W. Hung, M K Aguilera, M Walfish. Detecting Failures in Distributed Systems with the FALCON Spy Network, ACM Symposium on Operating System Principle, 2011.
 - [133] M Le, Y Tamir. Using Virtualization to Validate Fault-Tolerant Distributed Systems, Proceedings of the 7th ACM SIGPLAN/SIGOPS international conference on Virtual execution environments, 2011.
 - [134] Taesoo Kim, Xi Wang, Nickolai Zeldovich, M Frans Kaashoek. Intrusion Recovery Using Selective Re-execution, In Proceedings of the 9th Symposium on Operating Systems Design and Implementation (OSDI 2010) , Vancouver, Canada, October 2010.
 - [135] Ramesh Chandra, Taesoo Kim, Meelap Shah, Neha Narula, Nickolai Zeldovich. Intrusion Recovery for Database-backed Web Applications, In Proceedings of the 23rd ACM Symposium on Operating Systems Principles, October 2011.
 - [136] Qin F, Wang C, Li Z, Kim H, Zhou Y, Wu Y. LIFT: A Low-Overhead Practical Information Flow Tracking System for Detecting Security Attacks, Proceedings of Micro, 2006 135-148.
 - [137] Crandall J, Chong F. Minos: Control Data Attack Prevention Orthogonal to Memory Model, Procedding of Micro, 2004, 221-232.
 - [138] Myers A, Liskow B. Protecting Privacy Using the Decentralized Label Model, ACM Transactions on Software Engineering and Methodology, 2000 410-442.
 - [139] Efstathopoulos P, Krohn M, VanDeBogart S, Frey C, Ziegler D, Kohler E, Mazières D, Kaashoek F, Morris R. Labels and Event Processes in the Asbestos Operating System, Proceedings of the twentieth ACM symposium on Operating systems principles, ACM New York, NY, USA, 2005, 17-30.
 - [140] Nickolai Zeldovich, Silas Boyd-Wickizer, Eddie Kohler, David Mazières. Making Information Flow Explicit in HiStar, In Proceedings of the 7th Symposium on Operating Systems Design and Implementation (OSDI) , Seattle, WA, November 2006.
 - [141] Krohn M, Brodsky M, Kaashoek M, Morris R. Information Flow Control for Standard OS Abstractions,

- Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles, ACM Press New York, NY, USA, 2007 321-334.
- [142] Venkataramani G, Doudalis I, Solihin Y, Prvulovic M. FlexiTaint: Programmable Architectural Support for Efficient Dynamic Taint Propagation, Proceedings of HPCA, 2008.
- [143] Erik Bosman, Asia Slowinska, Herbert Bos. Minemu: The World's Fastest Taint Tracker, Recent Advances in Intrusion Detection, 2011.
- [144] <http://aws.amazon.com/contact-us/ec2-email-limit-request>.
- [145] Xiaoxin Chen, Tal Garfinkel, E Christopher Lewis, Pratap Subrahmanyam, Carl A. Waldspurger, Dan Boneh, Jeffrey DwoSkin, Dan R K Ports, Overshadow: a Virtualization-Based Approach to Retrofitting Protection in Commodity Operating Systems, Proceedings of the 13th international conference on Architectural support for programming languages and operating systems, 2008.
- [146] Garfinkel T, Pfaff B, Chow J, Rosenblum M, Boneh D. Terra: a Virtual Machine-Based Platform for Trusted Computing, ACM SIGOPS Operating Systems Review, 2003, 37(5): 193-206.
- [147] Seshadri A, Qu N. SecVisor: a Tiny Hypervisor to Provide Lifetime Kernel Code Integrity for Commodity OSes, Proceedings of twenty-first ACM SIGOPS symposium on Operating systems principles, ACM Press New York, NY, USA, 2007, 335-350.
- [148] Ahmed M Azab, Peng Ning, Zhi Wang, Xuxian Jiang, Xiaolan Zhang, Nathan C Skalsky. HyperSentry: Enabling Stealthy in-context Measurement of Hypervisor Integrity, Proceedings of the 17th ACM Conference on Computer and Communications security, 2010.
- [149] Wang Z, Jiang X. Hypersafe: A Lightweight Approach to Provide Lifetime Hypervisor Control-Flow Integrity, IEEE Symposium on Security and Privacy, IEEE, 2010 380-395.
- [150] David Lie, Chandramohan A. Thekkath, Mark Horowitz. Implementing an Untrusted Operating System on Trusted Hardware, Proceedings of the nineteenth ACM symposium on Operating systems principles, 2003.
- [151] Suh G, Clarke D, Gassend B, van Dijk M, Devadas S. AEGIS: Architecture for Tamper-Evident and Tamper-Resistant Processing, Proceedings of the 17th Annual International Conference on Supercomputing, ACM New York, NY, USA, 2003, 160-171.
- [152] R Lee, P Kwan, J McGregor, J DwoSkin, Z Wang. Architecture for Protecting Critical Secrets in Microprocessors, In International Symposium on Computer Architecture (ISCA), pages 2-13, IEEE, 2005.
- [153] D Champagne, R Lee. Scalable Architectural Support for Trusted Software, In International Symposium on High Performance Computer Architecture (HPCA), 2010, pp. 1-12.
- [154] Jakub Szefer, Ruby B Lee. Architectural Support for Hypervisor-Secure Virtualization, Proceedings of the International Conference on Architectural Support for Programming Languages and Operating Systems (AS-PLoS), March 2012.
- [155] Keller E, Szefer J, Rexford J, Lee R. NoHype: Virtualized Cloud Infrastructure Without the Virtualization, Proceedings of the 37th annual international symposium on Computer architecture, ACM, 2010 350-361.
- [156] Common Vulnerabilities and exposures, <http://cve.mitre.org>.
- [157] Kernel base Virtual Machine, http://www.linux-kvm.org/page/Main_Page.
- [158] Extended Page Table, <http://www.intel.com/technology/>.
- [159] T. Healthcare, TC3 Healthcare, <http://www.tc3.health.com>, 2008.

- [160] Raluca Ada Popa, Catherine M S Redfield, Nickolai Zeldovich, Hari Balakrishnan. CryptDB: Protecting Confidentiality with Encrypted Query Processing, ACM Symposium on Operating System Principle, 2011.
- [161] Thomas Ristenpart, Eran Tromer, Hovav Shacham, Stefan Savage. Hey, You Get Off of My Cloud: Exploring Information Leakage in Third- Party Compute Clouds, Proceedings of the 16th ACM conference on Computer and communications security, 2009.
- [162] Kuniyasu Suzuki, Kengo Iijima, Toshiki Yagi, Cyrille Artho. Memory Deduplication as a Threat to the Guest OS, Proceedings of the Fourth European Workshop on System Security, 2011.
- [163] Geambasu R, Kohno T, Levy A, Levy H. Vanish: Increasing Data Privacy with Self- Destructing Data, Proceedings of the USENIX Security Symposium, 2009.
- [164] Yinqian Zhang, Ari Juels, Alina Oprea, Michael K Reiter. HomeAlone: Co- Residency Detection in the Cloud via Side- Channel Analysis, IEEE Symposium on Security and Privacy, 2011.
- [165] Rutkowska. Introducing Blue Pill, The Official Blog of the Invisible Things, 2006.
- [166] L Singaravelu, C pu, H Hautig, C Helmuth. Reducing TCB Complexity for Security- Sensitive Applications: Three Case Studies, In proc. Eurosyst, 2006.
- [167] Vallee G, Naughton T, Ong H, Scott S L. Checkpoint/restart of Virtual Machines Based on Xen, Proceedings of the High Availability and Performance Computing Workshop (HAPCW 2006), Santa Fe, New Mexico, USA, 2006.
- [168] Google Apps 拖累 Gmail 又当机，中断持续将近 22 小时，<http://www.cnetnews.com.cn/2009/0313/1355592.shtml>.
- [169] Amazon S3 存储中断，影响企业营运，<http://news.csdn.net/n/20080218/113649.html>.
- [170] Mingyuan Xia, Miao Yu, Qian Lin, Zhengwei Qi, Haibing Guan. Enhanced Privilege Separation for Commodity Software on Virtualized Platform, ICPADS 2010.
- [171] Miao Yu, Zhengwei Qi, Qian Lin, Xianming Zhong, Bingyu Li, Haibing Guan. Vis: Virtualization Enhanced Live Forensics Acquisition for Native System, Digital Investigation 9(1): 22-33 (2012).
- [172] Jun Zhu, Wei Dong, Zhefu Jiang, Xiaogang Shi, Zhen Xiao, Xiaoming Li. Improving the Performance of Hypervisor- Based Fault Tolerance, the 2010 IEEE International Symposium on Parallel and Distributed Processing, IPDPS 2010.
- [173] Jun Zhu, Zhefu Jiang, Zhen Xiao, Xiaoming Li. Optimizing the Performance of Virtual Machine Synchronization for Fault Tolerance, IEEE Transactions on Computers, 2011.
- [174] Jun Zhu, Zhefu Jiang, Zhen Xiao. Twinkle: A Fast Resource Provisioning Mechanism for Internet Services, in Proc. of IEEE Infocom, Shanghai, China, April 2011.

作者简介

本文由罗英伟（北京大学）、廖小飞（华中科技大学）、陈海波（上海交通大学）负责组织撰写及整理，执笔人包括：罗英伟、肖臻、汪小林、宋维佳、陈琪（北京大学），廖小飞、郑龙、邵志远、邹德清（华中科技大学），陈海波、夏虞斌、戚正伟（上海交通大学），张彬彬（云南大学）。

罗英伟 北京大学信息科学技术学院教授、博士生导师，中国计算机学会高级会员，系统软件专委委员。主要从事计算系统虚拟化、云计算、地理信息系统等方面的研究。E-mail: lyw@ pku. edu. cn。



廖小飞 华中科技大学计算机学院教授、博士生导师，中国计算机学会高级会员，系统软件专委委员。主要从事计算系统虚拟化、对等计算、集群计算、流媒体服务等方面的研究。E-mail: xfliao@ hust. edu. cn。



陈海波 上海交通大学软件学院教授、博士生导师，中国计算机学会高级会员，系统软件专委委员。主要从事计算系统虚拟化、计算系统性能与安全等方面的研究。E-mail: haibochen@ sjtu. edu. cn。



社会化协同计算研究进展报告

CCF 协同计算专委会

摘要

社会化协同计算是根据社会科学的计算理论和方法，以计算技术为工具，研究人们社会关系、结构、层次、社会活动及其作用和效应的计算理论与方法。近年来随着以微博、社交网站为代表的社会媒体的迅速发展，形成了规模巨大的虚拟社会，对人们日常生活中的沟通、协作、学习、维系关系、交换信息等活动产生了巨大影响；同时，也为研究大规模社会环境中的人际关系、人群结构、信息传播和群体行为等问题提供了方便。本文将针对社会化协同计算领域的几个关键问题，包括面向协同工作的社会网络分析、群体思维与社会智能、情感挖掘与人人协同以及移动协同环境等，综述近期的国内外研究成果和研究现状，介绍社会化协同系统研究和开发方面的国内研究团队和研究成果，探讨未来协同计算发展趋势。

关键词：协同计算，社会智能，社会网络，情感计算

Abstract

Collaboratively social computing is a general term for an area of computer science that is concerned with the intersection of social behavior and computational systems. It has become more widely known because of its relationship to a number of recent trends. These include the growing popularity of social software and Web 2.0, increased academic interest in social network analysis, the rise as a viable method of production, and a growing conviction that all of this can have a profound impact on daily life. This paper reviews the key problems in this area that include social network analysis for collaborative target, collective intelligence, affective tendency mining, collaborative mobile computing etc. We survey the latest related works and introduce some research groups and their work in China. Some discussions are presented on future directions.

Keywords: cooperative computing, social intelligence, social network, affective computing

1 引言

社会化协同计算是根据社会科学的计算理论和方法，以计算技术为工具，研究人们社会关系、结构、层次、社会活动及其作用和效应的计算理论和方法。现实社会中的数据都是隐性的，获取、分析大规模的社会数据是一件很困难的事情，近年来随着以微博、社交网站为代表的社会媒体的迅速发展，虚拟社区大量出现，组成了规模巨大的虚拟社会，从而为研究大规模社会环境中的人际关系、人群结构、信息传播和群体行为等问题提供了方便。社会计算不仅仅是面向社会的，而且还是基于社会的——群体智慧，即利

用群体协作的方式解决问题。因此，随着网络沟通工具的普及、用户共享内容的涌现以及在线社交网络的繁荣，社会化协同计算对人们日常生活中的沟通、协作、学习、维系关系、交换信息等活动将产生深远影响。

本文将针对社会化协同计算领域的几个关键问题，综述近期的国内外研究成果。具体组织为：第2节分析当前协同计算的国际研究现状和发展趋势，包括面向协同工作的社会网络分析、群体思维与社会智能、情感挖掘与人人协同以及移动协同环境5个方面；第3节从国内社会化协同系统的开发与应用现状以及国内的主要研究团队和研究成果两个方面介绍国内协同计算的研究进展。

2 国际研究现状

2.1 面向协同工作的社会网络分析

社会网络正在从个人之间的交流工具演变为协同人们行为的重要信息来源。人们身处真实和虚拟的社会中，通过社会网络感知周边环境的状态变化，自觉/非自觉地形成某种意义上的行为一致性（默契或者对抗）。

社会网络分析的一般过程是先从应用环境中抽取信息，构建社会网络模型，进而采用统计学方法或图论方法，洞察社会网络的结构特性或信息流向。最后把这些特性映射回应用环境中活动的参加者之间的关系，发现问题并探索解决问题的方法，或者准确有效地促成或控制信息的传播。

新应用领域和应用模式层出不穷，导致了新型网络和网络新特性的出现。研究者探索了社会网络分析的创新方法。比如，多关系网络、多重网络、基于位置的网络、事件驱动的网络、过度流行结点、富计算模型等。

社会网络分析方面的多数研究成果重点关注于分析方法能否得到有趣的结论，而未关注大规模社会网络的分析效率，因此，其对在线动态大规模社会网络分析的实时性支持能力有限。社会网络以图为底层数据结构，已有实验表明，对于6467顶点（用户）和26467条边（关系）的图边介数串行算法在单CPU上耗时约10 s。按全源最短路径APSP（All-Pairs Shortest Paths）算法复杂度 $O(E * V)$ 估计，拥有1~2亿用户的新浪微博的社区分析将超过100年！因此，针对大规模网络的分析方法开始受到研究者的关注。

2.1.1 协同社会网络构建和分析

社会网络中的结点是社会中实体的抽象，链接则表示实体之间的某种关系。研究者通过各种手段从网络中提取实体和链接的抽象，从而构建以数学和图论为基础的社会网络模型。比如，Xia Hu 2012首次分析了公布的美国阿拉斯加州州长Sarah Palin的邮件通信集，通过邮件的发送和接收关系建立通信关系网络^[1]。Xiao-Yong Wei通过分析课堂照

片，自动识别学生和学生的位置，基于多媒体技术进行学生关系的数据采集，进一步建立学生之间的社会网络^[2]。在社会网络分析和统计的基础上，发现学生中的互助学习（co-learning）模式，为教学效果提供量化手段。Pengyan Yao 研究了维基百科中用户协同创建内容的方式，提出了一种以用户协同生成服务（Collaborative User-Generated Services, Co-UGS）的模式探索社会网络中的用户行为。通过分析用户对服务的反应，将用户分类^[3]。Elnaz Davoodi 首先在互联网上爬取研究人员的基于内容的个人档案，并用维基百科中的背景知识丰富个人档案内容的语义。然后根据个人档案计算研究者的相似性，构建社会网络，提出了一种针对学术研究者的基于语义的社会网络构建框架。该方法进而检测研究人员社区，并度量特征向量的中心度，识别社区中的典型代表。该成果可用于在分发学术会议的评审论文时选取正确的研究者^[4]。Sunju Oh 针对社交网络抽取时缺少语义信息的现状，提出了一种基于本体的社会网络抽取方法，能够计算实体之间的联系强度^[5]。

社会网络分析法是对社会网络的关系结构及其属性加以分析的一套规范和方法，主要分析不同社会单位（个体、群组或社会）所构成的社会关系的结构及其属性。社会网络分析法作为一种定量的分析方法，其测度指标主要有结点、连接、密度、中心度、中心势、成分、核、派系、位置、角色、聚类以及维度等。现代研究者根据网络特点，在传统方法的基础上进行了研究和创新。比如，Yukihiro 分析了社会网络中关系的量化对传统的中心度（结点度、边介数、亲密度、特征向量中心度）度量结果的影响。针对模拟生成的图数据实验结果表明，针对中心度的边的量化对发现最重要结点的效果不明显，反而，边被量化为 5~8 级对中心结点的发现和宽范围的结点排序是必要的^[6]。Juliano 以 Twitter 用户对历史话题发表的观点倾向标记该用户，建立社交网络。然后用关系挖掘的办法，推断对特定话题未发表观点的用户可能的观点倾向。这种方法在 Twitter 用户群中得到了证实，在党派选择方面很有应用前途^[7]。

社会网络分析得到的测度指标定量反映了目标网络的特性，但其抽象性仍不能直观呈现该群组的整体性。可视化的社群图有效弥补了这一缺憾，通过路径、距离、维度等属性，帮助使用者直观理解群组结构。Guo Zhang 为解释社区和在线社区的概念、问题和 Web 上的操作化，讨论了社区特性的四维框架：形状、结构、上下文、经验。针对社区分析的方法和量化测量被集成为工具，用于探索社区的本质和功能^[8]。Thiago Rúbio 提出了一种基于人工神经网络和模糊逻辑的社会网络行为分析方法^[9]，使用 SOM (Self Organized Maps) 模型，把 Twitter 用户分为 4 类，即无印象用户（Unimpressive User）；活跃用户跟随者（Desired User: Follower）；活跃用户：跟随者与发布者（Desired User: Follower and Publisher）；活跃用户：发布者（Desired User: Publisher）。Alan Said 等人^[10]为产生更丰富的推荐结果，除在常见的协同过滤推荐系统中用最邻近方法推荐候选项之外，提出用 k -Furthest 邻居识别最不可能的候选项。该方法一方面根据训练集/测试集评估，另一方面通过在线的用户学习识别用户的感知质量。测试结果表明，该方法的推荐结果好于传统方法，并且反面感知结果也与用户学习结果没有明显差别。

2.1.2 新型社会网络和网络新特性

互联网信息日益丰富，研究者获得研究目标之间的关系信息也更加容易，并且现代计算能力迅猛提高，研究者更加关注研究目标构成的社会网络的复杂细节，从而建立蕴含更丰富信息的社会网络模型，得到更细腻的度量和刻画。比如，Jian Chen^[11]认为网络结点之间不应该只有一种关系，他提出了一种基于多关系网络分析（Multi-Relational analysis）的推荐算法，用于减轻数据稀少问题和冷启动（Cold-start）问题。Jun-Ki Min^[12]也认为大多社会网络中把人们之间的各种关系归结为“朋友”而出现了信息损失，因此尝试了一种富计算模型。该模型利用移动电话中的呼叫记录和短信记录，按照生活侧面（如家庭、工作、社会等）对联系分类。其实现方法是先抽取联系的多种特征（如通信强度、规律性、媒介、时间倾向等），再利用机器学习技术对联系分类。实验针对40个用户展开，分类正确率达到90.5%，且实验结果表明，通话时长、频道选择、通话时间等是最相关的联系特征。

现实中，人们同时处于线下的真实社会和线上的虚拟社会中，同时构成了两个社会网络。针对这种异构网络，Xingjie Liu 研究了一种事件驱动的新型社会网络（Event-Based Social Network，EBSN），同时关联了线上和线下的社交世界^[13]。根据 Meetup 社交网络的数据分析，发现其存在结点度的重尾分布和强局部社交等独特性质。随后对这种异构的 EBSN 进行社区检测和信息流特性分析，发现其社区更具凝聚性，且同时考虑线上/下交互的社区信息扩散模型更具预测力。

手持设备的普及使用户的位置也成为社会网络中的重要信息。Miltiadis 提出了一种基于位置的在线社会网络的演化分析模型。作者通过跟踪目标网络4个月的发展演化数据，分析用户访问的地点和建立关系的时间，发现存在共同朋友的用户之间，以及访问过相同地点的用户之间更容易建立关系^[14]。

在虚拟社会中，催生了许多知名人物。知名人物的关注者（粉丝）数量远远大于普通人，成为社会网络中的流行结点（Popular）。Youngchul 针对现有的知名 LDA（Latent Dirichlet Allocation）不能处理社会网络中的流行结点入边过多问题，提出了两阶段标注（two-step labeling）和阈值噪声过滤（threshold noise filtering）等改进措施，最小化流行结点引入的噪声^[15]。在 Twitter 数据集上的实验表明，这些措施提高了标注质量。

隐私保护成为社交网络中的另一个被高度关注的问题。针对网络用户暴露的多重社会关系，Saajan Shridhar 开发了一套称为 SNAIR 的基于社会网络分析的个性化推荐框架^[16]。该框架从用户使用的多重社会网络应用中挖掘用户的位置、兴趣、工作领域、性别、朋友、交互本质等元素，以保护隐私的方式形成用户个性化推荐。

2.1.3 大规模社会网络的分析方法

对于一般规模的社会网络，研究者可自己开发程序，统计计算网络测度指标，分析网络特性；也可以配合网络可视化分析工具（如 UCINET、PAJEK 等），针对不同的侧面

构建网络模型，人工观测网络结构特性，推理社会网络中实体之间的关系特征。

对于大规模社会网络，往往需要采取基于图论的方法（如中心度、图匹配、图模拟、最短路径），开发专门的并行程序，提高计算性能，或者构建近似网络，降低计算的复杂度。这些方法主要有：基于云的大数据并行挖掘、图模式匹配和图查询、多尺度关系预测等。

云计算在工业界和学术界的研究热度，自然成为应对大规模社会网络分析的平台。Le Yu 为应对大规模数据的分析效率，设计了一款基于云的大数据并行挖掘工具 BC-PDM (Big Cloud-Parallel Data Mining)^[17]。BC-PDM 支持并行 ETL 处理、统计分析、数据挖掘、文本挖掘和基于 Hadoop 的社会网络分析。基于 BC-PDM 能完成商务推荐、顾客社区检测和用户偏好分类三项功能。

计算需求的无止境提高，永远无法真正解决计算需求与问题规模之间的矛盾。从图论方法入手，探索解决大规模图计算问题才是根本。Wenfei Fan 针对传统的图模式匹配技术（子图同构、图模拟等）在大规模图中过度消耗计算资源问题，提出了修订图模式匹配和图查询的相关概念，以便提高社区发现的效果和效率^[18]。这些修订包括：1) 为频繁更新的社会网络提出了增量图模式匹配分析的有界性；2) 为应对大规模现实图提出了查询保持的图压缩概念，只保留特定图查询的必要信息；3) 解决分布式图的模式匹配问题，提倡局部评价技术。

