

面向数据体系结构初探

苗放 周相兵 / 著

The Study of Data Oriented Architecture



科学出版社

面向数据体系结构初探

苗 放 周相兵 著

科学出版社

北 京

内 容 简 介

本书主要介绍了大数据软件工程——面向数据体系结构的相关思想以及实践基础、案例,集中反映了作者近来对面向数据的软件工程理解和相关思考,阐述了大数据软件工程的支撑基础、实践思想、技术要领和应用方法。全书共分4章,包括面向数据的软件发展历程,软件技术发展阶段,软件模型、开发方法、软件开发建模,数据方法、技术进展及其比较;面向数据体系结构(DOA)的基本思想、技术概况、支撑基础,DOA的基本描述、构成、安全机制;面向数据的DOA分类、体系结构;最后在互联网+思维下,以面向数据的智慧旅游公共服务平台构架为例进行了分析。本书大部分大数据软件工程的相关思考及理解均为作者的成果。

本书可以作为高等院校计算机、信息管理及相关专业本科、研究生指导用书,也可供企事业单位、科研机构等部门从事大数据的相关人员和教师参考。

图书在版编目(CIP)数据

面向数据体系结构初探 / 苗放 周相兵著. —北京: 科学出版社, 2017. 3
ISBN 978-7-03-052167-5

I. ①面… II. ①苗…②周… III. ①数据处理
IV. ①TP274

中国版本图书馆 CIP 数据核字 (2017) 第 054187 号

责任编辑: 冯 铂/责任校对: 韩雨舟
责任印制: 罗 科/封面设计: 墨创文化

科学出版社 出版

北京东黄城根北街16号

邮政编码: 100717

<http://www.sciencep.com>

成都锦瑞印刷有限责任公司印刷

科学出版社发行 各地新华书店经销

*

2017年3月第 一 版 开本: 787×1092 1/16

2017年3月第一次印刷 印张: 11.5

字数: 280千字

定价: 79.00元

(如有印装质量问题, 我社负责调换)

序

大数据时代，正在改造信息技术体系结构和产业发展模式，成了新一代信息产业革命的驱动器。为了促进互联网经济向各领域的全面渗透，需要更多智能化、易用性、实时性的信息软件系统支撑平台和产业安全平稳运行。大数据作为信息革新的血脉，正在促进软件工程发生互联网化变革，但以数据为导向的软件工程正处于发展阶段。数据作为互联网软件的核心内容，怎样用数据合理表达软件体系结构、怎样用数据有效描述软件自动生成过程成为了面向数据的互联网软件发展的重要课题。

苗放教授作为业界较早研究面向数据的软件体系结构方面的专家，从面向服务体系结构（SOA）的软件构架模式中获得启示，以云计算软件为基础，提出了一种面向数据的软件体系结构（DOA）模式及思想，指出以数据激活软件，以数据价值为软件运行动力，以数据流动更新软件，以数据审计确保软件体系安全；从而推动互联网软件体系变革，引导面向数据的软件体系结构能落地开花。更难能可贵的是，作者还以构建智慧旅游公共服务平台