建立近似网络也是应对大规模网络的有效途径之一。Donghyuk 针对社会网络分析中影响用户亲密程度的邻近关系估计问题，提出了一种多尺度关系预测方法 MSLP (Multi-Scale Link Prediction)。基本思想是以高性能方式构建低阶近似网络，解决大规模网络的关系预测的健壮性和准确性^[19]。在超过百万结点的实际网络的实验中，算法表现出了性能和可扩展性方面的优越性。

2.1.4 社会网络分析的典型应用领域

许多应用领域借用社会网络分析结果改善人们之间的协同效果。这些协作既有针对正面协作的推进，也有针对负面影响的避免。正面协作的例子有科研合作者的选取、邮件关系推测社会地位、学生之间协同学习、维基百科的协同编辑等；负面协作有恐怖行动、恶意软件等。

(1) 推荐合作者

选取正确的合作者，是成功的开端。Ge Gao 针对人们如何利用社会网络信息选取正确合作者问题，展开了基于情形的研究^[20]。参加实验的 459 人中，一组采用近似封闭模型，另一组采用非封闭模型。实验数据表明了社会网络信息中的文化差异在选取合作者时的估值方法。该研究结果为如何洞察差异化社会网络信息中的不同文化和社交网站如何支持人们在相同文化内和跨文化选取合作者提供了新的视角。Tamara Heck 针对科研人员严重依赖于与同事的合作这一事实，提出了一种通过从科研人员作品的引用关系和参考关系收集社会信息，聚类相似研究者的方法^[21]。比如，社会书签系统 Web 服务提供了

新的功能及更丰富多样的社会信息。采用正确的信息抽取方法，综合建立学术社会网络，能够提供更好的合作者推荐结果。

(2) 基于电子邮件的协作关系

电子邮件是人们最常用的通信方式。电子邮件不仅传递了协同工作的直接信息，同时也蕴含其他丰富信息，比如通信者的社会地位等。Xia Hu 通过分析美国阿拉斯加州州长 Sarah Palin 的 email 通信集，根据邮件的发送和接收关系构建了通信关系网络，并用减边技术对比了包含 Sarah Palin 与否的网络的中心度和聚类系数等特性，佐证了 Sarah Palin 在该网络中具有最高的社会地位和最重要的社会角色^[1]。

(3) 协同编辑和协同学习

基于互联网络的协同编辑与写作，在互联网内容建设方面起到了独特作用。Pengyan Yao 扩展了类似维基百科中的用户协同创建内容的成功方式，提出了一种以用户协同生成服务（Collaborative User-Generated Services, Co-UGS）的模式探索社会网络中的用户行为^[3]。通过分析用户对服务的反应，将用户分为三类：忽略用户（ignoring users）、共享用户（sharing users）和共同创新用户（co-creating users）。

学生之间的团体关系特性分析，有利于教师主动促成互助学习小组，提高教学效果。Xiao-Yong Wei 针对学生的社会关系信息提取难度较大问题，提出了一种基于多媒体技术的班级学生社会网络数据采集、建模和分析方法^[2]。该方法通过分析课堂上学生座位关系的照片，自动识别学生和学生的位置，建立社会网络。通过分析和统计，发现学生中的互助学习模式，为教学效果提供量化手段。

(4) 反恐与抵制恶意软件

通过互联网组织社会事件的社会安全问题受到全世界关注。研究者也通过社会网络分析，发现其前兆和发作路径，及早预防。Sebastián 针对反恐网站 DarkWeb portal，采用文本挖掘技术，设计了一种基于话题模型的重叠社区检测方法^[22]。重叠社区中允许一个结点同时属于多个社区，从而保留了传统社区检测中损失的可能有用信息，有利于准确预测可能出现的地区威胁。

恶意软件成为互联网上令人头疼的问题。Ting-Kai Huang 针对 FaceBook 上的 300 万用户进行了近 5 个月的观测，目的是对在线社会网络（Online Social Networks, OSN）上的恶意软件和垃圾软件的特性进行更好的理解^[23]。这些特性主要包括时空特性、底层目的和机制、欺骗用户的社会工程学陷阱等。研究结果表明，黑客常用的社会工程学陷阱主要有两种：以免费产品引诱用户，或者唤起用户的社交好奇而使恶意软件得逞。这些发现有利于为 Facebook 找到一条减少恶意软件的道路，尽管也将面临严重挑战。

2.1.5 发展趋势

大规模社会网络的高性能分析方法成为可能的重要研究方向之一。比如，图索引、图摘要、图压缩算法、分布式图模式匹配、不严格或近似图匹配算法、图重写算法等。

网络媒体资料和现实生活中的多媒体资料日益丰富，基于图像和视频媒体的社会网

络构建和分析方法，也可能成为重要发展方向之一。比如通过海量交通信息和道路信息构建的社会网络分析，可以实现对车辆的行为预测、实时追踪等，保障交通和社会安全。

对于线上线下实体的映射关系、社区依时间的演变特征分析、不同类型社团的演变特征与规律等，也有强烈的社会需求，这将成为进一步的重要研究趋势。

2.2 群体思维与社会智能

在社会信息化和经济全球化的大趋势下，随着协同技术的成熟与普及，人们越来越多的沟通、交流与协作活动通过协同计算平台和应用来展开，并逐步呈现出鲜明的社会化特征。在这个社会化交互的过程中，个体思维在这些协同平台进行汇聚，通过一定的机制形成有组织的群体思维，并在不断认知、理解、解释和解决各种社会问题的过程中，逐渐升华为体现社会价值的社会智能。

群体思维和社会智能一直是计算机支持的协同工作（CSCW）和人机交互（HCI）研究领域的重点和热点，特别是伴随着近年来新交互媒体和智能终端等的普及，涌现了一大批利用群体思维和社会智能来解决社会化协同计算问题的研究成果。通过对协同计算、社会计算和人机交互领域权威国际会议 CSCW’13 和 CHI’13 相关文献的调研，围绕众包（Crowdsourcing）、问答系统（Questions & Answers）和公民参与（Citizen Participation）等群体思维和社会智能典型应用，本文给出了国外该研究领域的研究现状、发展动态和未来趋势。

2.2.1 众包系统设计

众包是一种分布式的生产方式，通过在互联网上公布项目或问题，利用社会大众的参与来完成。各种在线众包网站的出现推动了群体智慧的发展，一方面，众包的价格低廉，甚至没有费用，全靠志愿参与，雇佣容易。另一方面参与者的多样性使得其能够胜任大规模的项目，比如人文学科的社会调研等。近几年以亚马逊在线市场（Amazon’s Mechanical Turk, MTurk）为代表的在线众包网站或社区的兴起，推动了众包的发展。

众包项目大都通过在线的方式进行发布，项目发起人与参与者之间的联系较为松散，缺乏有效沟通，项目在执行过程中得不到有效的控制和管理，因此，如何提高质量和效果是众包系统设计研究的重要方面。哈佛大学的 Komarov 等人，就线上和线下众包的效果进行了对比试验。他们将 MTurk 上招募的用户分为两组进行软件界面测试，一组在实验室内有专人的指导和监督下进行操作，另一组则通过在线的方式，通过发放指导手册，由用户自行完成实验。分析实验结果发现，尽管实验环境不同，但不论是有人指导还是自行完成，两组实验的结果基本相似。从而说明，虽然缺少有力的监督和管理，但在线众包方式并不影响一些项目的执行效果^[24]。UC Irvine 的 Irani 则指出通过众包建立起来的雇佣关系过于简单，存在很多潜在的问题^[25]。一方面，对项目发起人缺乏了解，群众可能会在不知觉的情况下参与到一些违背道德规范的事情。另一方面，对参与者的不知

情，也可能导致项目在一种违背发起人初衷的情况下进行。因此，如何加强众包双方的沟通和了解，是当前各个众包系统需要改进的地方。在工作质量方面，研究表明增强参与者间的信息交互，对提高众包的效果有一定的帮助。Kinnaird 发现向参与者提供“相同任务中其他合作者的数量”信息，确实会对参与者的积极性和工作质量产生重要的影响，但是这种影响不总是正面的。例如，当参与者感受到合作者的数量增加会最终降低工作的质量，这会让他们觉得这项工作的重要性正在降低并且降低了自己对这份工作的责任心^[26]。Huang 指出增加社会透明性可以提高众包项目的执行效率与产出^[27]。一方面，社会透明性的增加，意味着彼此熟悉，从而信息可以得到更好的分享；另一方面，社会透明性也提高了团队成员的责任感，工作不再是自己一个人的事情，要对他人负责。由于需要多人共同完成，众包中的隐私问题也越来越受到重视。邮件中的任务管理是一种典型的众包任务，但是这需要将邮件中可能存在的隐私数据分享给协助者。Kokkalis 设计了一种名为 EmailValet 的系统，该系统允许用户在邮件分享时指定具体的内容，同时也可以看到协助者对每封邮件都做了些什么，从而保护雇主的隐私和问责权^[28]。此外，众包的流行也给软件开发行业带来了冲击，O’Neill 对软件众包和外包进行了分析比较，指出众包的优势主要在三个方面：1) 价格低廉，不存在办公室租赁等方面的开销；2) 选择范围广，更容易招募到不同类型的雇员；3) 当需求变化，工作量增加时，可以轻易的增加人力予以支持。因为可以在全球范围内进行招聘，项目甚至可以 24 小时不间断的运行，不受时区限制^[29]。

2.2.2 众包质量的评估和控制

众包过程中需要对任务完成质量以及参与者的贡献度进行评估，需要通过合适的方法和机制对任务质量进行控制。FamilySearch Indexing 项目是世界上最大的众包项目之一，它雇用了成千上万的志愿者对上亿的录音材料进行转录。Hansen 等人通过实地试验的方法，比较了仲裁（Arbitration）和同行评审（Peer Review）这两种常见的质量控制机制在该项目中的有效性和效率^[30]。同行一致性评估（Peer Consistency Evaluation）是另一种质量控制方法，该方法使用之前参与者所提供的结果对当前参与者提供的结果进行评价，如果前后结果相同或相似，当前结果将被接受，否则将遭拒绝。Huang 和 Fu 的研究结果显示同行一致性评估方法能够加强人本计算系统（Human-based Computation System）的可靠性。作者还认为，由于产生标准结果的代价往往很大，因此同行一致性评估方法可以为保证人本计算系统（包括众包系统）的可靠性提供一个更加经济、有效的选择^[31]。Mashhadi 等人则对任务结果质量的相关因素进行了研究，他们分析了在 OpenStreetMap 众包项目中，地图信息的覆盖率这一质量信息与人口密度、人口流动、市中心和贫民区之间的距离等上下文信息之间的关系。作者认为，提高这些上下文信息的可感知性，不仅有利于街景地图绘制这类众包应用的任务规划，同时，对最终地图用户也很有价值，可以让他们知道哪些区域的信息是可以信赖的^[32]。在众包参与者的贡献度评估方面，Waterhouse 将信息论方法引入到了问答系统的用户贡献评价中^[33]。他们设计了一

个点式互信息（Point-wise Mutual Information）的度量值，来衡量对每个问题从贡献者那边得到的平均信息量。此外，他们使用了条件和多元互信息以捕捉不同贡献者之间因相互评价而产生的交互。作者认为，信息论方法在贡献评价方面很具潜力。

2.2.3 众包的新形式和新模型

随着智能手机的普及，通过移动终端来参与众包成为传统方式的一个很好的补充，特别是对于那些耗时短并且对移动数据有特殊需求的项目。Musttag 等人对移动众包与在线众包之间区别进行了研究。研究结果显示在移动众包市场中，大部分的项目是由那些年轻且教育背景较好的男性用户来完成的。这些人更愿意去距离较远的地方参与项目，并且可以同时承担多个项目。通过规划出行线路来节省出行的开销，降低项目的成本。其次，移动数据的采集主要通过智能手机来完成，数据的可信度相比于在线参与的方式要可靠。因此，相对于在线众包，移动众包市场应该更强调项目执行的效率与成本支出^[34]。

由于传统的众包模型通常要求将一个问题划分为一个个简单的小任务，而一个设计问题通常太过复杂而难以划分，因此，Park 等人为设计问题的众包提出了一种新的模型：CvC (Crowd vs. Crowd)。在该模型中，一个设计问题被分给多个众包小组独立完成，通过他们之间的竞争来产生出最终的设计结果。作者通过定性和定量的方法对该模型进行了评价，发现这种竞争机制不仅能够促进人们的参与热情，还能在很大程度上提高最终设计的结果的质量^[35]。

2.2.4 新的众包应用领域

除了用于数据收集、用户调查等传统领域，众包的问题解决方式还在很多新的场景中得到应用。Dow 提出利用众包方式解决大学教育脱离实际应用导致学生缺乏相关的从业知识等问题。通过对飞行员课程引入在线众包，学生可以在不同的学习阶段与飞行员、乘客或者机组成员进行必要的交互，有助于帮助他们了解今后的工作内容，对提高学生的学习积极性，增强责任感有很大的帮助^[36]。公众出行方面，虽然谷歌街景能够很大程度上满足人们的出行需求，但对于残障人士来说，街景所展现出来的信息依旧有限，比如，人行道上的斜坡、道路的宽度都会影响到轮椅的使用。Hara 在 MTurk 上招募志愿者，让他们在谷歌街景上将那些可能会影响残障人士的障碍进行标注，为特殊人群提供便利。结果显示，虽然这些人没有经过专业的培训，但所标注的数据通过一定的质量控制方法，可以达到 93% 的准确率，这为地理数据的收集提供了一种新的思路^[37]。同声传译作为实时众包的代表，已经被普遍采用。该方法将演讲的信息进行拆分，每个人只翻译其中一部分，最后把所有人的翻译结果进行汇总，可以弥补单人传译容易造成信息遗漏的不足。Lasecki 提出一种时间扭曲（Warp Time）的方法，对译员负责的部分，降低播放的速度，同时加快其他部分的播放速度，使其能在有限的时间内更好的解读演讲内容^[38]。Chilton 设计了一个众包分类系统（Cascade），将需要分类的信息在网上进行公

开，由用户参与分类，最后将用户的分类信息进行汇总。该方法的分类结果与专业人员的相比，有 80% ~ 90% 的正确率，而费用开销则要小得多^[39]。Chilana 开发了一个基于情景的软件众包帮助系统，通过搜集大量的操作日志，抽取出相关的操作信息。当用户在软件使用过程中碰到困难时，帮助系统根据分析其他用户在相似情境下的行为和选择，进行提示，提高软件的可用性和实用性^[40]。

2.2.5 问答系统

在线问答系统能够通过众多贡献者共同构建有价值的知识库。Furtado 等人根据用户贡献的数量和质量，对 5 个传统在线用户的贡献进行了分析，并为用户构建了贡献档案。此外，他们对其中最大的问答系统中用户贡献的动态性进行了分析。除了对传统问答系统的研究外，社交网络和众包也成为用户咨询问题和获取信息的新的途径。在社交网络中用户可以通过对问题的转发来获得信息和答案^[41]。Gray 等人对 Facebook 中用户对问题转发所产生的效果进行了研究，这种方式可以允许用户在同时包含强连接和弱连接的网络中寻找信息。该项研究结果表明，提问者连接社会资本的能力和在 SNS 中得到回答的有用性成正相关，然而，有用的回答基本来自于那些弱连接^[42]。Starbird 也同时指出信息技术的发展丰富了人与人之间的交互方式，为社会智能的发挥提供了更好的平台。例如，在 Twitter、Facebook 等社交网络上发布信息，通过对问题的转发，可以让更多的人接触问题，参与问题的解决^[43]。

Lee 等人对移动电话有偿问答的众包服务进行的调研显示，用户参与问答主要是从经济回报和个人兴趣爱好出发，而传统的在线问答，用户则参与问答则主要是受社会因素的影响，比如社会责任感、社会地位等。其次，有经验的用户会对回答的问题进行一个筛选，并且有一些普遍通用的答题策略或答案，而在传统在线问答中，用户对问题没有特别的偏好，并且作答也相对较快。但另一方面，从长期来看，用户在移动有偿问答中给出的答案并不会因时间的改变而变化，这也说明该方式在数据收集方面的鲁棒性^[44]。

还有研究者则为盲人这类特殊群体设计和评估新的问答系统。Brady 等人针对盲人用户设计了一个 iPhone 照片帮助软件，用户可以对其拍摄的照片在该社区中进行提问，通过他人的描述来了解照片的内容。研究者通过对用户进行访谈，发现该应用对盲人的辅助效果主要取决于照片的质量、问题回答速度以及回答内容长短这三个方面的因素^[29]。他们还对其中 191 名盲人成人进行了调研，研究了盲人使用社会网络进行问题咨询的潜在可能，同时通过日志分析了盲人是否觉得 VizWiz 是一个适合 Q&A 的在线社区。结果表明，有很大一部分盲人用户会使用社交网络，但是由于存在很高的代价，很少有盲人用户觉得适合在上面咨询问题并寻找答案^[45]。

2.2.6 公民参与

公众项目是社会智能体现的重要方式，只有提高民众的参与度，让更多的人投身公

众项目，才能最大限度的发挥社会智能的优势。Nov 从用户的角度出发，提出不同个性的人会被不同类型的 UI 所吸引，因而在线公众社会项目的界面设计以用户个性为导向，灵活多变，而不只是一个面向大众的设计^[46]。Denny 将虚拟奖励制度引入在线交互学习网站后，发现学生的积极性得到了提高，对网站的贡献也比之前要大，进而提出在在线参与项目中，根据贡献大小设置不同的等级和授予相应的称号，可以提高公民参与度^[47]。Massung 对公民参与的动机进行了分析发现，提供报酬或强制参与，虽然可以调动公民积极性，但无法形成长期的刺激，甚至还会产生排斥心理。较好的方式是通过排名的方式，在参与者之间形成良性竞争机制来激发公众的参与热情^[48]。Yoo 等人对公民参与的公共交通设计项目进行了跟踪调研，发现民众的参与可以帮助政府更好地了解当前社会公共系统存在的问题。同时，公众参与设计可以较好的减少决策中潜在的争议，更容易获得支持^[49]。Dalsgaard 对一个长达 8 年的公民参与的公众图书馆建造项目进行了定量和定性的调研，发现公众参与需要有效的管理和监督，统一安排，明确公民参与的部分及其作用，才能有效地让社会智能发挥应有的作用^[50]。

2.2.7 现状分析与发展趋势

从上面对众包、问答系统和公民参与等利用群体思维来实现社会智能的典型应用的分析现状来看，这些研究主要集中在以下几个方面：1) 群体思维模型，即描述和建模群体行为中聚合智能的呈现方式与过程，如众包中的线上和线下模型、在线和离线模型等；2) 群体思维质量评估与控制方法，即通过对群体中个体行为进行控制和评估，从而提高由此聚合产生的社会智能的质量，如通过仲裁和同行评议方法提高众包的质量等；3) 社会智能系统界面与设计方法，即通过界面设计与系统构建来支持和承载不同群体智能行为，如以用户个性为中心的公众参与应用的界面设计；4) 群体智能激励机制，即利用群体中个体之间的协同与竞争关系来实现群体思维与智能的低耗、高效的执行与产出，如基于社会半透明的众包激励方法等；5) 特定领域社会智能应用，如软件众包、盲人用户问答系统、公共交通设计的公民参与等。未来随着众包、问答系统等规模的不断扩大，应用领域范围越来越广泛，服务质量控制和保障要求愈加严格，特别是社交媒体、新媒体等技术的飞速发展与普及，群体智能系统研究将呈现的趋势包括：融合社会科学与计算科学的群体思维与社会智能基础理论，基于计算技术的群体智能涌现系统的构建技术，基于新兴交互技术的群体智能激励机制，以及对特定领域群体思维行为的分析与理解等。

2.3 情感挖掘与人人协同

在传统意义的协同计算模式中，更重视协同环境中计算机之间的数据共享、协作方式等问题，而忽略了以人为本，人与自然的协同和谐这一基本要素。将人的情感信息引入到 CSCW，通过对协同工作中参与者的情感进行挖掘，让参与者相互感知对方的情感

状态，实现类似于现实中面对面的人人协同工作，已经成为 CSCW 研究的一个热点。

情感协同计算研究体系自底向上主要包括情感信息获取、上下文感知建模、情感反馈与表达等几个层面，为教育、医疗、娱乐等不同的领域的情感协同应用提供支持，如图 1 所示。

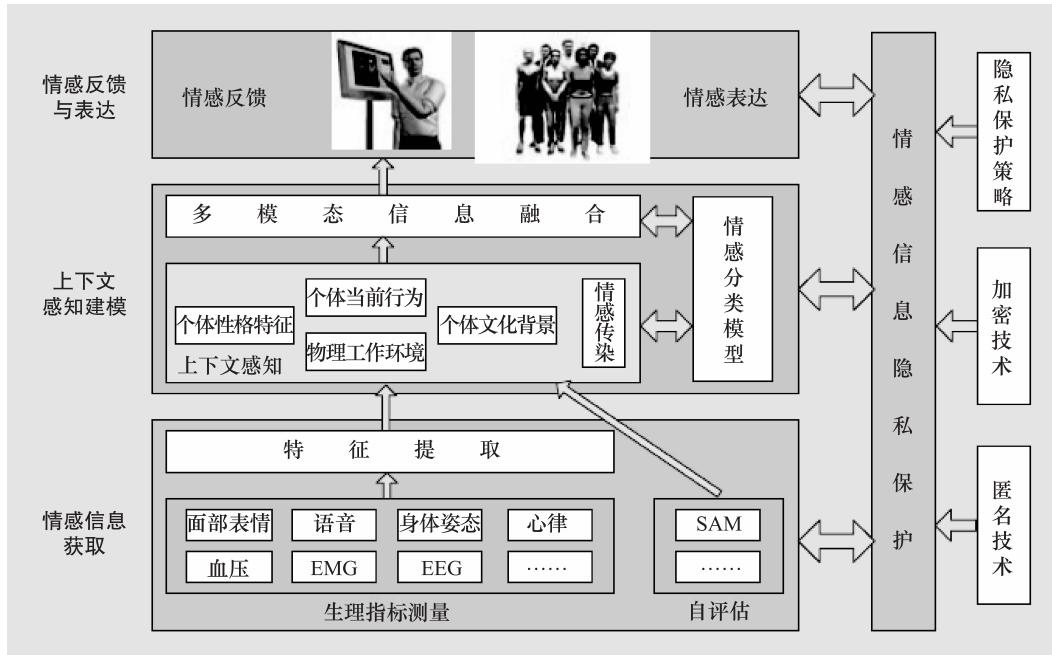


图 1 情感协同计算研究体系

2.3.1 情感信息获取

情感协同计算中首要问题是获取参与者自然真实的情感信息。目前对情感信息数据的采集主要通过自评估和生理指标测量。

自评估是由个体通过口头描述或者填写文本、图片量表的形式确认自身的情感体验。早在 1980 年，就有研究者开发出一种评估情绪的图片量表（Self-Assessment Manikin, SAM），在 SAM 中，用户在使用产品时的情感体验分为三个独立的维度（愉悦度、唤醒度和优势度），用户需要在这三个维度中选择一张适合自己当前状态的图片。不过，使用这种图片量表来评估情感也具有一定缺陷，例如存在用户对图片理解的偏差和个体差异，还有用户反映其情绪状态难以用当前的任意一幅图来准确描述。

生理指标测量是以生理变化作为情感的客观指标。其依据在于情感可以激活自主神经系统。当人处于比较激烈的情感状态时，许多生理指标（包括面部表情、语言、身体姿态、心率、血压、呼吸的频率和深度、肌电信号以及脑电信号等^[52]）就会出现变化。借助各种生理信号测量设备，特别是新型的可穿戴设备，可以自然准确地获取这新情感信息。例如，IBM 发明的 Emotion Mouse^[53]就可以测量手指压力、手掌温度等多种数据。

由于情感的表达是一个多成分的复合过程，包括内在体验、外显表情和生理激活这三种成分。故而，采集情感数据时往往需要使用多种设备，同步记录多个通道上情感信号的变化，再辅以参与者填写的情感主观体验问卷就可以从内在体验、外显表情和生理激活三方面反映情感的全貌。

对于采集到的各种原始形态的生理指标数据，还需要从中提取最能反映情感状态的特征，以方便情感识别。面部表情、身体姿势以及语音的特征提取，可以借助计算机视觉和语音技术的研究成果。而肌电图（EMG）、脑电图（EEG）等电信号的特征提取，则可以借助信号分析的研究成果，从时域、频域线性及非线性动力学等多个方面进行处理^[54]。

2.3.2 上下文感知建模

人类学研究表明，情感不是单纯地由生物进化、个体特征等内在因素引起，还在很大程度上依赖于外部社会的因素。因此，对于协同工作环境中个体情感状态的准确感知需要在情感上下文中进行，主要包括物理工作环境、个体性格、个体文化背景、个体当前行为、当前协同交流状况等信息。

上下文感知从情感信息感知设备处获得原始情感数据（底层上下文），再结合情感上下文对底层上下文进行封装、存储、推理（得到高层上下文）等操作，为上层应用提供支持。

协同工作的多用户、群体性，以及参与者之间为完成工作目标所需要进行的不断的交互等特点推动了对情感传染这一关键问题的研究。情感传染指个体情感传播、蔓延到其他个体、或整个协同工作组的过程。Hareli 和 Rafaeli^[55]曾经提出一个情感循环模型，并指出个体强烈的情感状态会对旁观者的情感产生很大的影响。Barsade^[56]通过实验证明了情感传染对于小组工作的影响，积极的情感传染可以提高写作能力、降低冲突，而消极的情感传染结果则恰恰相反。

上下文感知的目的是要建立起外在的情感表现和内在的情感状态之间的映射关系，其本身是一个典型的模式识别问题。由于情感状态是一个反映内在体验的、隐藏在多种类别的情感数据中的量，所以情感表现与情感状态之间是一个复杂的多对多映射。考虑到这一点，于是在上下文中感知中引入多模态信息融合技术，使计算机同步分析视频、音频、生理指标等多种信息。最大限度地提高识别的准确率。例如，麻省理工学院情感计算研究小组综合利用面部表情、坐姿、游戏难度等多个模态的信息，识别游戏者的感兴趣、厌倦、困惑等情感，与只用面部表情相比，识别率有较大的提高。

情感分类模型描述了不同情感之间的相似性和差异性。整个上下文感知过程都必须以情感分类模型为参考，将识别出的情感状态对应到不同的情感分类中去。目前使用最广泛的情感分类方法是维度法。

维度法认为几个维度组成的空间包括了人类所有的情感，并把不同情感看成是逐渐的、平稳的转变，不同情感之间的相似性和差异性根据彼此在维度中间中的距离来显示。Wundt 的情感三维理论认为，情感由三个维度构成：愉快 - 不愉快；激动 - 平静；紧张 -

松弛。各种具体情感分布在三个维度的两极之间的不同位置上。Plutchik 从强度、相似度和两极性三个方面进行情感划分，得出 7 种基本情感：狂喜、警惕、悲痛、惊奇、恐惧、接受、憎恨。Russell 把人的情感划分为两个维度：愉快度和强度。愉快度可以分、正、负极，强度可分为中、高强度，这样就形成了 4 种类型，即愉快 + 高强度 = 高兴，愉快 + 中强度 = 轻松，不愉快 + 高强度 = 惊恐，不愉快 + 中强度 = 厌烦，由此形成了情感环型模式。

2.3.3 情感反馈与表达

情感反馈是指协同工作系统能够分析参与者情感产生的原因，并对此作出合理、恰当的情感反应。

情感表达是指通过协同工作系统的交互设备向参与者表达其他人或者整个用户群组的情感状态。拟人化的情感表达是目前研究的主流，例如，通过语音、动画等方式表达情感。此外，借助于音乐、色彩等手段，系统也可以传递情感信息。随着情感计算研究的深入，情感表达的广度和深度都有了巨大的进步。现在，计算机可以综合运用面部表情、肢体语言、语言等模态，表达快乐、愤怒、悲伤等情感，甚至能够在一定的情境下表达出同情、困惑等细微情感。在面部图像合成领域，计算机已经能够合成出混合表情的面部图像，计算机合成的语音也不再是机械单调的，能够体现出快乐、愤怒、悲伤等不同情感基调。例如，穆罕默德·本·阿马尔^[57]等科研人员在其协同学习系统中使用了虚拟动画来表达用户的情感状态。

2.3.4 情感协同计算典型应用

(1) 协同医疗

协同医疗是指运用计算机、通信、医疗技术和设备，通过数字、文字、语音和图像资料的远距离传输，实现专家与病人、专家与医务人员之间异地“面对面”会诊，它可以使身处偏僻地区或“行动不便”的个人获得良好的诊断和治疗。

Helena M. Mentis 等研究人员^[58]就曾提出医院急诊室通常是一个快节奏、高强度、高压力的环境。在这个环境中，病人的各种体征都会发生快速的变化，医护人员要经常面对突发情况并在极短的时间内做出决定。在整个过程中，患者与医护人员之间的情感表达非常频繁并很可能影响决策的制定；同时，如果能及时对医护人员的情感状态进行调整，还可以使整个团队的工作更加有效。比如，主治医生可以通过情感协同医疗随时掌握护士的情感状态，指导和规范护士的护理方式。A. Luneski 等研究人员^[59]则分析了情感因素在治疗、康复工程中的重要作用，提出可以将情感信息合理引入医疗信息系统中，并列出了多项正在欧洲开展的相关研究项目。

Alison Benjamin 等研究人员^[60]发现，如果将对慢性病的传感器应用到与病人的交互中，那么慢性病患者就可以参与到社会交际中，而这大大降低了慢性疾病在他们生活中的破坏性。最近，Sherri Jean Katz 等研究人员^[61]提出在移动医疗应用中引入开放终端的社会意识的概念，使得使用者能够了解他人和自己的医疗决定，这样使用者就能够更灵

活、更自由地判别哪些因素对健康来说更为重要。

(2) 协同学习

协同学习是一种为了促进学习、由多名学生协作完成某个给定学习目标的教学方法，是一种通过小组或者团队的形式组织学生进行学习的一种策略。在协同学习过程中，小组成员个人学习的成功与他人的成功密不可分，协同工作是实现整体学习目标的有机组成部分。小组协作活动中的个体可以将其在学习过程中探索、发现的信息和学习资料与小组中的其他成员共享。Ann Jones 和 Kim Issroff^[62]讨论了认知和情感的关系，强调了情感在认知过程中的重要性。在这个过程中，人的情感状态可以对师生之间、学生之间的交流产生极大的影响。引入情感信息的协同学习可以在教育和认知过程中测量学习者的感情状况，并通过调整学习资料或者学习方法等形式，及时地对学习和认知过程进行干预，取得最满意的学习和认知效果。

(3) 协同设计与协同编辑

协同设计和协同编辑是指各设计、编辑人员在共享信息空间中，相互协调、合理分配设计资源，共同合作完成设计目标和编辑任务的过程，CSCW 的主要应用方向。Mauro 等研究人员指出注意力集中程度在协同编辑环境中的重要性，并设计了一种名为“电子眼”的技术帮助协同编辑人员及时感知各自的注意力集中程度。该技术使用眼动仪对参与者的眼部活动进行跟踪，分析在系统编辑工作中参与者的注意力集中在何处。Jeremy Birnholtz 等研究人员^[63]发现协同编辑系统中的编辑和评论等行为往往包含有一定的社会意义，而这样的社会意义能够影响团队协同工作的情绪和效率。

(4) 游戏娱乐

普适计算以及人工智能等新技术的引入，使得协同游戏环境中游戏玩家之间的交互和用户体验得到了新的发展。但是，对于一款协同游戏好坏的评价却不能只从性能等方面来考虑，还应该从用户的情感角度出发，根据用户的喜爱程度进行。Regan L. Mandryk 等研究人员^[64]将玩家相关的生理数据引入协同游戏环境中，以此识别玩家的情感状态，并将其作为评价游戏好坏的客观标准。

2.4 移动协同环境

2.4.1 定义

随着无线网络技术和移动便携设备终端技术的发展，人与人、人与物、物与物之间“随时随地”协同通信、协同工作等更加普遍，由此构成无线的“移动协同环境”。2000 年 Wiberg 等人首次提出移动协同的概念^[65]，将 CSCW 移植到移动环境中，为传统 CSCW 带来了更多新型应用，也同时增加了新特点和新的挑战。

移动协同环境的优点主要表现为“服务的随时随地性”，即移动无线通信环境为更广泛的应用服务、更灵活的服务方式提供了有力支持。但同时，它带来的挑战，也是它的三个主要特点体现为：

1) 环境的动态变化性。移动性使得结点间连接不稳定，结点构成的拓扑结构随时间不断变化。

2) 通信的不稳定性。移动协同环境中的通信为无线通信，结点的移动性使得移动协同环境中的无线传输具有不可靠、抗干扰能力差的特征。

3) 结点能力的有限性。由于移动终端设备的电源、体积的限制，因此，移动协作环境中单感知结点的能量和处理能力有限，结点间通信带宽有限。

由此，如何让移动结点之间信息互通、结点之间协同感知环境成为构造移动协作环境的重要支撑。下面将对移动协同通信、协同感知这两个方面的国内外研究现状进行综述和分析。

2.4.2 移动协同通信的相关研究

移动协同环境中，协同通信是移动协同环境成立的重要基础。当前移动协同通信的研究主要有两大类：一类是针对通信频率资源对通信频谱的分配进行研究，一类是对不确定通信路由机制的研究。其中利用端到端的位置相近性和短距离通信能力的完全分布式的移动通信路由研究成为当前一项热点研究，它在移动社会网络信息交流、旅游路线推荐、广告定向发送等商业应用、灾难现场信息传递、感知人类群体交互行为模式、车联网等方面具有巨大的应用价值。

2004 年，麻省理工学院的 Biswas 等人^[66]将这种利用结点移动性进行短距离通信构成的自组织网络称为机会网络。机会网络是一种重要的移动协同通信网络，它是一种无结构的对等网络，各结点可充当不同角色，既可能是信息发送者，也可能是信息接收者，还可能是中继转发者，通过各结点的自主智能判断进行下一跳结点选择，多个短距离通信协作达到远距离通信的目的。它充分利用了信道广播特性，通过设备和设备间的协作通信，提高吞吐量和传输可靠性。

机会网络的主要特点是依靠移动实现协同通信，如何进行消息转发以减少延迟是其关键问题。主要有两种方法，一种是基于网络编码的转发，一种是基于复制的转发。基于编码的转发机制是将待传输数据编码成相互冗余的消息，目标结点仅需要接收到部分编码后的消息，即可通过消息之间的运算重建原数据。基于编码的转发需要转发结点进行编解码计算，加大了计算开销。基于复制的方法主要是通过移动关系转发信息备份达到传送目的。由于移动网络环境是以人为中心的环境，机会网络的路由转发不可避免地与人的社会网络结构具有紧密的联系，所以基于社会结构分析、复杂网络分析的机会路由成为当前主流。按照计算方法以及成果发表时间分类，基于社会结构进行机会路由的相关研究可以分为如下四类：

第一种是着眼于微观层面的路由设计研究，这也是机会网络发展初期的主要研究方法，主要根据结点历史移动信息的相遇概率、通信上下文信息等进行多转发设计和移动模型设计。最早也是最简单的是 Epidemic 协议^[67]，它是一种泛洪协议，覆盖率高但是耗费资源；Prophet^[68] 和 spay and focus^[69] 协议是两种较早将社会关系引入机会网络的路由协议，它们的一个隐含假设就是结点的移动具有周期性，通过对结点和目标结点的相遇

历史记录等计算，决定哪个结点为转发点。这类方法的主要特点是根据结点个体信息或结点间的信息进行决策。

第二种是基于中观层面的路由设计，鉴于移动网络的动态变化和与人类社会活动的紧密相关性，以及基于结点局部感知信息的决策在应对人们社会网络动态变化的非有效性、社会结构的相对稳定性等^[70,71]，一些学者开始从中观尺度考虑机会路由问题。典型做法就是引入了社区结构，然后基于这种粗粒度拓扑进行分层路由。SimBet 系统^[72]引入社区和中心性，首先根据结点间的相似性划分社区，然后将中心性和相似性综合计算形成效用评价，信息转发给效用度高的结点。Panhui^[72]等同样引入社区和中心性建立了BuBBLE Rap，结点将消息传送给公共度更高的点或与目标点属于同一社区的点；牛建伟等人^[73]根据结点之间的通信频繁程度，自动将结点划分成不同的社区，自适应地控制消息的复制数量并依靠活跃结点将消息传输到目标社区。Wang 等人^[74]根据通信频次建立社区并设定社区间的连接权值，同一社区内采用 Prophet 协议通信，不同社区间根据权值进行消息转发。鉴于移动环境动态性引发社区的动态变化，Nam P. Nguyen 等人^[75]提出一种快速发现动态社区的算法，M. J. Williams 等人针对具有周期性的活动场景提出一种周期社区的发现算法^[76]。Sabrina Gaito 等人^[77]引入通信时长这一因素，根据联系时长和频度将结点分为 4 种不同角色：联系时间长且频度高的点构成社区，联系时间短且频度高的点为熟悉的陌生人，联系时间短且频度低的点为陌生人，联系时间长且频度低的为朋友。最后发现短时联系可有效增强网络连通性、传送的覆盖性以及低延迟。可以看出，基于社区结构进行机会转发，在一定程度上减少了延迟，但仍存在以下局限性：①缺乏对机会网络环境中社区结构粒度、重叠性等如何影响机会路由的深入探讨，而这些是决定社区是否真正有用的关键。②社区结构的动态演化性及社区结构形态的多样性。随着时间变化，社区结构是在不断演化的，在一些场景中是平稳变化，但在另一些场景中，很可能是突兀的周期性变化，也可能是随机变化，所以，当前基于社区中观结构的机会传送的研究忽视了这些不同场景“社区结构存在与否”以及“若存在是什么形态”的发现和判别方法的研究，具有很大的局限性。③在机会网络中，也许有大部分的结点具有（短时或长时的）聚集或社区结构属性，但还有其他的一些结点，它不属于任何一个社区，却又会游离于一些社区之中。这些点与重叠点不同，因为重叠点是那些与不同社区联系都紧密的点，而这些点是偶尔出现在一些社区但又与社区联系不紧密的点，在基于社区的机会转发研究中，大都忽略了这类点。但实际上这些点在机会路由中的作用也是不应忽略的，文献^[77]对此作了一些初步的工作，已有一些有趣的结论，仍然需要进一步的工作。④当前基于中观结构的机会路由研究所关注的主要还是社区和中心性，而实际上社会网络结构还有许多其他可能影响路由的结构属性，例如 closeness、三角闭合性、自相似性以及通信活动中的幂率等，这些属性在移动社会网络中的表现如何以及是否可以用来进行路由设计，也是值得注意的。由此也引来了一类新的研究，即宏观层次上的关于移动模型的研究。

第三种是在宏观层面进行统计意义上的理论研究，发现移动社会网络的结构特点。挖掘和探测统计意义上的人类移动结构特征是最近的研究热点，通常采用复杂网络分析

的方法，分析和发现移动通信过程的各种结构参数的分布规律；也有利用概率统计理论和近似算法等方法，对移动社会网络的分布密度、转发成功率等关系进行研究。这类研究对于部署服务、设计路由策略（或内容共享、信息传送等）具有重要价值。复杂网络结构发现的主要结论有：A. Chaintreau 等人^[70]发现一定有限时间范围内移动通信间隔为幂率分布，并理论证明了分布参数如何影响机会转发延迟。瑞典 T. Hossmann 和法国 T. Spyropoulos 等人^[78]提出，机会网络路由设计不仅需要对微观层面的联系规律分布的统计分析，而且需要宏观层面关于联系关系的结构分析。他们对四种不同规模的场景进行分析，发现移动符合小世界模型，具有社区结构，但是社区内结点并非是同构点，社区内的权重和度分布与全局的权重分布规律相似。Zhao 等人^[79]对无线 ad hoc 网络中的结点密度和延迟之间的关系进行了理论研究。

第四种是从机会网络的社会理性角度进行的研究，由于结点设备使用者的社会理性以及机会网络完全分布式的通信拓扑，使得该网络模式不一定是合作模式，这类研究主要从结点的自私性方面研究结点分布、转发策略与性能的关系^[80]。

还有一些学者从别的角度进行机会路由的研究，例如，北邮的学者们提出根据消息间的关联性在传递过程中将消息集成进行传送，提升通信性能^[81]。

2. 4. 3 移动协同感知的研究现状

移动协同通信机制使结点之间能够互相连通，构成了协作互联网络的基础设施；结点状态的相互感知、协同感知外部环境等是多结点协同工作的第二个重要内容，即移动协同感知。移动协同感知面对的主要挑战为：如何解决结点间频繁短接带来的感知信息的质量不稳定问题，如何将异质网络信息集成到统一框架中并进行有效融合，如何优化协同结点计算能力、能量和资源，协同感知中的安全和信任问题等。

针对移动环境频繁短接带来的信息感知质量不稳定问题，学者们提出多种中间件技术。Mirco Musolesi 等人^[82]对信息中间件 JMS（Java Message Service）进行改进，添加了消息传送的路由机制，提出 EMMA（Epidemic Messaging Middleware for Ad hoc networks）机制。A. Murphy 等人^[83]基于信息共享机制提出了一种中间件 LIME（Linda in a Mobile Environment）。Fok 等人^[84]提出了 LIME 的改进版本 LIM ONE，它提供了一个高度不确定网络状态下的强社交模型，其特征是上下文管理、外在数据访问、代码移动和代理移植。比起 LIME，它考虑了可测量性，有限硬件资源和安全的问题。L Capra^[85]等人针对数据共享，提出 XMIDDLE 中间件，它提供了处理频繁断接操作的机制，解决了移动中的设备资源缺乏和频繁断接的问题。Levis 等人^[86]提出了 Mate（mahtay）方法，用于 ad hoc 无线传感器场景，提出了从简单参数调整到使用 VM 方法的卸载完全程序的频谱改编方法，克服了限制带宽、能量耗损大的问题。Curino 等人^[87]提出的 TinyLIME 是另一种专门针对传感器网络的中间件，它从 LIME 发展而来，在 TinyOS 下支持移动协同感知中间件，能很好支持客户端的移动性。

针对异质网络协同感知信息建模和集成问题，Jakob E. Bar dram 等人^[88]提出了上下文中介的社会感知概念，介绍上下文感知计算是如何帮助社会意识的。Yushun Li 等

人^[89]提出 3 DCAM 三维感知模型。该模型可以用在不同的协作领域描述感知信息，为需要协作的感知结点提供信息种类、信息方式、构建信息的方式。

3 国内研究进展

3.1 国内社会化协同系统开发与应用现状

由于相同的价值观、态度、兴趣而把一个人同其亲戚、邻居和朋友等社会性地联系起来，形成了社会网络。基于社会网络概念形成的各类软件系统，一般称为社会化协同系统。在过去的一年中，国内的社会化协同系统取得了长足的发展，成为互联网的主流趋势。本文首先介绍主流的社会化协同系统，然后针对性的展示学术领域有代表性的社会化协同系统，并进行比较。

3.1.1 主流的社会化协同系统

从早期的开心网、人人网，再到后来加入的腾讯朋友网等，都是社会网络的典型应用，吸引了众多用户的参与。在目前的业界，各类软件都呈现社会化的发展趋势，新浪微博、QQ 圈子等陆续吸引了国内广大用户，微信等产品更是迈出国际化的步伐，开始了国外的发展布局。另外，各类即时通信软件（如 QQ、MSN 等）、视频分享软件（如优酷土豆等）也是社会化协同软件发展的代表。下面介绍几种典型的社会化协同系统。

(1) 开心网

开心网由北京开心人信息技术有限公司创办于 2008 年 3 月，是国内第一家以办公室白领用户群体为主的社交网站。开心网为广大用户提供包括日记、相册、动态记录、转帖、社交游戏在内的丰富易用的社交工具，使其与家人、朋友、同学、同事在轻松互动中保持更加紧密的联系。成立 4 年多以来，网站注册用户已突破 1.3 亿（截至 2012 年 4 月底），已发展成为中国领先和最具影响力的实名化社交网站。

(2) 人人网

人人网是由千橡集团将旗下著名的校内网更名而来。2009 年 8 月 4 日，将旗下著名的校内网更名为人人网，社会上所有人都可以来到这里，从而跨出了校园内部这个范围。人人网是为整个中国互联网用户提供服务的 SNS 社交网站，给不同身份的人提供了一个互动交流平台，提高用户之间的交流效率，通过提供发布日志、保存相册、音乐视频等站内外资源分享等功能，搭建了一个功能丰富高效的用户交流互动平台。

(3) 朋友网

朋友网，原名 QQ 校友，是腾讯公司打造的真实社交平台，为用户提供行业、公司、学校、班级、熟人等真实的社交场景。2011 年 7 月 5 日，腾讯公司正式宣布旗下社区腾讯朋友更名为朋友网。

(4) 新浪微博

新浪微博是一个由新浪网推出，提供微型博客服务的类 Twitter 网站。用户可以通过网页、WAP 页面、手机短信、彩信发布消息或上传图片。可以把新浪微博理解为“微型博客”或者“一句话博客”。用户可以将看到的、听到的、想到的事情写成一句话，或发一张图片，通过电脑或者手机随时随地分享给朋友，一起分享、讨论；还可以关注朋友，即时看到朋友们发布的信息。

(5) 微信

微信是腾讯公司推出的，提供类 Kik 免费即时通信服务的免费聊天软件。用户可以通过手机、平板、网页快速发送语音、视频、图片和文字。微信提供公众平台、朋友圈、消息推送等功能，用户可以通过摇一摇、搜索号码、附近的人、扫二维码方式添加好友和关注公众平台，同时微信将内容分享给好友并将用户看到的精彩内容分享到微信朋友圈。

3.1.2 社会化协同系统的特征

社会化协同系统在使用过程中反射了社会网络，使用者映射到社会中的个人，具有个体性，个体可以组成不同大小的团体，具有层次性，个体团体动态演化，又具有自组织性和运动性特征。

(1) 个体性

社会化协同系统通常以自我为中心，松散耦合，个人主动参与到群体中，是多人对多人的系统。其中的人际关系具有匿名性和不在场性，存在大量的陌生人关系，而使用者的身份和信任在系统中逐渐体现，虚拟的网络活动与现实的社会活动逐渐融合。软件本身也会在发展过程中不断更新和自我完善。

(2) 层次性

社会化协同系统具有层次性，可分成多个级别。级别划分的颗粒度可根据需要灵活设置。小级别是大级别的组件，大级别是小级别的环境。现实社会的级别由小到大可以分为个人、家庭、社团、城市、国家和人类社会。虚拟社会的级别可以分为个人、群组、虚拟社区、社会媒体和虚拟社会。

层次是系统由元素整合为整体过程中的涌现等级，不同性质的涌现形成不同的层次，不同的层次表现不同质的涌现性。出现新的涌现性不一定产生新的层次。层次分析是结构分析的重要方面。系统是否划分层次，层次的起源，分哪些层次，不同层次的差异、联系、衔接和相互过渡，不同层次的相互缠绕，层次界限的确定性与模糊性，层次划分如何增加系统的复杂性，层次结构的系统学意义，层次结构设计的原则等，是层次分析要回答的问题。

(3) 运动性和自组织性

社会化协同系统的结构、状态、行为、功能是动态演变的，可用一组外化的社会指标反映出来。这些社会指标的综合指数的走势曲线反映了社会系统的运动轨迹。通过分析社会指标走势曲线可以预测社会未来的趋势。社会化协同系统具备一定的自组织性，

低级社会系统是高级社会系统的组件，低级社会系统可通过自组织性，生长为高级社会系统。不同级别的社会系统，尽管级别不同，但各个级别的演变过程遵循大自然法则，均可分解为若干个阶段（生、兴、衰、病、死），呈现自同构性。根据系统论的观点，社会系统这种自同构性相当于社会系统的基因。级别具有自组织能力，根据级别的自组织性，按着基因图谱，低级系统如同生命般自动生长出高级系统。

（4）社会化协同系统的发展趋势

计算机网络技术对传统产业技术升级产生了巨大的影响。1990 年至 2000 年，PC 软件时代，推动了传统企业的自动化；2000 年至 2010 年，Web1.0 时代，推动了传统企业的网络化；2011 年至 2020 年，社交媒体时代，必将推动传统企业的社交化。云计算、移动网、物联网、社会网将主导今后的十年。云计算的应用场景不应是收水电费模式，而是作为云端技术，驱动社交媒体的蓬勃发展。基于移动网、物联网、社会网的云服务相互连通。随着移动设备的普及，各类社会化协同系统的移动版纷纷上线。

社会化协同系统具有超大规模的结构化数据，包括用户信息、用户关系和资源信息。大规模用户的数据库并发访问，每秒需要处理上万次读写请求，用户忍耐时间为 2 秒。数据库有高可扩展性和高可用性的需求，还需要数据库结构、存储结点的可扩展性及软件架构实现的结点容错能力。

关系型数据库难以适应数据处理要求。关系数据库应付上万次 SQL 查询还勉强顶得住，但是应付上万次 SQL 写数据请求，硬盘 I/O 就已经无法承受了。数据库的水平扩展能力差，需要云计算平台的 NoSql 数据库。例如 Hadoop HBase、Google BigTable、Amazon Dynamo、Facebook Cassandra。绝大多数的社会化协同系统选择开源的 Hadoop 搭建自己的云计算平台。

所有的社会化协同系统本质上都是一种基于云计算的 Web 应用。未来的架构将是云计算和社交服务。后台的云计算平台是大规模数据中心，由大量的服务器集群组成，提供强大的存储、计算能力。前端则提供基于 Web 的社交服务。

3.1.3 学术社会化协同系统

在学术界，中国知网、学者网、天玑学术网、学术空间（C-DBLP）、科研之友、Arnet-Miner 等得到了进一步发展，成为了针对学术领域的有代表性的社会化协同系统，为广大的学者用户提供了良好的服务。

（1）中国知网

中国知网是国内最知名、最权威的学术文献检索系统，包含了丰富的中英文文献资料。在过去的一年中，除了继续丰富中英文文献资料外，中国知网又在个性化服务方面发力，推出了 CNKI 学者成果库功能。它可以为用户自动收集整理学术成果，根据姓名、工作单位实时推送最新论文，方便展示用户的最新研究成果。美中不足之外是用户个人的管理功能较弱。

（2）学者网

学者网是华南师范大学汤庸教授团队研发的专注于学术服务的网站。在过去的几年

中，经过不断的持续开发，学者网的功能日渐丰富，形成了学者的个人主页、社交网络、学术搜索、在线信息服务等几大核心功能。具体功能有：

1) 个人主页：为学者提供发布个人主页功能。学者可以在主页上显示自己的简历，介绍自己的论文、项目、专利、课程等信息，分享学术资源、发布公告等，并且提供了多套中英文模板，为学者增强自己在国内外的影响力提供了一个便利的开放平台。

2) 社交网络：

- 学者之间可以互相结为好友。
- 好友之间可以通过多种方式进行交流，如站内通信、在线交流等。在常规的交流功能之外，这两种方式都实现了文件的传送功能，学者之间的交互更加顺畅。相比较而言，站内通信较为正式，可以保存所有的交互历史，目前已具有正式邮件的雏形；而在线交流更注重即时性和便捷性。
- 学者可以加入一个团队进行协作。团队为学者提供了一个学者间进行讨论以及资源共享的平台。
- 在用户的学术空间中，用户可以很及时地了解到好友的动态及团队中的最新信息。
- 学术分享：好友之间可以分享感兴趣的论文等学术作品。
- 教学课程：针对教学型用户，提供了教学课程工具，可以实现与学生分享课件、在线答疑等常用的功能。

3) 学术搜索：目前，学者网学术搜索引擎索引进了近 9000 万篇中英文论文，覆盖比较全面，极大地方便了学者查阅感兴趣的论文，并在好友之间分享。另外，可以为注册学者自动推送他的学术成果，明显减轻了用户维护个人主页的负担。

4) 在线信息服务：可以为学术机构提供公开的信息服务接口，自动获取相关学者的最新动态。

5) 学术工具：学者网还为用户提供了学者日历、学术网盘、导出简历、常用链接等工具，为用户的科研活动提供了很大的便利。

6) 移动终端：学者网还开发了移动客户端的版本，进一步增强了用户使用学者网的便捷性。

(3) 天玑学术网

天玑学术网是中科院计算所近年新推出的一款类似于微博的学术服务网站，它具有很强的社交属性，用户可以自由地与其他学者交流学术观点，分享学术资源，形成了浓郁的学术交流气氛。此外，天玑学术网还提供了论文的在线阅读功能，并且用户可以在论文的在线阅读页面中进行很方便地交流，就论文中的某一问题进行讨论。

(4) 学术空间 (C-DBLP)

学术空间 (C-DBLP) 是中国人民大学孟小峰教授团队开发的学术服务网站。它的目标是建立以作者为中心的中文文献检索系统。开始的时候，C-DBLP 专注于计算机领域的中文文献，去年已扩展到经济、法学、物理、艺术等 25 个领域。它以“贵精而不贵多”的精神来选择收录的期刊，依照《中国人民大学期刊考核目录》进行收录，目前收录的

期刊增加至 466 本，收录文章数约有 172 万，收录作者超过了 96 万。

除了以作者为中心进行文献检索外，C-DBLP 还提供了 EasyScholar 功能，针对学者在建立学术主页时收集整理论文著作、科研项目等信息时诸多不便的情况，为国内学者提供个人主页的服务，利用 Web 数据集成技术，从网络上公开的信息中提取用户的学术信息，使用户可以较少的人工来建立并维护自己的主页。

(5) 科研之友

科研之友是具有官方背景的学术服务网站，由国家自然科学基金委员会指定在申报相关的基金项目时使用。因此，在科研之友中，最主要的功能就是为用户提供基金项目的申报功能。除此之外，用户还可以发布个人主页，并且与同行、好友建立联系，辅助基金申请和成果的推广，还可以建立群组，与其他学者就共同感兴趣的问题进行讨论。

(6) Arnet-Miner

清华大学计算机科学与技术系唐杰副教授在社交网络的语义信息抽取、基于话题的信息搜索、网络行为建模和影响力分析等方面进行了深入研究，研发了研究者社会网络 ArnetMiner 系统。该系统收集了 100 多万名研究者、300 万篇论文信息、3700 多万引用关系以及 8000 多个会议信息。从 2006 年运行以来，该系统吸引了 189 个国家 73 万个独立 IP 的访问（>6960 万访问日志），访问量还在以每月 10% 左右的速度增长。ArnetMiner 系统在国际顶级会议 WWW、KDD、ISWC、ICDM 中进行了演示和交流，在国际上具有一定影响力。该网站虽然是在国内开发的，但是目前仅针对英文文献进行检索，具有国际化的基因。

Arnet-Miner 提供了丰富的语义搜索工具，主要有：

- 用户资料提取：给定一位学者的名字，系统可以自动从互联网中提取有关他的信息资料，生成他的语义信息，例如，联系方式、用户照片、教育经历、研究兴趣、发表的论文等。
- 领域专家查找：给定一个学术领域，系统可以自动给出该领域有哪些著名的专家学者。
- 主题浏览：系统自动从论文中发现当前的研究主题，并自动以易读的文本方式表示出来。针对每个主题，可以给出相关的学者、会议/期刊，并且可以给出主题的演变情况，为研究者提供非常友好的帮助。
- 用户管理：学者可以在 Arnet-Miner 上注册，对系统自动提取的资料信息进行修改以建立正确的个人主页，此外，还可以关注其他学者的主页，方便了解其最新的研究进展。

3.2 国内主要研究团体和成果

传统 CSCW 的试验平台主要有白板类系统、协同图形编辑系统等，其中白板类系统最典型的如 Microsoft Netmeeting 中的白板，它可以简单仿真可擦写的绘图板，是一种最简单的共享图形工作空间。浙江工商大学姜波团队等开展了普适环境下图案协同编辑的

相关工作，提出了基于消息的动态体系结构 MASCA 及体系结构调整原语，基于消息路由实现了体系结构运行时的自适应动态调整，提高了实时协同编辑协作的效率。提出了基于协同任务和资源的系统性能评价模型。提出了动态自适应协同网络架构及基于多协议动态数据传输模型，实现多类型协同数据优化传输。提出了 AFEC 的控制点传输方法及比例可变更机器学习算法的控制点轨迹预测算法，提高了动态上下文协同感知质量。提出了基于扩展预测集的实时协同用户关注区预测算法，实现了资源受限情况下实时有效的共享工作空间上下文感知。提出了非对称多模态协同感知方法及基于认知的感知映射，提出了全局协作效率算法，同时以定量及可视化方式实现了适度感知及协作效率评价，上述方法可应用于在多类型设备参与和多类型网络条件下的协同编辑工作，具有重要的学术意义与应用价值^[90,91]。Tim Paek 针对大屏幕显示和移动设备（如手机、手提电脑、无线 PDA 等）之间的共享显示^[92]进行了深入研究。

在社会网络协同领域，中科院计算所程学旗研究员团队对在线社会网络中的协同推荐、协同影响力计算、排序学习等方面进行了系列研究，提出了群组稀疏图编码方法、协作因子分解推荐、基于 session-flow 的高实用查询词推荐等方法。清华大学的唐杰等人对社会网络中的链预测、影响力等进行了研究，提出一种无监督的动态因子图模型来推断个体间的关系类型，建立了集成社会学原理和三元组因子图模型的方法，从单元关系推断二元关系以及三元闭合关系的构成等，并提出一种协同进化模型来模拟社区动态变化过程。山东大学孙宇清教授团队研究了社会网络中的隐私保护问题，提出了基于历史数据分析的个性化隐私策略推荐方法，研究了整合策略验证、实施和审计的个性化隐私保护方案。

剑桥大学、清华大学、卡内基 - 梅隆等高校研究机构对无线网络视频和音频的协作处理、协作数据传输等方面进行研究和开发。IBM 和 Microsoft 将 CSCW 应用到移动企业级。

另一方面，可以看到，移动协同环境研究的持续高涨与短距离通信技术的发展、移动社会网络、物联网、普适计算等应用的快速发展有无法割裂的联系。意大利的 Marco Conti、美国 Mohan Kumar 等人^[93,94]分别在 2010 年、2012 年撰文展望未来移动机会网络，提出机会网络将促进移动社会网络的发展。西北工业大学的者与法国国立电信学院联合提出以物联网社会性为主题的“机会物联”^[95]，指出机会物联的研究将会促进“一系列以人为中心的创新性应用”的产生，“对于物理信息空间与人类社会的和谐融合、互促发展具有积极意义”。国内外研究团体纷纷开展相关研究工作，国外主要有瑞典的 ETH、意大利的 IIT、英国剑桥大学、美国 MIT、得克萨斯大学阿灵顿分校、IBM 等，国内主要有北邮、清华、北航、北大、重邮、上海交大、西北工大等；一些重要项目纷纷开展，主要有欧盟 7 框架中的 Awareness 子项目、Socialnet、Metrosense、法国国立电信学院的“社群智能”、美国大都市移动网络项目、国内北京邮电、清华大学和诺基亚研究院的移动社会网络等相关项目，移动协同通信、感知的研究是这些项目中的重点研究内容。

参考文献

- [1] Xia Hu, Huan Liu. Social Status and Role Analysis of Palin's Email Network. WWW 2012 Companion, April 16-20, 2012, Lyon, France: 531-532.
- [2] Xiao- Yong Wei, Zhen- Qun Yang. Mining In- Class Social Networks for Large- Scale Pedagogical Analysis. MM'12, October 29-November 2, 2012, Nara, Japan. : 639-648.
- [3] Pengyan Yao, Zheng Hu, Zhenzhen Zhao, Noël Crespi. A Study of Social Behavior in Collaborative User Generated Services. ICUIMC'12, February 20-22, 2012, Kuala Lumpur, Malaysia.
- [4] Elnaz Davoodi and Keivan Kianmehr. A Semantic-Based Social Network of Academic Researchers. IEA/AIE 2012, LNAI 7345, pp. 323-332.
- [5] Sunju Oh, Heon Y. Yeom. A social network extraction based on relation analysis. ICUIMC'12, February 20-22, 2012, Kuala Lumpur, Malaysia.
- [6] Yukihiko Matsumoto, Sho Tsugawa, Hiroyuki Ohsaki, Makoto Imase. Robustness of Centrality Measures against Link Weight Quantization in Social Network Analysis. SIMPLEX '12, April 17, 2012, Lyon, France: 49-54.
- [7] Juliano C B Rabelo, Ricardo C B Prudêncio, Flávia A Barros. Leveraging Relationships in Social Networks for Sentiment Analysis. WebMedia'12, October 15-18, 2012, São Paulo/SP, Brazil. : 181-188.
- [8] Guo Zhang, Elin K. Jacob. Community: Issues, Definitions, and Operationalization on the Web. WWW 2012 Companion(SWCS'12 Workshop), April 16-20, 2012, Lyon, France. : 1121-1129.
- [9] Thiago Rúbio, Rogério Gomes, Adriano Pereira. Behavioral Analysis in Social Networks: An Approach Based on Intelligent System. WebMedia'12, October 15-28, 2012, São Paulo/SP, Brazil. : 193-196.
- [10] Alan Said, Ben Fields, Brijnesh J Jain, Sahin Albayrak. User-Centric Evaluation of a K-Furthest Neighbor Collaborative Filtering Recommender Algorithm. CSCW'13, February 23-27, 2013, San Antonio, Texas, USA. : 1399-1408.
- [11] Jian Chen, Guanliang Chen, Haolan Zhang, Jin Huang, Gansen Zhao. Social Recommendation Based on Multi-relational Analysis. 2012 IEEE/WIC/ACM International Conferences on Web Intelligence and Intelligent Agent Technology : 471-477.
- [12] Jun-Ki Min, Jason Wiese, Jason I Hong, John Zimmerman. Mining Smartphone Data to Classify Life-Facets of Social Relationships. CSCW'13, February 23-27, 2013, San Antonio, Texas, USA. : 285-294.
- [13] Xingjie Liu, Qi Hey, Yuanyuan Tian, Wang-Chien Lee, John McPherson, Jiawei Han. Event-based Social Networks: Linking the Online and Offline Social Worlds. KDD '12, August 12-16, 2012, Beijing, China. : 1032-1040.
- [14] Miltiadis Allamanis, Salvatore Scellato, Cecilia Mascolo. Evolution of a Location-based Online Social Network: Analysis and Models. IMC'12, October 29-30, 2012, Austin, Texas, USA. : 145-158.
- [15] Youngchul Cha, Junghoo Cho. Social- Network Analysis Using Topic Models. SIGIR'12, August 12-16, 2012, Portland, Oregon, USA. : 565-574.
- [16] Saajan Shridhar, Mithleshwar Lakhanpuria, Abhishek Charak, Ankur Gupta, Swapan Shridhar. SNAIR: A Framework for Personalised Recommendations Based on Social Network Analysis. ACM SIGSPATIAL LB-

- SN'12, November 6, 2012. Redondo Beach, CA, USA: 55-61.
- [17] Le Yu, Jian Zheng, Wei Chong Shen, Bin Wu, Bai Wang, Long Qian, Bo Ren Zhang. BC-PDM: data mining, social network analysis and text mining system based on cloud computing. KDD'12, August 12-16, 2012, Beijing, China. : 1496-1499.
- [18] Wenfei Fan. Graph Pattern Matching Revised for Social Network Analysis. ICDT 2012, March 26-30, 2012, Berlin, Germany. 8-21.
- [19] Donghyuk Shin, Si Si, Inderjit S. Dhillon. Multi-Scale Link Prediction. CIKM'12, October 29-November 2, 2012, Maui, HI, USA. : 215-224.
- [20] Ge Gao, Pamela Hinds, Chen Zhao. Closure vs. structural holes how social network information and culture affect choice of collaborators. CSCW'13, February 23-27, 2013, San Antonio, Texas, USA. 5-17.
- [21] Tamara Heck. Combining social information for academic networking. CSCW'13, February 23-27, 2013, San Antonio, Texas, USA. 1387-1398.
- [22] Sebastián A. Ríos, Ricardo Muñoz. Dark Web Portal Overlapping Community Detection Based on Topic Models. ISI-KDD 2012, August 12, 2012, Beijing, China.
- [23] Ting-Kai Huang, Md Sazzadur Rahman, Harsha V. Madhyastha, Michalis Faloutsos. An Analysis of Software Cascades in Online Social Networks. WWW 2013, May 13-17, 2013, Rio de Janeiro, Brazil: 619-629.
- [24] Steven Komarov, Katharina Reinecke, Krzysztof Z Gajos. 2013. Crowdsourcing performance evaluations of user interfaces. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 207-216.
- [25] Lilly C Irani, M Six Silberman. 2013. Turkopticon: interrupting worker invisibility in amazon mechanical turk. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 611-620.
- [26] Peter Kinnaird, Laura Dabbish, Sara Kiesler, Haakon Faste. 2013. Co-worker transparency in a microtask marketplace. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 1285-1290.
- [27] Shih-Wen Huang, Wai-Tat Fu. 2013. Don't hide in the crowd!: increasing social transparency between peer workers improves crowdsourcing outcomes. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 621-630.
- [28] Nicolas Kokkalis, Thomas Köhn, Carl Pfeiffer, Dima Chornyi, Michael S Bernstein Scott R Klemmer. 2013. EmailValet: managing email overload through private, accountable crowdsourcing. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 1291-1300.
- [29] Erin Brady, Meredith Ringel Morris, Yu Zhong, Samuel White, Jeffrey P Bigham. 2013. Visual challenges in the everyday lives of blind people. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 2117-2126.
- [30] Derek L Hansen, Patrick J Schone, Douglas Corey, Matthew Reid, Jake Gehring. 2013. Quality control mechanisms for crowdsourcing: peer review, arbitration, & expertise at familysearch indexing. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 649-660.
- [31] Shih-Wen Huang, Wai-Tat Fu. 2013. Enhancing reliability using peer consistency evaluation in human com-

- putation. In Proceedings of the 2013 conference on Computer supported cooperative work (CSCW ' 13) . ACM, New York, NY, USA, 639-648.
- [32] Afra Mashhadi, Giovanni Quattrone, Licia Capra. 2013. Putting ubiquitous crowd-sourcing into context. In Proceedings of the 2013 conference on Computer supported cooperative work (CSCW ' 13) . ACM, New York, NY, USA, 611-622.
- [33] Tamsyn P Waterhouse. 2013. Pay by the bit: an information-theoretic metric for collective human judgment. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW ' 13). ACM, New York, NY, USA, 623-638.
- [34] Mohamed Musthag, Deepak Ganesan. 2013. Labor dynamics in a mobile micro-task market. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI ' 13). ACM, New York, NY, USA, 641-650.
- [35] Cheong Ha Park, KyoungHee Son, Joon Hyub Lee, Seok-Hyung Bae. 2013. Crowd vs. crowd: large-scale cooperative design through open team competition. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 1275-1284.
- [36] Steven Dow, Elizabeth Gerber, Audris Wong. 2013. A pilot study of using crowds in the classroom. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI ' 13). ACM, New York, NY, USA, 227-236.
- [37] Kotaro Hara, Vicki Le, Jon Froehlich. 2013. Combining crowdsourcing and google street view to identify street-level accessibility problems. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI ' 13). ACM, New York, NY, USA, 631-640.
- [38] Walter S. Lasecki, Christopher D Miller, Jeffrey P Bigham. 2013. Warping time for more effective real-time crowdsourcing. In Proceedings of the 2013 ACM annual conference on Human factors in computing systems (CHI ' 13). ACM, New York, NY, USA, 2033-2036.
- [39] Lydia B Chilton, Greg Little, Darren Edge, Daniel S Weld, James A Landay. 2013. Cascade: crowdsourcing taxonomy creation. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI ' 13). ACM, New York, NY, USA, 1999-2008.
- [40] Parmit K Chilana, Andrew J Ko, Jacob O Wobbrock, Tovi Grossman. 2013. A multi- site field study of crowdsourced contextual help: usage and perspectives of end users and software teams. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI ' 13). ACM, New York, NY, USA, 217-226.
- [41] Adabriand Furtado, Nazareno Andrade, Nigini Oliveira, Francisco Brasileiro. 2013. Contributor profiles, their dynamics, and their importance in five q&a sites. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 1237-1252.
- [42] Rebecca Gray, Nicole B Ellison, Jessica Vitak, Cliff Lampe. 2013. Who wants to know?: question-asking and answering practices among facebook users. InProceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 1213-1224.
- [43] Kate Starbird. 2013. Delivering patients to sacré coeur: collective intelligence in digital volunteer communities. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI ' 13). ACM, New York, NY, USA, 801-810.
- [44] Uichin Lee, Jihyoun Kim, Eunhee Yi, Juyup Sung, Mario Gerla. 2013. Analyzing crowd workers in mobile pay-for-answer q&a. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems

- (CHI'13). ACM, New York, NY, USA, 533-542.
- [45] Erin L Brady, Yu Zhong, Meredith Ringel Morris, Jeffrey P Bigham. 2013. Investigating the appropriateness of social network question asking as a resource for blind users. In Proceedings of the 2013 conference on Computer supported cooperative work(CSCW'13). ACM, New York, NY, USA, 1225-1236.
- [46] Oded Nov, Ofer Arazy, Claudia López, Peter Brusilovsky. 2013. Exploring personality-targeted UI design in online social participation systems. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 361-370.
- [47] Paul Denny. 2013. The effect of virtual achievements on student engagement. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 763-772.
- [48] Elaine Massung, David Coyle, Kirsten F. Cater, Marc Jay, Chris Preist. 2013. Using crowdsourcing to support pro-environmental community activism. InProceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 371-380.
- [49] Daisy Yoo, John Zimmerman, Tad Hirsch. 2013. Probing bus stop for insights on transit co-design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 409-418.
- [50] Peter Dalsgaard, Eva Eriksson. 2013. Large-scale participation: a case study of a participatory approach to developing a new public library. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 399-408.
- [51] Jacki O'Neill, Shourya Roy, Antonietta Grasso, David Martin. 2013. Form digitization in BPO: from outsourcing to crowdsourcing? In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems(CHI'13). ACM, New York, NY, USA, 197-206.
- [52] Picard R W, Vyzas E, Healey J. Toward machine emotional intelligence: analysis of affective physiological state. *IEEE transactions on pattern analysis and machine intelligence*. 2001; 23(10): 1175 ~ 1191.
- [53] Wendy S Ark, D Christopher Dryer, Davia J Lu. The Emotion Mouse, *Proceedings of HCI International (the 8th International Conference on Human-Computer Interaction) on Human-Computer Interaction: Ergonomics and User Interfaces-Volume I*, 1999, p. 818-823, August 22-26.
- [54] Zhao Q L, Peng H, Hu B, Liu Q Y, Liu L, Qi Y B, Li L L. Improving Individual Identification in Security Check with an EEG based Biometric Solution. *BI* 2010.
- [55] Hareli S, Rafaeli A. Emotion cycles: On the social influence of emotion in organizations[J]. *Research in organizational behavior*, 2008, 28: 35-59.
- [56] Barsade, S G. The Ripple Effect: Emotional Contagion in Groups. Working Paper 98. Yale School of Management, Yale University, New Haven, CT.
- [57] Mohamed Ben Ammar, Mahmoud Neji, ADEL M. ALIMI et Guy Gouardères. (2007). Emotional Agents for Collaborative e-Learning. *International Research Journal on Digital Future. (FormaMente)*, V. 2 n. 3-4, pp. 33 ~ 64.
- [58] Mentis H M, Reddy M, Rosson M B. Invisible emotion: information and interaction in an emergency room [C] //Proceedings of the 2010 ACM conference on Computer supported cooperative work. ACM, 2010: 311-320.
- [59] Luneski A, Bamidis P D, Hitoglou-Antoniadou M. Affective computing and medical informatics: state of the art in emotion-aware medical applications[J]. *Studies in Health Technology and Informatics*, 2008, 136: 517.

- [60] Benjamin A, Birnholtz J, Baecker R, et al. Impression management work: How seniors with chronic pain address disruptions in their interactions[C] //Proceedings of the ACM 201 conference on Computer Supported Cooperative Work. ACM, 2012: 799-808.
- [61] Baumer E P S, Katz S J, Freeman J E, et al. Prescriptive persuasion and open-ended social awareness: Expanding the design space of mobile health[C] //Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. ACM, 2012: 475-484.
- [62] Jones A, Issroff K. Motivation and mobile devices: exploring the role of appropriation and coping strategies [J]. Research in Learning Technology, 2007, 15(3).
- [63] Birnholtz J, Ibara S. Tracking changes in collaborative writing: edits, visibility and group maintenance [C]. Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work. ACM, 2012: 809-818.
- [64] Magerkurth C, Cheok A D, Mandryk R L, et al. Pervasive games: bringing computer entertainment back to the real world[J]. Computers in Entertainment(CIE), 2005, 3(3): 4-4.
- [65] Wiberg M, Gronlund. A Exploring mobile CSCW: Five areas of question for fur ther research[J]. IRIS23, Sweden, 2000.
- [66] Biswas S, Morris R. Opportunistic routing in multihop wireless networks. ACM SIGCOOM Computer Communication Review, 2004, 34(1): 69-74.
- [67] Amin Vahdat, David Becker. Epidemic Routing for Partially-Connected Ad Hoc Networks. Technical Report CS-200006, Duke University, April 2000.
- [68] Anders Lindgren, Avri Doria, Olov Schelén. Probabilistic routing in intermittently connected networks. ACM SIGMOBILE Mobile Computing and Communications Review. Volume 7 Issue 3, July 2003.
- [69] T Spyropoulos, K Psounis, C S Raghavendra. Spray and focus: efficient mobility-assisted routing for heterogeneous and correlated mobility, in: Proceedings of Fifth IEEE International Conference on Pervasive Computing and Communications Workshops.
- [70] Augustin Chaintreau, Pan Hui, Jon Crowcroft, Christophe Diot, Richard Gass, James Scott. Impact of Human Mobility on the Design of Opportunistic Forwarding Algorithms. INFOCOM'2006.
- [71] Pan Hui, Jon Crowcroft, Eiko Yoneki. BUBBLE Rap: Social-based Forwarding in Delay Tolerant Networks. IEEE Trans. Mob. Comput. 10(11): 1576-1589(2011).
- [72] E Daly, M Haahr. Social network analysis for routing in disconnected delay-tolerant MANETs, in: Proceedings of MobiHoc 2007, pp. 32-40.
- [73] 牛建伟, 周兴, 刘燕, 孙利民, 马建. 一种基于社区机会网络的消息传输算法[J]. 计算机研究与发展, 2009, 46(12).
- [74] Li Wang, Xue Geng. A Community-driven Hierarchical Message Transmission Scheme in Opportunistic Networks. Smart Computing Review[J] vol. 1, no. 1, October 2011 p85-94.
- [75] Nam P Nguyen, Thang N Dinh, Ying Xuan, My T Thai. Adaptive Algorithms for Detecting Community Structure in Dynamic Social Networks. Infocom 2011.
- [76] M J Williams, R M Whitaker, S M Allen. Decentralised detection of periodic encounter communities in opportunistic networks. Ad Hoc Networks Available online 23 July 2011.
- [77] Sabrina Gaito, Elena Pagani, Gian Paolo Rossi. Strangers help friends to communicate in opportunistic networks. Computer Networks. 55(2011)374-385.
- [78] Hossmann T, Spyropoulos T, Legendre F. A complex network analysis of human mobility. Computer Com-

- munications Workshops(INFOCOM WKSHPS), 2011 INFCOMW P876 - 881.
- [79] Shizhen Zhao, Luoyi Fu, Xinbing Wang. Fundamental Relationship between Node Density and Delay in Wireless Ad Hoc Networks with Unreliable Links. The 17th Annual International Conference on Mobile Computing and Networking MOBICOM2011.
- [80] Yong Li, Pan Hui, Depeng Jin, Li Su, Lieguang Zeng. Evaluating the Impact of Social Selfishness on the Epidemic Routing in Delay Tolerant Networks. IEEE Communications Letters, 2010, 14 (11) : 1026-1028.
- [81] Peiyan Yuan, Huadong Ma, Xufei Mao. The dissemination speed of correlated messages in opportunistic networks. IEEE Symposium on Computers and Communications(ISCC), 2011 309-315.
- [82] Musoles i M, Mascolo C, Hailes S. EM MA: Epidemic Messaging Middleware for Ad hoc networks[C] // Personal and Ubiquitous Computing. 2005.
- [83] Murphy A, Picco G, Roman G-C. Lime: A Middle ware for Physical and Logical Mobility [C]. International Conference on Distributd Computing System. 2001 : 524-536.
- [84] Fok C-L, Roman G-C, Hackmann G. A light weight coordination middle ware for mobile computing [C]. Proceedings of COORDINATION 2004. Volume 2949 of LNCS. , Springer Verlag, 2004 : 135- 151.
- [85] Mascolo C, Capra L, Zachariadis S, et al. Emmerich. XMIDDLE: A Data-Sharing Middle ware for Mobile Computing. Wireless Personal Communications. 2002, 21(1) : 77-103.
- [86] Levis P, Culler D Matce. A Tiny Virtual Machine for Sensor Networks[C]. International Conference on Architectural Support f or Programming Languages and Operating Systems. San Jose, CA, U SA, Oct. 2002.
- [87] Curino C, Giani M, Gioret t a M, et al. TinyLIME: Bridging mobile and sensor networks through middle ware[C]. 3rd IEEE International Conference on Pervasive Computing and Communications(PerCom). IEEE Computer Society, March 2005 : 61-72.
- [88] Sil va V H Z, la Cruz Salgado E I D, Quintana F R. The AWARE Architecture: Supporting Context Mediated Social Awareness in Mobile Cooperation[C]. Proceedings of the 2004 ACM CSCW. ACM Press, 2004 : 192-201.
- [89] Li Yushun, Gong Neng, Shi Meilin. A New Collaborative Awareness Model and Its Application[C]. 8th International Conference on Computer Supported Cooperative Work in Design(IEEE Cat. No. 04 EX709). 2004, 1 : 53-58.
- [90] Bo Jiang, Haixiang Zhang, Chun Chen, Jianxv Yang. Enable collaborative graphics editing in mobile environment. CSCWD(1)2005 : 313-316.
- [91] Jiajun Bu, Bo Jiang, Chun Chen, Jianxv Yang: On Demand Consistency Maintenance in Heterogeneous Collaborative Graphics Editing Systems. CSCWD 2006 : 539-543.
- [92] Tim Paek, Maneesh Agrawala, etc. Toward Universal Mobile Interaktion for Shared DisPlays, in Proeeding of CSCW04, 2004.
- [93] Marco Conti, Sajal K Das, Chatschik Bisikian, Mohan Kumar, Lionel M. Ni, Andrea Passarella, George Roussos, Gerhard Tröster, Gene Tsudik, Franco Zambonelli. Looking ahead in pervasive computing: Challenges and opportunities in the era of cyber- physical convergence Pervasive and Mobile Computing 8(2012)2-21. February 2012.
- [94] Marco Conti, Mohan Kumar. Opportunities in Opportunistic Computing. IEEE Computer 2010, 43 (1) : 42-50.
- [95] 郭斌, 於志文, 张大庆, 周兴社. 机会物联——兼谈物联网的社会性. 中国计算机学会通信 . 2011, 12 : 50-54.

- [96] Wikipedia, the free encyclopedia. Computer-supported cooperative work (CSCW) [DB/OL]. [2012], http://en.wikipedia.org/wiki/Computer-supported_cooperative_work.
- [97] ACM Computer-supported cooperative work (CSCW)'13 Homepage [DB/OL]. [2012], <http://cscw.acm.org>.
- [98] Computer-supported cooperative work (CSCW) Review Process Statement [DB/OL]. [2012], <http://cscw.acm.org/CSCW-review-process-statement.pdf>.
- [99] Steven Poltrock, CarlaSimone, Jonathan Grudin, et al. Computer Supported Cooperative work In: Proceeding Computer-supported cooperative work (CSCW)'12. ACM, 2012.

作者简介

胡斌 CCF理事、协同计算专委会主任，兰州大学教授。主要研究方向为协同计算和情感计算。



卢暾 博士，复旦大学计算机科学技术学院副教授，CCF会员，协同计算专委会委员。目前主要研究兴趣为计算机支持的协同工作(CSCW)、人机交互(HCI)、协同计算、云计算。



王莉 博士，太原理工大学计算机科学与技术学院教授，中科院计算所博士后，CCF高级会员。目前主要研究兴趣为社会网络分析、移动网络通信、移动社会网络等。



汤庸 CCF杰出会员，华南师范大学计算机学院院长，广东省服务计算工程计算研究中心主任，二级教授、博士生导师。主要研究领域为数据库、云服务、协同计算与社交网络。



孙宇清 博士，山东大学计算机科学与技术学院教授，CCF高级会员、协同计算专委会秘书长，系统软件专委会委员。主要研究方向为协同计算、系统安全、隐私保护。



张常有 博士，中国科学院软件研究所并行软件与计算科学实验室高级工程师。CCF 会员，CCF YOCSEF 委员，CCF 协同计算专委会委员，CCF 高性能计算专委会委员。目前主要研究兴趣为社会网络分析、并行与分布式计算。



唐卫清 博士，中国科学院计算技术研究所研究员，CCF 高级会员、常务理事、协同计算专委会副主任。目前主要研究方向为流程工厂的协同设计与管理。



李建国 2010 年获中山大学计算机软件与理论博士学位，现任华南师范大学计算机学院副教授，主要研究领域为协同计算与社交网络。



陈国华 2011 年获中山大学计算机软件与理论博士学位，同年 12 月进入华南师范大学博士后流动站，主要研究领域为信息检索与社交网络。



夏欢欢 复旦大学计算机科学技术学院博士生，主要研究方向为计算机支持的协同工作（CSCW）、乐观复制技术等。



黄 兴 复旦大学计算机科学技术学院博士生，主要研究方向为计算机支持的协同工作（CSCW）、数据世系与科学软件分享等。



多模态物理过程仿真研究进展与趋势

CCF 虚拟现实与可视化技术专业委员会

摘要

多模态物理过程仿真是指采用物理学方法对多种模态物质（如气体、液体、固体）间的耦合作用过程进行仿真，多模态物理过程仿真在计算机仿真、虚拟现实、工业制造和医疗等领域具有重要研究意义和应用价值。本报告通过对近几年来国内外多模态物理过程仿真的主要研究方法和研究成果进行分析与比较，总结了国内外多模态物理过程仿真的研究进展，并针对最新的科技发展，给出多模态物理过程仿真的研究发展趋势。

关键词：多模态物理过程，耦合仿真，虚拟现实

Abstract

Multimodal physical process simulation refers to the simulation of coupling process of multimodal materials (gas, liquid, solid) by using physics-based methods and it has important research significance and application value in computer simulation, virtual reality, industry manufacturing and medical fields, etc. This report analyzes and compares the main research methods and achievements of multimodal physical process simulation at home and abroad in recent years, summarizing the research progress of multimodal physical process simulation and giving the research trends of multimodal physical process simulation according to the latest development of science and technology.

Keywords: multimodal physical process, coupling simulation, virtual reality

1 引言

随着计算机仿真计算和虚拟现实技术的快速发展，近年来气体、液体、固体多种模态物质之间的物理过程仿真得到了快速的发展，在工业制造、医疗卫生、灾害防治、数字娱乐等领域具有重要研究意义和应用价值。多模态物理过程仿真提供了一种方便有效的方法对真实世界的复杂情景进行模拟、预测及评估。例如，虚拟手术应用中腔内气压变化、组织形变、流血复杂环境的手术场景模拟仿真，用于培训外科医生的手术操作；高温燃气作用下金属材料受力的数值模拟；流固耦合作用下上呼吸道内稳态气流运动特性的数值模拟等。

在多模态物理过程中，气体、液体和固体的建模及数值求解一直是人们重点关注的核心研究内容。近年来，研究人员提出了许多基于物理的方法用于气体、液体和固体之间的多模态物理过程仿真。它一般分为两类，一类是基于网格的方法或欧拉方法，另一类是基于粒子的方法或 Lagrangian 方法。

基于网格的方法如有限元法（Finite Element Method, FEM）等使用较为广泛，有限元模型是一种连续参数化模型，具有较高的精度，尤其对于几何形状极为复杂的区域，通过恰当地调节参数和构造不同形状的单元，得到几何和力学上都较为精确的模型^[1,2]。有限元模型计算量较大，实时性难以保证，虽然随着算法的不断改进和模型的不断优化，计算效率也在不断地提高。但是，其缺点也是显而易见的。对于三维问题，形成 FEM 网格时的计算成本太高，且自动生成不影响 FEM 分析质量的网格难以实现；由于 FEM 公式中的假设位移场是分片（或分单元）连续的，虽然计算得到的位移能够保证连续性，但是应力、应变通常有间断点，因此需要后处理修匀。在虚拟手术中，经常发生大变形情况或切割、缝合等操作，这时器官的拓扑结构会发生变化，采用有限元法需要不断进行网格重构，不仅耗时耗力，而且容易出现网格畸变问题，大大影响了 FEM 的求解精度。因此，后处理阶段需要采用某些特殊计算以提高其应力的精确性^[3]。FEM 在处理自由表面问题、边界运动问题、表面运动问题和大形变及碰撞传播问题时并不适用，并且复杂几何模型的高质量网格生成也是困难耗时的过程，无网格法则能较好地解决这些问题^[4]。

基于粒子的方法如光滑质点流体动力学（Smoothed Particle Hydrodynamics, SPH）是近年来无网格法中使用最广泛的一种。无网格法主要是将问题域分解为一系列任意分布的粒子或结点，以求解具有各种边界条件的积分方程和偏微分方程，从而得到精确稳定的数值解。无网格法因其没有网格依赖性的特点，常用于液体和气体的形变模拟，并取得了很好的效果^[5~7]。近年来，由于有限元方法的广泛应用，采用有限元方法结合无网格法处理流固耦合现象的方法层出不穷^[8]，但是它无法规整网格法和无网格法间的整体差异，尤其在处理两种模型接口过程中计算复杂度高的问题始终难以解决。

随着现代经济社会的发展，传统的串行编程技术已经不能满足科学和工程等领域迅猛增长的高性能计算需求，大规模并行计算已成为科学和工程计算等领域的发展趋势^[9]。多模态物理过程仿真的求解规模越来越大，求解精度越来越高。为了实现快速高效的多模态物理过程仿真，引入大规模并行计算技术能实现快速精确的仿真效果。

本报告通过对近几年来国内外多模态物理过程仿真的主要研究方法和研究成果进行分析与比较，总结了国内外多模态物理过程仿真的研究进展，并针对科学计算的最新发展状况，给出了多模态物理过程仿真的研究趋势。

2 国际研究现状

国外有较多学者对多模态物理过程仿真进行了深入的研究，针对气体、液体和固体的建模与仿真，国外学者提出了许多行之有效的方法。

对于气体仿真，气体属于可变形流体现象，外观形状极不规则，没有光滑的表面，且运动规律受外界影响明显。Fedkiw 等人^[10]提出了一种新的数值烟雾仿真方法，针对传

统的网格方法比较粗糙的缺陷，这种方法利用了烟雾的物理学特性构造更为精细的网格，运用非粘性的欧拉方程对烟雾气体进行仿真，非粘性欧拉方程相比于粘性的纳维-斯托克斯方程更适合于气体模型仿真，且计算量更小。另外，这种方法为了对烟雾小规模旋转特征进行建模，还引入了一般粗糙网格仿真所不具有的物理连续的漩涡限制条件。Foster 和 Metaxas^[11]采用显式积分方法生成了三维烟雾的漩涡运动，得到了极其逼真的效果，但气体在作用区域中速度较大时，运行速度相对较慢。Stam 引入了拉格朗日法求解对流项^[12]，其求解无条件稳定并允许大时间步长。H. Schechter^[13]等人介绍了一种简单的模拟烟雾动画的湍流模型，该模型能在一个经典仿真网格上，捕获湍流能量的传输、扩散以及它们在光谱上的重叠关系。对于每个网格单元，跟踪平均处于活动状态的八个湍流元，并且用一个新的网络旋转（Net Rotation）变量来对湍流漩涡的自水平运动进行建模。这些对标准流体的求解程序所附加的条件驱动的问题设置了后置处理，在大尺寸动画仿真问题上，动态的分层使得动画细节得以展现。另外，为了在转向规范子网格模型前充分利用已有仿真网格，H. Schechter 等还提出了一种新的多时间步预测方法，该方法能减少在标准图形学流体求解程序中由于角度动量所带来的非物理耗费。Song 等人通过采用约束插值剖面法（Constrained Interpolation Profile）来提高空间精度^[14]，但增加了执行复杂性和计算时间。Kim 等提出小波湍流的流体模拟方法^[15]。现有算法都是采用 N-S（Navier-Stokes）方程描述烟雾等物体的运动，严格遵循物理规律的约束，能较精确地模拟出真实效果。用 SPH 方法模拟高粘度的流体，在较大的外力作用下，为了进行稳定的流体仿真并得到较好的结果，需要对时间步长设置得很小。Francois 等人^[98]采用迭代预测-修正方法来优化施加在流体上的作用力。当流体粘度非常大甚至接近于刚体时，这种方法能很大程度上减少计算次数，在每一个时间步能预测出下一时刻的粒子位置。在变形过程中，将计算出的变形结果同目标变形结果进行比较，通过比较结果修正施加的作用力。

在固体建模方面，M. Müller 提出了一种基于粒子的方法并借助 SPH 方法对弹性体、塑料体和融化物体进行建模仿真^[4]，它能有效地仿真固体的形变。随后 Müller 提出基于位置的动力学过程^[16]，通过将模型表示成粒子系统，研究粒子间相对位置对于粒子速度、位移等的影响，从而计算出粒子的运动规律。2011 年，M. Müller 又提出了用存有方向和旋转信息的粒子来解决在小粒子团中模型形状失真的问题^[17]，在计算粒子弹力和形变向量场时保存粒子的运动性质，以确保模型不失真。该方法保证了粒子不依赖于网格，但是难以保证物理参数的准确性，它不能应用于对于物理精确性要求较高的多模态物理过程仿真。

之后关于固体和流体（液体和气体）建模的研究大多是基于 M. Müller 的工作，相关改进多集中于计算的简化^[18]和稳定性的提高^[19]。一些方法简单地把流体、固体粒子视为整体来计算邻域^[19]，但这将导致流体粒子进入固体的不真实现象。虚粒子法^[20]在固体、液体周围设置虚粒子用于耦合，但这并不能仿真双向物理过程，且虚粒子的设置与计算过程相当复杂。Afonso 等人^[21]提出了一种实现具有良好视觉效果的流固交互现象的方法，但是该方法未能实现物理参数测量，不具有物理真实性。Becker^[22]将有限元中的

同步旋转方法应用于光滑粒子流体动力学 SPH 方法中取得了较好的效果，但是由于 SPH 方法粒子的不连续性，固体边界不能保持平滑。基于位置的方法成功应用于固体与流体仿真中^[23,24]，该方法直接对物体位置或网格顶点进行控制，其主要优势是可控性且碰撞限制能够简便处理，并能完全处理穿透现象。虽然基于位置的方法有着视觉似然性，但是这类方法不能像基于力的方法那样精确。在基于力的方法中，内力与外力累积起来便可加速其计算，SPH 的粒子特性使得其非常适用于变形固体及液体的仿真^[25~28]。

最近的多模态物理过程仿真研究中，针对流体的仿真工作大多集中于不可压缩流体的仿真与效率方面。Solenthaler 等人^[29]提出了预测修正模型用于保证不可压缩性，将加速分为算法加速与硬件加速。Adams 和 Solenthaler 等人^[30,31]提出了多尺度概念。Ihmsen 等人^[32]则采用 GPU 和 CPU 联合并行加速。Müller 等人^[4]和 Richard Keiser 等人^[33]阐述了两种方法，分别构建可形变固体模型和刚性固体模型，通过可形变固体内部的应变能来计算可形变固体粒子间相互作用的弹性力，从而实现了一种可形变固体粒子仿真模型；而对于刚性固体，由于其仅有平移、旋转两种运动方式，故能计算其转矩得到刚体运动的角速度及更新位置。

在多模态物理过程仿真应用方面，国外学者也进行了较多的研究。气-液-固的耦合对医学领域的仿真有着举足轻重的作用。如在心脏功能方面，这种两相或多相耦合能使人们更加深入地理解心脏的作用。理解固-液耦合的基础反馈机制以及该机制在心脏功能中所扮演的角色对于塑造正常心脏功能和在疾病中的表现有着极其重要的意义。D. Nordsletten 等人^[34]为了强化这种理解，构造了一种精确的基于解剖学的固-液力学计算模型用于评定心脏的被动舒张期和主动收缩阶段的特性，为了整合描述心脏血动力和组织机械性能的多重数据，该方法允许最佳有限元组合中的不规则性。这种方法主要用来观察左心室固液耦合情况，能对左心室中血液流动和压力分布以及由于粘滞扩散造成的机械能量的激活和损耗进行定量的评价。Sebastian 等人^[35]提出了一种三维 Code-coupling 方法来模拟心血管中的血液流动情况，该方法用心肌复合模型来替代表示心血管流动情况的 KaHMo (Karlsruhe Heart Model) 模型，这种模型是通过流体作用力和心肌的反应来产生变形效果，利用有限体元方法 (Finite Volume Method, FVM) 对 N-S 方程的任意拉格朗日 - 欧拉公式进行离散化，利用有限元方法对固体力学有限弹性方程进行离散化，并在病人核磁共振图像集的基础上模拟左心室的流体结构间的相互作用情况。实验结果表明，采用这种方法得到的流体模型与实际观察得到的结果在很大程度上保持了一致性。

Rahul 等人^[36]提出了能够有效模拟颗粒物质的连续一致模型，该模型能充分地模拟颗粒物质间的内部压强和摩擦所产生的应力情况，从而使颗粒物质再生运动过程能可视化地展示出来。这些颗粒的再生过程包括在没有内聚力的情况下颗粒的自由扩散飞溅运动，以及摩擦力和内部压强间的耦合行为，内部压强的处理方式使颗粒物质能够与固体间进行双向作用，在模拟颗粒流动情况中，与基于粒子的模型相比，这种方法显著地减小了计算代价。

带有尖锐边缘的固体与流体的耦合运动是仿真一个重要研究内容。Michelin 等人^[37]

提出了一个降阶控制模型，它用于处理带有尖锐边缘的固体和高雷诺兹指数的流体间的二维相互作用，该模型能应用于模拟刚体卡片掉落在流体中的情况。

D. Gregory 等人^[38]考虑了在冲击波作用下的固液相互作用问题，其中固体产生了大体积的拉格朗日位移，在非结构化（松散的）流体网格内嵌入固液接口，使用一个水平集（等值面）作为固液耦合的一般性接口。该文提出了一种基于边缘的非结构化网格欧拉求解公式以及在非结构化液体网格上的固液接口识别问题，它采用超级采样投射技术结合拉格朗日接口的位置，能快速地识别接口和响应边界条件。接口的识别采用了窄带方法并能通过接口结构的细节展示出来；该文还提出了一系列二维、三维冲击体的计算来论证方法在静力学、动力学接口问题上的有效性，包括抛射体/冲击体相互作用的仿真。

针对二维和三维空间上的流固耦合问题，Yen 等人^[39]提出了一种简单且有效的流固耦合方法，该方法考虑了 N-S 方程在不规则域的数值逼近问题并解决了任意形状域的霍奇投影步问题。该方法为投影步和粘度的隐性处理生成了一种易于实现的对称的正定线性系统。Yen 等人采用该方法模拟了在二维空间中的流体流经柱体的情况，并且使用该方法来处理流体和移动刚体间的耦合作用。

然而，当前的仿真方法在处理流体和具有不规则边界形状的固体耦合问题上，总体效率较低且鲁棒性差。Christopher Batty 等人^[40]提出了一种能在粗糙笛卡儿网格下鲁棒并精确求解问题的方法，该方法能加快仿真模拟的速度。它通过将经典压强投影步改进为最小活动能量，该方法提出了一种鲁棒的流体与任意固体间的耦合仿真模型，该模型给出了适应相应问题的对称半正定线性系统。另外，Christopher Batty 等人还为自由边界流体问题扩展了一种新的边界限制条件框架，这种框架能允许流体自然真实地与固体分离。

在流变学、摩擦学和流体力学中，无滑移是一种基本的假设且被人们广泛接受的边界条件。然而，在很多场景中，尤其是非牛顿流体流动中，边界条件会经常出现相互矛盾的情况。如在流体喷射流动等场景模拟中，壁面滑移会产生很大的错误并导致流体流动不稳定，壁面滑移还会使流体系统分析复杂化，从而导致在实际情况下难以使用，Taha 等人^[41]就流-固相互作用中的滑移问题展开了深入的讨论。

F. Ilinca 等人^[42]提出了用沉浸式边界方法解决流固相互作用的问题。固体沉浸在流体中的运动情况是根据固体与流体间相互作用力情况而定。传统的方法是采用正多边形网格来模拟流体与固体的相互作用，这种网格在域形状发生改变或者被沉浸的固体位置发生改变时，网格需要重新适应这种改变，重新适应过程会耗费大量的时间，在连续网格间生成网格或插入新的网格也会带来巨大的时间耗费，若网格几何变形程度很大就会产生严重的错误。如果采用沉浸物体的方法，网格不需要在每个时间步长内去适应沉浸物的形状，从而能避免传统网格方法的缺陷。为了计算流体和沉浸的固体的相互作用力，F. Ilinca 提出了一种改进的沉浸边界（Immersed Boundary, IB）有限元方法，它采用固定的网格覆盖流体和固体区域，而沉浸固体的边界用随时间变化的水平集函数表示，沉浸固体表面的边界限制用 BCE（Body Conformal Enrichment）方法处理。

在气-固两相耦合应用如在能源生产工业中，热能传输是其中一项非常重要的内容。

热能传输的计算流体力学（Computational Fluid Dynamics, CFD）仿真是基于统计学的相关理论，它有一系列的平均方程，如欧拉 - 欧拉两流体模型。这些平均方程要求一些精确的模型参数，如平均气 - 固热能传输率，该参数是根据努塞尔特数进行建模求解的，在构造求解平均气 - 固热能传输率的闭合模型时，假设气 - 固两相流是局部同质的。然而，沿着流体流动方向持续地进行加热或冷却就会引起流体温度场的不同质。S. Tenneti 等人^[43]提出了一种粒子分解直接数值仿真方法（Particle-Resolved Direct Numerical Simulation, PR-DNS）来研究粒子数固定时的稳定流体流动时热能传递问题。通过类比管道中的完全展开流，文中提出了一种在平均流体温度场不同质环境下的热相似条件来确保统计学上的同质努塞尔特数。从 PR-DNS 实验结果看出，在两流体模型中的尺度分离隐式假设并不总是有效的。

Franck Lomine 等人^[44]提出了用于处理流体 - 固体交互耦合的数值方法，并用该方法模拟出粒子 - 流体系统。这种数值方法能广泛用于颗粒介质力学的离散元方法（Discrete Element, DE）和应用于流体动力学的格子玻尔兹曼方法（Lattice Boltzmann, LB）进行结合の場合，离散元方法用于对粒子间的相互作用进行建模，格子玻尔兹曼方法则用来表现间质牛顿流体的流动情况。这两种方法的结合能充分展现出作用于流体上的粒子的运动情况。文中采用这种综合数值仿真方法对管道的腐蚀进行了模拟，实验证明 LB-DE 结合方法是研究地质力学混合问题很好的工具。

在流体表面现象中，气泡和泡沫是一项重要的研究内容。但真实的气泡边缘膜很薄，并且相互作用关系复杂，导致难以模拟。在稠密泡沫中的小气泡（直径小于 1cm）很容易被表面张力所影响，因此小气泡与大气泡相比更容易变形。在上述小气泡假设的基础上，Oleksiy 等人^[99]提出了一种更精确、更有效的基于粒子的仿真算法来模拟气泡的动态变化和相互作用，该方法的核心在于通过将气泡粒子当作有权值的泰森多边形角点对泡沫的几何形状进行近似。在前面框架的基础上，Oleksiy 等人提出了一种气泡相互作用力集来处理复杂的泡沫行为，这些复杂行为包括 Plateau 法则下泡沫结构、流体表面气泡的簇拥、气泡 - 流体和气泡 - 固体耦合以及气泡喷发聚集等，实验表明该方法能与当前的流体方法直接组合，产生逼真的气泡模拟效果。

Markus 等人^[100]提出了一种用于物质扩散的新模型，该模型能与基于粒子的流体方法结合。扩散物质现象一般表现为喷射、泡沫和气泡等，在产生、传输及驱散这些扩散物质时，所采用的是基于物理的法则。该方法也是后置处理的一种通用且有效的方法。随着粒子间作用力和物质扩散到流体的影响被忽略，大量的扩散物质能有效地实现大尺度变形中高精细度的模拟仿真。

3 国内研究进展

针对多模态（固体、液体、气体）物理过程仿真，国内学者在诸多应用领域也进行了较深入的相关基础研究及应用研究工作。王丽等人^[45]通过联立偏微分方程组的控制方

程进行差分化来对泥石流进行数值模拟。陈日东等人^[46]采用处理对流项有较好计算稳定性的有限元特征分裂算法建立了与主河汇流的泥石流堆积数学模型，再现了特大山洪泥石流的堆积过程。杨志亮等人^[47]提出了一种基于粒子的固液两相流模型对泥石流的运动进行建模，并采用隐式表面定义来渲染液相浆体部分。

石莹等人^[48]提出了基于 Fluent 的气固喷射器三维数值模拟研究中根据 CFD 方法和仿真思想，采用了标准 K-8 湍流模型并利用拉格朗日法与欧拉法相结合的三维数值模拟技术对气固喷射器内的流场和颗粒行为进行数值模拟；针对固体颗粒，采用了离散相模型用于观察气固两相间的相互耦合作用。王勇等人^[49]在冲击载荷液压缸内压的流固耦合仿真与振动分析中采用了有限元分析法对内部流场压力进行流固耦合分析，并采用动力学波动方程对相同模型的内部压力进行分析。

孙剑韬等人^[50]在流固耦合分析中借助有限元分析法，采用 ANSYS Workbench 和 CFX 及 ANSYS FLOTTRAN 相结合的算法实现流体分析和结构分析的单向流固耦合计算和双向流固耦合计算，并对两种流固耦合方法进行分析和比较。在高速气流驱动下固体结构的变形相对较大，从而导致流场的边界形貌发生较大的改变，宜采用双向流固耦合方法进行仿真分析。任意拉格朗日欧拉法求解流固耦合的一个优点是它能实现欧拉网格的移动，这样就能使流体网格随着固体的变形而改变。

落葵寿等人^[51]在整流罩地面分离试验流固耦合分析与数值模拟中采用拉格朗日乘子法建立挠性整流罩分离动力学模型。基于该模型，借助 ADAMS (Automatic Dynamic Analysis of Mechanical Systems) 软件对整流罩的分离过程进行了仿真，初步得出不计气动作用时整流罩的分离过程和相关特性。该文还从气动理论出发，分析了大气阻力的作用机理，建立了整流罩地面分离试验的气动模型。基于该模型，以不计气动阻力作用时整流罩分离过程的运动特性作为初始条件，采用弹簧光顺与局部重划相结合的非结构化网格以及基于有限体积思想的 SIMPLE 算法，利用商业软件 Fluent 仿真计算了其分离过程中周围大气的阻力作用。

夏生林等人^[52]在基于 CFD 非线性气动力学的静气弹计算中，采用基于 CFD 非线性气动力静气弹计算方法，在飞机概念和初步设计阶段进行了静气动弹性的计算和分析工作，该方法通过 CFD 仿真技术对具有复杂翼型的飞机进行气动力学仿真，并采用成熟的流固耦合技术进行了基于气动力学的气动弹性静稳定性计算分析，以便快速地得出可靠的静气动弹性响应结果，为气动布局设计提供有价值的输入量。其数学模型包括平板气动模型和三维结构有限元模型。首先进行 CFD 非线性气动力进行三维样条插值和弦平面法向投影等数值计算，再通过气动/结构插值矩阵，得到基于结构的气动力影响系数矩阵，进而构成静气动弹性响应分析方程，实现静气动弹性的计算分析。基于 CFD 非线性气动力的静气弹性计算技术为前期初步设计阶段提供了快速、准确、可靠的静气弹性计算分析手段。在实际工程应用中，采用此方法的耗时大约为其他商用软件的 1/10，而计算结果满足前期设计要求，其三维欧拉方程数学模型设计简单，能在前期设计阶段中进行多轮气动布局优化设计，并能为结构的优化设计提供快速和准确的参考数据。

孙栋等人^[53]通过构建流固耦合力学模型，对流固耦合作用下包括前三级支气管在内

的上呼吸道内稳态气流运动特性进行数值模拟，系统地分析了流固耦合作用下上呼吸道壁面的形变特点、壁面剪切应力分布以及呼吸道内的气流运动特点。胡海峰等人^[54]采用流固耦合中的流动与换热模型实现了温度场解算与流场解算的耦合数值分析。姜海晨等人^[55]结合流-固耦合有限元的基础理论，利用有限元分析软件建立了阀系的流固耦合模型，对多片叠加节流阀片进行了流固耦合有限元分析，给出了减振器节流阀片的瞬态非线性特性、减振器节流阀系流场特性以及节流阀片动态响应特性。

郭军刚等人^[56]在计算时考虑了燃气流动与传热对离心拉伸应力的影响，体现各部分之间的相互影响关系以及从全局反映叶片的工作状态。该文基于 ANSYS 分析软件，应用材料非线性本构关系以及弹塑性力学相关理论对涡轮转子叶片在高温燃气作用下受离心拉应力的情况进行了数值模拟。崔涛等人^[57]通过参数传递以及同步控制实现了列车系统动力学与 CFD 间结合的联合同步仿真，但忽略了考虑轨道结构对列车动力学的影响。刘辉等人^[58]同时考虑油液、空腔与箱体的耦合作用，建立了油、固、气耦合的仿真模型，对在发动机激励和齿轮啮合激励作用下箱体振动产生的外部辐射噪声进行了仿真计算，从而更精确地预测出变速箱的辐射噪声。李小川等人^[59]以节流型自激式水幕除尘器为研究对象，通过紊流状况下除尘器脉动压力均值的测量，对除尘器内部气-液耦合状况进行了研究，得出接触腔的高度直接影响除尘器气-水两相耦合程度，进而影响除尘器除尘的效果。张广明等人^[60]基于多孔介质流-固耦合的控制方程，推导了相应的非线性增量有限元列式，并采用以结点位移和孔隙压力为自由度的黏结单元来模拟水力损伤造成的预设裂缝的起裂和扩展，在此基础上建立了油井水力压裂有限元模型。刘小民等人^[61]通过对结构和流体介质进行双向流固耦合的瞬态数值模拟，获得了罐体结构应力的时间历程，从而准确揭示了罐体在时间历程中最大应力分布情况。高兴龙等人^[62]基于任意拉格朗日-欧拉有限元方法对开缝救生伞进行了充气过程的流固耦合问题研究，建立了降落伞结构以及流场的计算力学和数值仿真模型，并利用任意拉格朗日-欧拉耦合方法计算得出了流固耦合作用下有缝降落伞的伞衣投影面积和充气过载的数值仿真结果。

彭正标等人^[63]在采用欧拉法处理气相场的同时利用基于拉格朗日法的直接数值模拟蒙特卡罗（Direct Simulation Monte Carlo, DSMC）法处理离散颗粒场，对循环流化床烟气脱硫塔内气固两相流动特性及其优化进行了数值模拟。文中使用非结构网格数值求解，采用了四向耦合方法严格地考虑颗粒-流体、颗粒-颗粒、颗粒-器壁间的相互作用，并应用高效的非结构网格搜索法，以实现气固相间映射的耦合作用与反馈。得到了加装导流板组前后、脱硫塔内气固流动特征、轴向速度分布以及固相颗粒浓度分布，同时对影响颗粒停留时间及颗粒碰撞率的关键因素进行了分析。

杨玉新等人^[64]基于流-固耦合方法，在充分考虑混合火箭发动机工作过程中诸多复杂物理过程的基础上，建立了一种可适用于不同工作状况下混合火箭发动机固体燃料表面退移速率预示的计算模型。计算结果与实验数据的对比验证了所建立计算模型的准确性。对模型发动机进行模拟的结果表明，混合火箭发动机中的燃烧、流动及固体燃料表面的退移速率具有明显的不均匀性，发动机中的固体燃料表面的退移速率沿轴向近似地呈“W”形状的曲线变化；在混合发动机中，突扩形状的预燃室和补燃室有利于燃料热

解气体和氧化剂气体的扩散混合可强化对固体燃料表面的换热，从而提高固体燃料表面的退移速率。孙可明等人^[65]针对深部煤层处于较高的应力和孔隙压力环境下，煤层渗透率降低和瓦斯运移表现出非达西渗流及受煤岩体变形耦合作用明显的特点，建立了反映深部低渗透煤层特征和瓦斯流动特性的气液两相流流固耦合模型。以此耦合模型为基础，通过数值模拟对深部低渗透井群开采煤层气进行了较系统地研究，得出不同渗透率和不同井群间距条件下开采煤层气的储层压力、气和水产能大小、压降漏斗、水饱和度、甲烷浓度和井群干扰的变化规律，考虑较高的应力储层变形影响的流固耦合模型模拟结果比不考虑耦合作用的结果偏小，对于深部煤层气开采必须重视耦合作用对产量造成不利影响，制定合理的生产制度和布井方案，尽可能确保储层渗透特性受弱化的程度最小。

程载斌等人^[66]应用 LS-DYNA 显式程序建立了三种包含水体、空气、导弹、发射井和筒盖的多物质耦合 ALE 网格模型，使用罚函数流体 – 固体耦合方法对导弹水下无攻角潜射过程进行了数值模拟，给出了弹体质心的轴向加速度、轴向流体阻力及阻力系数、空泡区压力分布。张俊鸽等人^[67]借助流体动力学计算软件 CFD2000 对介质不同流速下试样的淬火过程进行了流固耦合数值模拟。将实测换热系数曲线数值化并线性插值得到了不同温度、流速下的换热系数，并赋给流 – 固换热边界，从而实现流固耦合计算，同时得到了淬火介质流场和试样温度场。数值模拟结果与实测结果的对比表明，用该数值方法来模拟淬火介质存在相变的复杂淬火过程是可行的，能够显示出淬火过程的蒸汽膜、沸腾换热、对流冷却三个阶段以及各自的特点。

陈明等人^[68]在针对工程中的多相流动问题，以最常见的气液固三相流为对象，建立了一种考虑流固耦合效应的水力瞬变模型。通过改变流体中固相或气相的体积分数，模型可适用于气液、固液两相流以及纯液体单相流的耦合水力瞬变，它是单相流耦合瞬变模型向多相流耦合瞬变模型的拓展。曹树刚等人^[69]讨论了煤层瓦斯流动特性，认为游离瓦斯和吸附瓦斯、突出空洞内的瓦斯和突出空洞附近煤体中的瓦斯均在煤与瓦斯突出过程中起重要作用，因而提出原煤吸附瓦斯贡献系数 k 并建立了煤层瓦斯流动的质量守恒方程；基于煤岩流变力学实验，提出了讨论煤岩流变力学性质的广义弹粘塑性组合模型。最后，建立了可以用来研究煤与瓦斯延迟突出机理的含瓦斯煤固 – 气耦合分析的数学模型。

朱洪来等人^[70]分析了在流固耦合工程实际问题中的大多数情况下，弹性薄壁构件的变形为几何非线性以及流体方程的非线性将导致流体和弹性体相互作用界面上的强非线性，可以在界面上结合拉格朗日法和欧拉法建立方程及接触条件。其方法可分为单一拉格朗日法、单一欧拉法、相容拉格朗日 – 欧拉法和任意拉格朗日 – 欧拉法四种方法。在流体弹性力学理论的基础上，介绍了流固耦合问题界面相互作用的描述方法，并根据诺沃日洛夫 BB 在非线性弹性力学中，从几何非线性方面对弹性力学问题的分类方法出发，对流体弹性力学中的流固耦合问题进行分类，由此，可按类分别对运动学条件、动力学条件及界面上的接触条件进行相应的简化提供可靠的依据。陶海亮等人^[71]主要研究了气热固耦合场对涡轮模态参数的影响，文章采用了基于 $k-\varepsilon$ 湍流模型理论建立了涡轮的流场模型，并进行网格划分和边界条件的加载，通过气热固耦合分析计算获得了流场内部温度

和压力的分布，把气动力及温度载荷映射到涡轮结构上，并在此基础上进行了涡轮模态的计算。计算结果表明，气热固耦合场主要影响涡轮结构的模态固有频率，对模态振型的影响较小。

陈建栋等人^[72]为了解决微型空压机缸过热的问题，在基于流固耦合传热理论的基础上，建立了一种微型空压机内部流场和温度场以及轴流式风扇的数值计算模型。文章实验结果表明通过改变轴流式风扇的吹风方向解决了冷却空气在机架处回流的问题，通过优化出风口尺寸以及风扇参数，提高了冷却空气的流量，解决了缸盖处过热的问题。岳迎春等人^[73]分析了在油田老井挖潜、稳产增产方面，重复压裂已成为重要的技术手段，而地应力转向机制与压裂时机选择一直是制约该技术增产效果的关键问题。文章还对ABAQUS有限元软件平台进行了二次开发，实现了流体压力变化与岩石物性参数的全面耦合，并分析了重复压裂转向机制。结果表明，地应力转向现象普遍存在，人工裂缝对地应力的影响范围有限，而孔隙压力变化是造成地应力转向的主要因素；随着生产的持续进行，地层压力下降变缓，岩石体积应变变化趋缓，导致渗透率下降趋于平稳，应力转向距离逐渐增大并最终趋稳；应力差越大、应力转向距离越小，越难形成重复压裂转向裂缝。其研究结果实现了流-固耦合作用下地应力转向的可视化描述，直观地模拟结果有利于指导重复压裂的应用实施。

王月红等人^[74]提出了针对由于采空区内非均匀孔隙介质的存在，造成采空区风流动态流动的复杂性，在渗流力学与气-固理论相结合的基础上，采用理论分析、软件模拟以及实例验证等方法，分析了采空区内孔隙度变化规律以及多参数渗流连续数学模型；基于高效的有限体积算法进行了数值解算，并编制了模拟软件，研究结果表明采空区内渗透率整体呈现“U形”分布，在进风口和回风口处最大，随着采空区深度的增加渗透率逐渐变小；压力呈“扇形”分布，在进风口处最大，随着向采空区延伸，压力变化逐渐趋近于零；采空区内渗流速度呈“弧形”分布，靠近工作面的区域漏风较强，随着向采空区延伸，漏风速度逐渐趋近于零；同时对模拟结果进行了二维和三维演示，使问题更清晰、更直观，为辅助分析采空区瓦斯排放、自然发火、注氮灭火、大面积均压调压等诸多安全问题提供了保障。吴朋朋等人^[75]在针对固体火箭发动机大膨胀比喷管出现的分离流动的问题上，采用数值仿真方法进行分析，并与试验进行对比。通过集成软件平台MpCCI，连接计算流体动力学软件FLUENT和有限元软件ABAQUS，对燃气流动与喷管结构运动变形进行了耦合计算。耦合计算结果发现，此大膨胀比喷管发生气流分离，且分离处斜激波后的气流温度与压力变化较大，采用流固耦合数值方法能体现喷管的结构变形，从而更准确地反映喷管与燃气流相互影响的真实环境。通过耦合计算结果与试验对比，得出了耦合计算获取的分离位置能很好地拟合实验测得的气流分离位置，说明了流固耦合数值方法的有效性，为更深入研究大膨胀比喷管分离流动现象提供了支撑。

廖家麒等人^[76]阐述了反应堆压力容器（RPV）支座在核电厂设计中起到支承 RPV 的作用，属 ASME 核安全 1 级支承，是关系到核电厂安全的重要设备。文章采用响应面法实验设计，结合建立的 RPV 支座“流-固”耦合模型，系统分析支座中竖板开孔位置和入口空气流速对支座温度场和空气流场的影响规律；进一步面向支座底面最高温度限

制和减小中竖板开孔应力集中的双目标要求，获得了优化的中竖板开孔位置和入口空气流速，为 RPV 支座结构设计和入口空气流速选择提供了指导。唐铭等人^[77]分析了作用在索膜结构上的风荷载，除与气流本身的特性有关外，还与结构在风荷载作用下的位移、速度、加速度等有关。索膜结构与风环境之间存在着较强的流固耦合作用，并且这种耦合作用往往对索膜结构的振动起控制作用。文章将两者的耦合作用分为静力耦合作用和动力耦合作用，并主要针对静力耦合作用所涉及的几个因素运用解析公式的说明方法进行了初步分析总结，为进一步的风致动力响应研究和数值风洞试验模拟研究奠定了理论基础。

郭衍茹等人^[78]针对气井生产过程中存在的流固耦合作用，在充分考虑真实天然气的偏差因子、密度以及黏度随压力变化的基础上建立数学模型，利用 COMSOL Multiphysics 软件对应力场、渗流场进行耦合计算，得出了流固耦合作用下的渗流场分布。从中提取井底上方 20m 处切线上的地层压力、孔隙度、渗透率、渗流速度、气体密度、气体黏度等数据，进一步分析耦合渗流场的分布特点。郭君等人^[79]采用谱单元方法建立水下爆炸瞬态流固耦合的三维数值模型，探讨了水下爆炸瞬态流固耦合作用的机理，用经典的平板冲击问题对数值模型进行验证，数值结果与解析解吻合良好，并根据数值结果绘制了流体中的气穴区域，对气穴效应进行分析，分析表明，气穴效应能对结构响应产生很大影响，在计算中应予以考虑。基于所建立的数值模型，在不同网格细化的条件下，分别采用谱单元方法和有限元方法对弹簧—平板模型进行水下爆炸瞬态流固耦合问题的求解，并在此基础上对谱单元方法和有限元方法进行了对比研究。

蔡定国等人^[80]为了模拟对流换热过程，摆脱气道参数对变压器温升计算结果的影响，从空气质量、动量、能量守恒方程出发，建立了干式变压器流固耦合场数值模型，采用有限差分法对温度场进行了计算，并与实验结果进行了比较，模拟结果与实验吻合较好。考虑空气流场的流固耦合模型能准确分析干式变压器工作中的温升问题，可分析变压器内部温度分布规律及影响温度分布的因素，指导变压器优化设计。陈华磊等人^[81]在考虑高温气体对燃气轮机旋转对称支承机座的作用的前提下，利用计算流体动力学软件 CFX 对支承板机座进行热流固耦合分析，并运用 Workbench 进行热固耦合计算。通过变换热气通道进口温度的计算方法，得到变工况条件下旋转对称支承板机座表面温度场分布，进而得到变工况下机座的热变形、热应力的变化关系。研究结论具有一定的理论意义和实用价值。

吴作伟等人^[82]以高速动车组车下水箱为研究对象，采用双向流固耦合方法，对车载水箱箱体内液体冲击晃动问题进行了建模与数值分析。通过研究列车变速过程中水箱内气液两相分布状态，获得了采用单向耦合方法不能获得的水箱结构应力随时间的变化历程，实现了对水箱结构强度更为准确的分析。文章为高速动车组水箱结构的设计提供了实用可靠的分析方法。尚翠霞等人^[83]研究钢厂废煤气流场作用下大型压力控制阀的稳压问题，分析了影响压力稳定的结构因素，利用 Pro/E 建立了控制阀的几何模型，采用 CFX 进行流场计算，ANSYS 进行结构计算，并利用两者间的数据交互平台传递流场压力载荷和结构位移数据，实现了流固耦合数值仿真计算。研究结果为控制阀结构的优化设

计和测试分析提供了重要依据。吴正人等人^[84]采用数值计算软件 CFX 和 Ansys 对离心风机叶轮进行了流固耦合模拟研究，对叶轮的强度、模态和振动特性进行了计算和分析。

4 国内外研究进展比较

多模态物理过程仿真涉及多个交叉学科，存在于各个不同的领域，与单个目标分析相比，多模态物理过程接近于真实物理世界。尽管近年来国内外研究人员在气体、液体和固体的多模态物理过程仿真领域取得了一定的成绩，但是要实现精确、高效率的多模态物理过程仿真，还需要解决现有方法所面临的如下主要问题：

1) 随着多模态物理过程仿真场景复杂度要求的不断提高，现有的处理方法难以适应复杂场景下的多模态物理过程仿真。以虚拟手术中的多模态物理过程仿真为例，人体器官组织形状大小复杂多变，现有方法无法对任意复杂的器官组织模型精确地生成内部体数据，也没有一个真实性高、适用性强的统一多模态物理过程仿真方法来方便地实现复杂场景的多模态物理过程仿真。因此，需要研究有较强真实性、适用于任意复杂模型以及多种复杂场景混合仿真的通用多模态物理过程仿真理论与方法。

2) 现有大部分基于物理仿真方法通常采用类似于 SPH 方法的无网格法用来对气体和液体建模，采用弹性模型对固体建模，这类方法往往缺乏对气体、液体和固体的材料性质的考虑。例如，在真实世界中，几乎所有的生物组织都是由固态和液态两种成分共同构成，生物组织是最典型的流变体，体液具有松弛行为而组织具有蠕变行为，因此从真实性的角度考量，在建立流体和固体模型的同时需要结合流变力学来进行多模态物理过程的分析。

3) 现有方法在处理多物体、多模态物理过程（包括碰撞检测和响应）较为复杂，在碰撞检测和响应过程中耗费较多资源，且通常选取与运动粒子最近的粒子作为碰撞点，缺乏精确性，容易产生穿透现象，因此需要研究精确高效的碰撞检测和响应算法来实现多物体、多相的耦合过程。

针对这其中的具体问题，国内外学者都做了一定的研究，并取得了一些研究成果。Solenthaler 等人^[19]将固体、流体粒子模型的接触面上的粒子视为统一整体，依然采用 SPH 方法的邻域搜索计算接触面上粒子受力，此方法在流固耦合时一旦相互速度差过大，将导致流体与固体粒子相互进入的非真实现象发生。Schechter 等人^[20]在固体和流体粒子模型表面粒子支持域内均设置一层特质虚粒子，保持固体粒子模型中粒子的稳定性，防止流体和固体粒子相互进入，但虚粒子的设置与计算过程却相当复杂。Afonso Paiva 等人^[21]在碰撞处理时将粒子的速度分为切向速度与法向速度分别进行修饰处理。Akinci 等人^[85]提出一种新颖的、通用的能同时用于单向双向耦合的流固耦合的边界处理方法，该模型将固体建模成刚体，只需要与固体最外层粒子进行计算，且不需要粒子法向量，同时计算的边界粒子不需要均匀采样，能稀疏也能浓密，流固边界粒子的受力是对称的。但是，该模型只适用于刚体，不能用于流体与实体可变形的固体间的物理过程仿真计算，

且这种方法严重的依赖预采样对象，并且很难被推广到处理大变形的固体。Chentanez N. 等人^[86]对于固体变形中欧拉法求解的流体和网格法的固体交互现象提出了一种解决方案，但是不能很好地处理网格法和无网格法间的整体差异。Losasso F. 等人^[87]提出一种能处理燃烧和融化现象中固体与液体、气体等流体的耦合问题解决方法，但是对于处理更复杂的固体的凝固现象缺少有效的方法。

为了实现高速列车多模态物理过程的动力学行为的联合仿真，西南交通大学牵引动力国家重点实验室的李田等人^[88,89]在高速列车流固耦合的平衡状态方法中为实现横风下车辆-轨道耦合动力学和列车空气动力学之间的联合仿真，提出一种快速高效计算的平衡状态方法，在忽略轨道不平顺情况下，轮流交替求解稳定气动力作用下的车辆-轨道耦合动力学响应以及稳定姿态下的列车流场和气动力；通过收敛准则判定横风下高速列车的平衡状态；将平衡状态下的稳定气动力加载车辆-轨道耦合动力学并计算高速列车动力学响应。通过将车辆-轨道耦合动力学计算程序嵌入流体力学程序中，避免了流体与固体两个求解器之间的数据交换通信，并且采用耦合迭代步内流固边界条件气动力线性插值的方法实现了流固两个求解器时间迭代步长的独立选取，提高了计算速度。

浙江大学岩土工程研究所的陈云敏等人^[90]建立了一维固、液、气三相耦合的大孔隙垃圾固结数学模型，把垃圾视为大孔隙介质，假定基质吸力等于0，即孔隙气压等于孔隙液压，把填埋场简化为非稳定单向渗流场，模型中结合了达西定律、气体状态方程、有效应力原理和多孔介质流体动力学理论。填埋场的沉降按封场前和封场后两个阶段分别计算，封场前填埋场接受垃圾，导致堆载压力增加，而封场后的填埋场堆载压力保持不变。

北京航空航天大学虚拟现实国家重点实验室 Lipeng Yan 等人^[91]提出一种新颖的方法用于模拟液体和可变形固体间的双向耦合作用，其中液体采用 SPH 方法建模，可变形固体采用非线性有限元建模。为了实现精确的双向耦合，该文在可变形固体的边界处生成代理粒子，用于处理可变形固体和液体之间的交互，提出一种有效的方法将交互过程中作用力施加到可变形固体模型的非线性有限元结点上，并采用 TLED (Total Lagrangian Explicit Dynamics) 有限元算法求解可变形固体非线性有限元模型，使用 GPU 并行求解整个流体和固体的双向耦合仿真过程。该文所提出的方法主要缺陷是有时大量液体粒子会在固体表面堆积，从而影响数值计算的稳定性。和 Müller^[92]提出的边界粒子方法比较，该文章代理粒子是动态形成的且时间和内存消耗小，其速度可以按照网络拓扑关系上的周围顶点的插值来获取，因此当固体网格元素是相对稀疏的流体粒子时，它避免了动态细分该多边形。

针对国内外研究中所面临的主要问题，武汉大学计算机学院也做了一些相关的工作。以虚拟手术仿真中的多模态物理过程仿真为例，结合 SPH 方法和流变力学研究了一种新颖的虚拟手术中多模态物理过程仿真通用统一粒子框架，提出了一种对任意器官模型生成粒子化体数据的方法，可适用于任意复杂模型、多种复杂场景下进行虚拟手术中多模态物理过程仿真。为了仿真人体组织及体液的物理性质，结合流变力学对体液和组织的物理行为进行建模，在处理多模态物理过程中多个模态物质的交互作用时，提出

了基于边界粒子的虚拟球碰撞检测算法以及基于多模态耦合核函数的碰撞响应方法，为多模态粒子之间的碰撞响应赋予流变体性质，以便能精确、高效地对虚拟手术中的多模态物理过程进行仿真。

这些方法与国内外同类工作相比较，具有一定的特色。与 Schechter 等人^[20]的方法相比，提出的碰撞检测方法只需对可变形固体粒子模型和刚体粒子模型的一层边界粒子进行处理，并不需要对所有的粒子进行碰撞检测，提高了碰撞检测的效率。Schechter^[20]需要通过建立虚粒子并对其进行复杂的设置来实现流体固体耦合现象，我们的方法通过虚拟球碰撞检测算法只需要对小范围的边界粒子进行碰撞检测，这样就可以实现高效的碰撞检测同时能有效地防止穿透现象的发生。Müller、Bender、Macklin 等人^[16,23,24]提出了基于位置的方法实现了流体及固体耦合的仿真，然而其物理真实性有待于进一步提高，我们的方法虽然计算过程相较于基于位置的方法更为复杂、计算时间也更长，但是通过高性能计算能弥补这种不足且符合物理真实性，适用于实际应用中多模态物理过程仿真对仿真精度的要求。Afonso 等人^[21]在处理流体和固体交互过程的碰撞响应时，对碰撞后的流体粒子速度乘以一个系数以模拟流体和固体粒子之间的耦合作用力，而我们引入了多模态粒子之间的摩擦力，并通过流变粘滞系数的使用为多模态粒子间的作用赋予流变体特性，在求解完每种模态粒子的摩擦力之后采用动量守恒定律对多相粒子进行物理量的求解，这样更具有物理真实性。

虽然国内外在多模态物理过程仿真应用的许多领域进行了大量深入的研究，然而国内外学者的研究却各有着不同的特点。我们发现，从多模态物理过程仿真的研究成果上来看，国外学者原创性的理论方法相对较多，如基于网格法和无网格法的新方法层出不穷，新方向和研究内容不断涌现；而国内学者所关注的研究内容大多集中于热点问题上，在方法改进及具体应用方面的成果较多，即通过在现有的多模态物理过程仿真方法上针对其缺陷进行相应的改进后将其应用于某个具体领域中，而在新方法、关键技术及前沿领域的研究学者相对较少。从近几年发表的高水平论文来看，国外学者在一流期刊上发表的研究成果相对较多，而国内学者在同类研究成果方面仍处于弱势，相比国外学者还有一定的差距。但是，我们也应注意到，随着国家对科研及教育投入的不断增加和优秀人才激励机制建设的不断完善，国内与国外在相关研究领域的差距会逐渐缩小。

5 国内外研究发展趋势展望

随着科学与工程应用技术的发展，多模态物理过程的研究一直得到了国际学术界的广泛关注。多模态物理过程的研究对象是固体在流体（气体、液体）作用下的各种行为以及固体变形或运动对流体影响。多模态物理过程的重要特征是不同介质之间的交互作用，固体在流体动载荷作用下会产生变形或运动，而固体的变形或运动又反过来影响流体，从而改变流体载荷的分布和大小，正是这种相互作用将在不同条件下产生许多多模态物理过程^[50]。

目前，多模态物理过程仿真取得了较大的进展，主要体现在如下几个方面：1) 线化理论日益完善、程序化。通过不断研究，提出了基于考虑自由面线性波的两个变分公式，提出了混合型多模态物理过程分析有限元模型的三种对称化方法（位移型，平衡型，混合型），完成了可求解大量工程耦合问题的线性多模态物理过程的动力分析；2) 非线性耦合问题研究取得显著成果。非线性耦合问题同其他耦合问题一样，由于非线性，导致叠加原理失效。在固体力学非线性有限元方法和流体力学非线性 N-S 方程求解技术发展的基础上，多模态物理过程非线性求解取得了较大的进步^[51]。

随着科学技术的快速发展，多模态物理过程仿真通过结合最新科学技术、网络技术及通信技术的发展，可能呈现出如下发展趋势：

(1) 基于高性能计算的多模态物理过程仿真

随着多模态物理过程仿真的数据规模和精度的提高，现有的在单台计算机上实现的串行编程方法无法快速、高效地对多模态物理过程进行仿真，多模态物理过程仿真将向高性能计算方向发展，通过求解效率及求解精度的提高，实现高效逼真的多模态物理过程仿真。其中多模态物理过程仿真的高性能计算可进一步分为如下三个层次。

第一个层次是在单台计算机上利用基于 GPU 的并行计算技术实现多模态物理过程仿真。其中 CUDA (Compute Unified Device Architecture) 和 OpenCL (Open Computing Language) 是近几年提出的大规模并行计算技术解决方案，在科学和工程领域具有广阔的应用前景。CUDA 是由 NVIDIA 公司提出的一种新的并行编程模型和指令集架构的通用计算架构^[93]，它能使得编程人员很方便地在 NVIDIA 显卡芯片上开发出大规模并行计算程序。OpenCL 是一种开放的针对异构系统进行并行编程标准，使用它可以在 CPU、GPU 和其他处理器上进行通用目的的并行编程。OpenCL 具备良好的跨平台特性，可以在多种硬件平台方便地开发出大规模并行计算程序而无需知道特定的芯片指令是什么，得以最大性能地发挥异构系统的计算性能^[94~97]。

第二个层次是分布式并行计算环境下的多模态物理过程仿真，它采用一组由多种不同结构功能组成的、网络互连的计算机来实现多模态物理过程仿真。通过将多模态物理过程仿真分解成许多小的计算部分，然后将它们分配给多台计算机处理，再将得到的结果综合起来获得最终的多模态物理过程仿真结果。

第三个层次是互联网云计算环境下的多模态物理过程仿真。云计算由一系列可以动态升级和被虚拟化的资源组成，这些资源被所有云计算的用户共享并且可以方便地通过网络访问。云计算是分布式计算、并行计算和网格计算的发展，对大数据量的操作性能非常好，不但包括分布式计算还包括分布式存储和分布式缓存。云计算通过将多模态物理过程仿真分割成多个进程，并对云内的基础设备进行动态按需分配与管理，以便在多台服务器上并行执行得到最终的多模态物理过程仿真结果。

(2) 基于大数据分析的多模态物理过程仿真

人们对多模态物理过程仿真的精确性要求越来越高、场景复杂度及场景规模不断增大，常规的多模态物理过程仿真方法将不能满足真实性要求，需要结合真实实验数据来修正多模态物理过程仿真结果，以达到最接近真实的仿真结果。研究者需要使用大量的

实验数据，这些实验结果数据量巨大，可能是不同研究者多年所积累的实验数据，研究者难以在合理时间内达到撷取、管理、处理，并整理成为有用的信息，因此研究者必须结合大数据分析来获取最有价值的实验数据，将其与仿真结果进行对比，对仿真模型进行有效的修正，以实现接近真实实验结果的多模态物理过程仿真。以虚拟手术中多模态物理过程仿真为例，为了达到虚拟手术训练的最佳效果，需要用最佳的手术方案来训练初学医生。为了找到最佳的手术方案对医生进行培训，研究人员需要通过记录大量真实手术过程中，包括手术路径、体液、组织和腔内环境变化的数据，通过大数据分析寻求一种最佳的手术方案，并将此过程进行多模态物理过程仿真，同时修正仿真过程中出现的误差，以达到医生培训的最佳效果。

(3) 物联网环境下的交互式多模态物理过程仿真

在实际应用中，多模态物理过程仿真通常需要引入用户的交互作用，以对不同情形下的多模态物理过程进行交互式仿真。以虚拟手术中的多模态物理过程仿真为例，虚拟手术是利用各种医学影像数据，利用虚拟现实技术在计算机中建立一个模拟环境，医生借助虚拟环境中的信息进行手术计划、训练，以及实际手术过程中引导手术。然而，虚拟手术系统价格昂贵，在培训医生时，现有的系统只能通过单台计算机设备进行虚拟手术训练，这样效率太低，无法用于批量的外科医生培训。因此，我们需要寻找到一种更高效率的外科医生培训方案，如借助 WebGL 和物联网技术建立一种信息物理系统环境下的 Web 虚拟手术系统，虚拟手术交互设备作为传感器接入到物联网中，医生借助物联网实现远程虚拟手术操作训练，这样便能大大地提高医生培训的效率。

6 结束语

多模态物理过程仿真在计算机仿真、虚拟现实、工业制造、医疗等领域具有重要研究意义和应用价值。现阶段国内外学者在多模态物理过程仿真中取得了丰硕的研究成果，并将多模态物理过程仿真应用于许多领域，解决了一系列实际工程与应用问题。多模态物理过程仿真属于多学科交叉研究，当前多模态物理过程仿真还有许多问题没有很好地得到解决，需要国内外相关研究人员不断努力和协作，才能更高效地解决相关工程应用中所面临的新问题，从而将多模态物理过程仿真研究和应用推向新的高度，取得更多、更好的研究成果。

参考文献

- [1] Brian Wheeler, Philip C Doyle, Shamir Chandarana. Interactive Computer-Based Simulator for Training in Blade Navigation and Targeting in Myringotomy. *Computer Methods and Programs in Biomedicine*, 2010, Vol. 98, No. 2, Pages: 130-139.

- [2] Yi-Je Lim, Suhranu De. Real Time Simulation of Nonlinear Tissue Response in Virtual Surgery Using the Point Collocation-Based Method of Finite Spheres. Computer methods in Applied Mechanics and Engineering, 2007, Vol. 196, No. 31, Pages: 3011-3024.
- [3] Francios Faure, Benjamin Gilles, Guillaume Bousquet, Dinesh K. Pai. Sparse Meshless Models of Complex Deformable Solids. ACM Transactions on Graphics, 2011, Vol. 30, No. 4, Article 73: 1-9.
- [4] Müller M, Keiser R, Nealen A, Pauly M, Gross M, Alexa M. Point Based Animation of Elastic, Plastic and Melting Objects. Proceedings of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2004, Pages: 141-151.
- [5] Pauly, M, Keiser, R, Adams, B, Dutre, P, Gross, M Guibas, L J. Meshless Animation of Fracturing Solids. ACM Transactions on Graphics, 2005, Vol. 24, No. 3, Pages: 957-964.
- [6] Gerszewski, D, Bhattacharya, H, Bargteil, A W. A Point-based Method for Animating Elastoplastic Solids. Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2009, Pages: 133-138.
- [7] Yu J, Turk G. Reconstructing Surfaces of Particle- based Fluids Using Anisotropic Kernels. ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2010, Pages: 217-225.
- [8] Knight R Pyrak- Nolte L J, Slater L. Geophysics at the Interface: Response of Geophysical Properties to Solid-Fluid, Fluid- Fluid, and Solid- Solid Interfaces. Reviews of Geophysics, 2010, Vol. 48, No. 4, RG4002: 1-30.
- [9] Alastair Nottingham, Barry Irwin. GPU Packet Classification Using OpenCL: a Consideration of Viable Classification Methods. Proceedings of the 2009 Annual Research Conference of the South African Institute of Computer Scientists and Information Technologists, 2009, Pages: 160-169.
- [10] Fedkiw R, Stam J, Jensen H W. Visual Simulation of Smoke. Proceedings of the 28th Annual Conference on Computer Graphics and Interactive Techniques, 2001, Pages: 15-22.
- [11] Foster N, Metaxas D. Modeling the Motion of a Hot, Turbulent Gas. SIGGRAPH 97 Conference Proceedings, Annual Conference Series, August 1997, Pages: 181-188.
- [12] Stam J. Stable Fluids. SIGGRAPH 99 Conference Proceedings, Annual Conference Series, August 1999, Pages: 121-128.
- [13] Schechter H, Bridson R. Evolving Sub-grid Turbulence for Smoke Animation. Proceedings of the 2008 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2008, Pages: 1-8.
- [14] Song O Y, Shin H, Ko H S. Stable but Nondissipative Water. ACM Transactions on Graphics, 2005, Vol. 24, No. 1, Pages: 81-97.
- [15] Kim Y, Thurey N, James D, et al. Wavelet Turbulence for Fluid Simulation. ACM Transactions on Graphics, 2008, Vol. 27, No. 3, Article 50: 1-6.
- [16] Müller M, Heidelberger B, Hennix M, Ratcliff J. Position Based Dynamics. Journal of Visual Communication and Image Representation, 2007, Vol. 18, No. 2, Pages: 109-118.
- [17] Müller M, Chentanez N. Solid Simulation with Oriented Particles. ACM Transactions on Graphics, 2011, Vol. 30, No. 4, Article 92: 1-9.
- [18] Gerszewski D, Bhattacharya H, Bargteil A W. A Point- based Method for Animating Elastoplastic Solids. Proceedings of the 2009 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, August 2009, Pages: 133-138.
- [19] Solenthaler B. , Schläfli J, Pajarola, R. A Unified Particle Model for Fluid- solid Interactions. Computer

- Animation and Virtual Worlds, 2007, Vol. 18, No. 1, Pages: 69-82.
- [20] Schechter H, Bridson R (2012). Ghost SPH for Animating Water. ACM Transactions on Graphics, 2012, Vol. 31, No. 4, Article 61: 1-8.
- [21] Afonso Paiva, Fabiano Petronetto, Thomas Lewiner, Geovan Tavares. Particle-based Viscoplastic Fluid/solid Simulation. Computer-Aided Design, 2009, Vol. 41, No. 4, Pages: 306-314.
- [22] Becker M, Ihmsen M, Teschner M. Corotated SPH for Deformable Solids. Proceedings of the Fifth Eurographics Conference on Natural Phenomena, 2009, Pages: 27-34.
- [23] Bender J, Müller M, Otaduy M A, Teschner M. Position-based Methods for the Simulation of Solid Objects in Computer Graphics. Eurographics 2013-State of the Art Reports, 2012, Pages: 1-22.
- [24] Macklin M, Müller M. Position Based Fluids. To Appear in ACM Transactions on Graphics, 32 (4), http://www.matthiasmueller.info/publications/pbf_sig_preprint.pdf.
- [25] Monaghan J. J. Smoothed Particle Hydrodynamics. Reports on Progress in Physics, 2005, Vol. 68, No. 8, Pages: 1703-1759.
- [26] Lucas Braune, Thomas Lewiner. An Initiation to SPH, Department of Mathematics, PUC-Rio-Rio de Janeiro, Brazil, 2009.
- [27] Monaghan, J J. Simulating Free Surface Flows with SPH. Journal of Computational Physics, 1994, Vol. 110, No. 2, Pages: 399-406.
- [28] Müller M, Charypar D, Gross M. Particle-based Fluid Simulation for Interactive Applications. Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, July 2003, Pages: 154-159.
- [29] Solenthaler B, Pajarola R. Predictive- corrective Incompressible SPH. ACM Transactions on Graphics, 2009, Vol. 28, No. 3, Article 40: 1-6.
- [30] Adams B, Pauly M, Keiser R, Guibas L J. Adaptively Sampled Particle Fluids. In ACM Transactions on Graphics, 2007, Vol. 26, No. 3, Article 48: 1-7.
- [31] Solenthaler, B, Gross, M. Two- scale Particle Simulation. ACM Transactions on Graphics, 2011, Vol. 30, No. 4, Article 81: 1-8.
- [32] Ihmsen M, Akinci N, Becker M, Teschner M. A Parallel SPH Implementation on Multi- core CPUs. Computer Graphics Forum, 2011, Vol. 30, No. 1, Pages: 99-112.
- [33] Richard Keiser, Bart Adams, Dominique Gasser, Paolo Bazzi, Philip Dutré et al. A Unified Lagrangian Approach to Solid- fluid Animation. Eurographics/IEEE VGTC Symposium Proceedings, 2005, Pages: 125-148.
- [34] D Nordsletten, M McCormick, P J Kilner, P Hunter, D Kay, N P Smith. Fluid-solid Coupling for the Investigation of Diastolic and Systolic Human Left Ventricular Function. Numerical Methods in Biomedical Engineering, 2011, Vol. 27, No. 7, Pages: 1017-1039.
- [35] Sebastian Kritian, Uwe Janoske, Herbert Oertel, Thomas Bohlke. Partitioned Fluid- solid Coupling for Cardiovascular Blood flow. Annals of Biomedical Engineering, 2010, Vol. 38, No. 4, Pages: 1426-1441.
- [36] Rahul Narain, Abhinav Golas, Ming C. Lin. Free- flowing Granular Materials with Two- way Solid Coupling. ACM TransActions on Graphics, 2010, Vol. 29, No. 6, Article 173: 1-10.
- [37] Michelin, Sébastien, Llewellyn Smith, Stefan G. Falling Cards and Flapping Flags: Understanding Fluid-solid Interactions Using an Unsteady Point Vortex Model. Theoretical and Computational Fluid Dynamics, 2010, Vol. 24, No. 1-4, Pages: 195-200.

- [38] D. Gregory Tipton, Mark A Christon, Marc S Ingber. Coupled Fluid-solid Interaction Under Shock Wave Loading. International Journal for Numerical Methods in Fluids, 2011, Vol. 67, No. 7, Pages: 848-884.
- [39] Yen Ting Ng, Chohong Min, Frederic Gibou. An Efficient Fluid-solid Coupling Algorithm for Single-phase Flows. Journal of Computational Physics, 2009, Vol. 228, No. 23, Pages: 8807-8829.
- [40] Christopher Batty, Florence Bertails, Robert Bridson. A Fast Variational Framework for Accurate Solid-fluid Coupling. ACM Transactions on Graphics, 2007, Vol. 26, No. 3, Article 100: 1-7.
- [41] Taha Sochi. Slip at Fluid-solid Interface. Polymer Reviews, 2011, Vol. 51, No. 4, Pages: 309-340.
- [42] F llinca, J-F Hetu. Numerical Simulation of Fluid-solid Interaction Using an Immersed Boundary Finite Element Method. Computers & Fluids, 2012, Vol. 59, Pages: 31-43.
- [43] S Tenneti, B Sun, R Garg, S Subramaniam. Role of Fluid Heating in Dense Gas-solid Flow as Revealed by Particle-resolved Direct Numerical Simulation. Heat and Mass Transfer, 2013, Vol. 58, No. 1, Pages: 471-479.
- [44] Frank Lomine, Luc Scholtes, Luc Sibille, Philippe Poullain. Modeling of Fluid-solid Interaction in Granular Media With Coupled Lattice Boltzmann/discrete Element Methods: Application to Piping Erosion. Numerical and Analytical Methods in Geomechanics, 2013, Vol. 37, No. 6, Pages: 577-596.
- [45] 王丽, 陈嘉陵. 泥石流的数值模拟和试验模拟方法[J]. 水土保持科技情报, 2004, (1): 7-10.
- [46] 陈日东, 刘兴年, 曹叔尤, 郭志学. 泥石流与主河汇流堆积的数值模拟[J]. 中国科学: 科学技术, 2011, 41(10): 1305-1314.
- [47] 杨志亮. 泥石流现象的真实感建模与绘制[D]. 浙江: 浙江大学, 2008.
- [48] 石莹, 林显臣. 基于 Fluent 的气固喷射器三维数值模拟研究[J]. 选煤技术, 2011, 4(6): 19-22.
- [49] 王勇, 王阳阳. 冲击载荷下液压缸内压的流固耦合仿真与振动分析[J]. 液压与气动, 2011, 10 (9): 25-27.
- [50] 孙剑韬, 姜琦, 王军红, 徐伟, 王昊. 基于引信 MEMS 射流发电机的流固耦合分析[J]. 弹箭与制导学报, 2012, 32(4): 84-88.
- [51] 落葵寿, 等. 整流罩地面分离试验流固耦合分析与数值模拟[[M]]. 哈尔滨工业大学, 2008.
- [52] 夏生林, 赵利霞, 李伟, 唐克兵. 基于 CFD 非线性气动力的静气弹计算[D]. 第十二届全国空气弹性学术交流会会议论文集, 2011, 2148: 251-256.
- [53] 孙栋, 李福生, 徐新喜, 赵秀国, 谭树林. 流固耦合作用下人体上呼吸道内气流运动特性数值仿真研究[J]. 中国生物医学工程学报, 2012, 31(1): 89-95.
- [54] 胡海峰, 鲍福廷, 王艺杰, 蔡强, 刘旸. 喷管分离流流动 - 热 - 结构顺序耦合数值模拟及试验研究[J]. 宇航学报, 2011, 32(7): 1534-1541.
- [55] 姜海晨, 张娜, 于振环. 减振器叠加节流阀片非线性特性仿真分析[J]. 计算机仿真, 2012, 29 (5): 338-342.
- [56] 郭军刚, 王春侠, 胡丽国. 基于流 - 固耦合方法的涡轮叶片热应力仿真[J]. 强度与环境, 2012, 39(1): 22-28.
- [57] 崔涛, 张卫华. 基于姿态变化的列车侧风安全性研究[J]. 铁道学报, 2010, 32(5): 25-29.
- [58] 刘辉, 陈福忠, 项昌乐. 油 - 固 - 气耦合作用下车辆变速箱辐射噪声估测[J]. 农业工程学报, 2011, 27(11): 56-60.
- [59] 李小川, 胡亚非, 蒋炎炎, 熊建军. 基于脉动压力时均值的除尘器气液耦合研究[J]. 煤炭学报, 2011, 36(8): 1412-1416.
- [60] 张广明, 刘合, 张劲, 彪仿俊, 吴恒安, 等. 油井水力压裂流 - 固耦合非线性有限元数值模拟

- [J]. 石油学报, 2009, 30(1): 113-116.
- [61] 刘小民, 王星, 许运宾. 运动罐体内液体晃动的双向流固耦合数值分析[J]. 西安交通大学学报, 2012, 46(5): 120-126.
- [62] 高兴龙, 高庆玉, 张青斌, 唐乾刚. 基于 ALE 方法的开缝降落伞充气过程研究[J]. 航天返回与遥感, 2013, 34(1): 13-19.
- [63] 彭正标, 袁竹林. 基于蒙特卡罗法的脱硫塔内气固流动数值模拟[J]. 中国机电工程学报, 2008, 28(14): 6-14.
- [64] 杨玉新, 胡春波, 孙得川, 蔡体敏, 陈灏, 等. 基于流 - 固耦合的混合火箭发动机固体燃料表面退移速率计算[J]. 固体火箭技术, 2007, 30(3): 214-218.
- [65] 孙可明, 潘一山, 梁冰. 流固耦合作用下深部煤层气井群开采数值模拟[J]. 岩石力学与工程学报, 2007, 26(5): 994-1001.
- [66] 程载斌, 刘玉标, 刘兆, 申仲翰. 导弹水下潜射过程的流体 - 固体耦合仿真[J]. 兵工学报, 2008, 29(2): 178-183.
- [67] 张俊鸽, 张伟民, 郝晓伟, 马烨. 动态淬火过程的流 - 固耦合数值模拟[J]. 材料热处理学报, 2008, 29(3): 176-180.
- [68] 陈明, 焦光伟, 雍岐卫, 冯剑. 管道气液固三相流的耦合瞬变模型研究[J]. 后勤工程学院学报, 2011, 27(3): 30-34.
- [69] 曹树刚, 鲜学福. 煤岩固 - 气耦合的流变力学分析[J]. 中国矿业大学学报, 2001, 30(4): 362-365.
- [70] 朱洪来, 白象忠. 流固耦合问题的描述方法及分类简化准则[J]. 工程力学, 2007, 24(10): 92-99.
- [71] 陶海亮, 郭宝亭, 谭春青. 基于气热固耦合的涡轮模态分析[J]. 振动、测试与诊断, 2012, 32(6): 941-944.
- [72] 陈建栋, 孙蓓蓓. 基于流固热耦合仿真的微型空压机风冷系统[J]. 东南大学学报(自然科学版), 2013, 43(1): 65-70.
- [73] 岳迎春, 郭建春. 重复压裂转向机制流 - 固耦合分析[J]. 岩土力学, 2012, 33(10): 3189-3193.
- [74] 王月红, 温佳丽, 秦跃平, 张九零. 采空区多参数气 - 固耦合渗流模拟[J]. 辽宁工程技术大学学报(自然科学版), 2012, 31(5): 760-764.
- [75] 吴朋朋, 杨月诚, 高双武, 张昆鹏. 固体火箭发动机喷管分离流动流固耦合数值仿真[J]. 固体火箭技术, 2012, 35(3): 344-347.
- [76] 廖家麒, 刘冬安. 反应堆压力容器支座“流 - 固”耦合分析及优化设计[J]. 制造业信息化, 2013, 5: 87-89.
- [77] 唐铭. 张拉索膜结构与风场的流固耦合作用分析[J]. 城市设计理论研究, 2013, 14: 1-5.
- [78] 郭衍茹, 练章华, 魏臣兴, 蔺刚, 王磊. 气井的流固耦合渗流场分析[J]. 断块油气田, 2012, 19(4): 481-484.
- [79] 郭君, 张阿漫, 杨文山, 李佳. 基于谱单元方法的平板冲击流固耦合研究[J]. 振动与冲击, 2012, 31(2): 81-85.
- [80] 蔡定国, 姚育成, 徐业彬, 唐金权. 基于流固耦合分析的干式变压器温度场数值分析[J]. 华北电力大学学报, 2012, 39(2): 34-38.
- [81] 陈华磊, 买买提明·艾尼, 王豫鄂, 虞烈. 旋转对称支承机座的热分析与变工况计算[J]. 机床与液压, 2013, 41(1): 35-37.

- [82] 吴作伟, 梁闯, 郭海霞. 基于双向流固耦合的动车组水箱强度分析[J]. 北京交通大学学报, 2013, 36(6): 42-46.
- [83] 尚翠霞, 纪兴华, 孙贵斌, 王勇. 流固耦合作用下控制阀的稳压分析与设计[J]. 齐齐哈尔大学学报(自然科学版), 2012, 28(4): 85-88.
- [84] 吴正人, 王松岭, 戎瑞, 孙哲. 基于流固耦合的离心风机叶轮动力特性分析[J]. 动力工程学报, 2013, 33(1): 53-60.
- [85] Akinci N, Ihmsen M, Akinci G, Solenthaler B, Teschner M. Versatile Rigid-fluid Coupling for Incompressible SPH. ACM Transactions on Graphics, 2012, Vol. 31, No. 4, Article 62: 1-8.
- [86] Chentanez N, Goktekin T G, Feldman B E, O'Brien J F. Simultaneous Coupling of Fluids and Deformable Bodies. Proceedings of the 2006 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, 2006, Pages: 83-89.
- [87] Losasso F, Irving G, Guendelman E, Fedkiw R. Melting and Burning Solids Into Liquids and Gases. IEEE Transactions on Visualization and Computer Graphics, 2006, Vol. 12, No. 3, Pages: 343-352.
- [88] 李田, 张继业, 张卫华. 高速列车流固耦合的平衡状态方法[J]. 机械工程学报, 2012, 49(2): 95-102.
- [89] 李田, 张继业, 张卫华. 横风下高速列车流固耦合动力学联合仿真[J]. 振动工程学报, 2012, 25(2): 138-145.
- [90] 陈云敏, 谢焰, 詹良通. 城市生活垃圾填埋场固液气耦合一维固结模型[J]. 岩土工程学报, 2006, 28(2): 184-190.
- [91] Yang L, Li S, Hao A, Qin H. Realtime Two-way Coupling of Meshless Fluids and Nonlinear FEM. Computer Graphics Forum, 2012, Vol. 31, No. 7, Part 1, Pages: 2037-2046.
- [92] Müller M, Schirm S, Teschner M, Heidelberger B, Gross M. Interaction of Fluids with Deformable Solids: Research Articles. Computer Animation and Virtual Worlds, 2004, Vol. 15, No. 3-4, Pages: 159-171.
- [93] NVIDIA Corporation. NVIDIA CUDA C Programming Guide, version 3.2, 2010.
- [94] Khronos OpenCL Working Group. The OpenCL Specification. Version: 1.0.43, 2009.
- [95] Zhiyong Yuan, Weixin Si, Xiangyun Liao, Zhaoliang Duan, Yihua Ding et al. Parallel Computing of 3D Smoking Simulation Based on OpenCL Heterogeneous Platform. Journal of Supercomputing, 2012, Vol. 61, No. 1, Pages: 84-102.
- [96] Jens Breitbart, Claudia Fohry. OpenCL-An Effective Programming Model for Data Parallel Computations at the Cell Broadband Engine. 2010 IEEE International Symposium on Parallel & Distributed Processing, Workshops and Phd Forum(IPDPSW 2010), April 2010, Pages: 1-8.
- [97] Wenyu Zhang, Li Zhang, Shangmin Sun, Yuxiang Xing, Yajie Wang et al. A Preliminary Study of OpenCL for Accelerating CT Reconstruction and Image Recognition. 2009 IEEE Nuclear Science Symposium Conference Record(NSS/MIC 2009), 2009, Pages: 4059-4063.
- [98] Francois Dagenais, Jonathan Gagnon, Eric Paquette. A prediction-correction Approach for Stable SPH Fluid Simulation From Liquid to Rigid. CGI2012, 2012, Pages: 1-10.
- [99] Oleksiy Busaryev, Tamal K. Dey, Huamin Wang, Zhong Ren. Animating Bubble Interactions in a Liquid foam[J]. ACM Transactions on Graphics, 2012, Vol. 31, No. 4, Article 63: 1-8.
- [100] Ihmsen M, Akinci N, Akinci G, Teschner M. Unified Spray, foam and Bubbles for Particle-based Fluids. Visual Computer, 2012, Vol. 28, No. 6-8, Pages: 669-677.

作者简介

袁志勇 CCF 高级会员，CCF 虚拟现实与可视化技术专业委员会委员，武汉大学计算机学院教授、博士、博士生导师。主要研究方向为计算机仿真与虚拟现实、图形图像处理、嵌入式系统等。E-mail: zhiyongyuan@whu.edu.cn。



廖祥云 武汉大学计算机学院博士研究生。主要研究方向为计算机仿真与虚拟现实、图形图像处理。E-mail: xiangyun-l@163.com。



关键词索引

- GPU 虚拟化 263,273,274,276,308
OpenFlow 63-71,76,77-82,88-90
安全软件 131,134,136,140-144
存储系统 83,238-240,244,245,247-256
多模态物理过程 351-354,362-366
恶劣环境 50,57,129-132,134,136-139,
141-144
高可靠微处理器 173,174,178,196
高效能计算 38
高性能计算机系统 38,39,44,51,55,
57,59
基于知识的建模 101,107,113,117
集中控制 63,81,84,87
计算机辅助设计 99,100
计算机图形学 99,100,107
加固技术 136
检索 1,2,6-28,103,107-109,113-114,
250
可靠性评估 174,181,185,186,192-194
可靠性设计 174,175,177-179,192-195
可视媒体 99,101,107,112,113,116,119
控制平面 63,65,68-72,79,80,84,89-91
跨媒体 1,2,11-17,20,21,24,27,28
冷却技术 39-47,49-51,56-61,136
耦合仿真 354,357,363
普适服务 159
普适计算 148-152,158-161,165,
167-169,222,333,342
嵌入式系统 129-131,137,141-143,169,
205,206,213-215,222,227,264-266,
270-272
嵌入式虚拟化 263-269,271-273,308
情感计算 331,332
人机交互 148,154,156-158,166-169,325
容错技术 130,131,132,137,142-144,
177,182,190,195,196
软件测试 203-207,209,211,212,216,
217-219,222-224,227,228
软件定义网络 63-65,67,68,80,85,89
软件工程 203,204,206,214,219,227,229
社会网络 101,113,150,153,320-325,
328,334-338,341
社会智能 320,325,328,329
事件识别 2-6,17,18,21,22,25-27
视频 1-4,11,12,17-27,100,103,107,
113,152,153,195,267,290,325,331,
337,338,342
数据平面 63,65,69,72,74-80,90,91
数据中心 59,64,79-86,89,91,174,218,
246,248-251,253-256,262,263,284,
289,290,292,293,300,339
数据组织 239,252
数字城市建模 99,100,103,105,110,
115,118
瞬态故障 174-178,222
索引 2-28,246,325,
体网格生成与编辑 99,105,111,115,118
系统结构 132,137,219,244,255-257
协同计算 319,325,329,330,332
虚拟化 65,66,71-78,83-86,248,253,
255,262-284,289,293-308
虚拟化安全 263,294,295,307,308

虚拟现实 2,351,363,366

虚拟资源调度与管理 262

虚拟资源性能计量 263,279,282,308

永久故障 178,180

用户模型 148

云计算 118,149,153,154,158,168,174,

205,218,219,223,249,252,263,279-

302,305-308,323,339,365

智能终端 148,149,158,165-167,325

作者索引

- | | | | |
|-----|-----|-----|-----|
| 艾列富 | 36 | 罗英伟 | 318 |
| 白晓颖 | 236 | 牛文生 | 146 |
| 鲍虎军 | 127 | 阮 琳 | 62 |
| 曹 强 | 261 | 孙宇清 | 349 |
| 陈宝权 | 128 | 孙允明 | 147 |
| 陈国华 | 350 | 汤 庸 | 349 |
| 陈海波 | 318 | 唐卫清 | 350 |
| 陈 鸣 | 97 | 王 峰 | 62 |
| 陈 旭 | 62 | 王 乐 | 147 |
| 冯结青 | 128 | 王 莉 | 349 |
| 高曜明 | 128 | 文鹏程 | 147 |
| 郭 斌 | 172 | 吴朝晖 | 172 |
| 胡 斌 | 349 | 吴春明 | 97 |
| 黄 兴 | 350 | 吴 飞 | 36 |
| 江建慧 | 202 | 夏欢欢 | 350 |
| 兰巨龙 | 97 | 谢高岗 | 96 |
| 李 丹 | 97 | 谢建春 | 147 |
| 李华伟 | 202 | 徐宝文 | 237 |
| 李建国 | 350 | 徐常胜 | 37 |
| 李石坚 | 172 | 徐 凯 | 128 |
| 李 伟 | 98 | 徐明伟 | 98 |
| 李晓维 | 202 | 鄢贵海 | 201 |
| 李 星 | 98 | 于俊清 | 36 |
| 李亚晖 | 147 | 於志文 | 172 |
| 廖祥云 | 372 | 袁志勇 | 372 |
| 廖小飞 | 318 | 张常有 | 350 |
| 刘 超 | 236 | 张天柱 | 37 |
| 刘永进 | 128 | 张喜民 | 147 |
| 卢 暇 | 349 | 赵 莎 | 172 |
| 罗军舟 | 98 | 赵小冬 | 147 |
| 罗腊咏 | 98 | 庄越挺 | 36 |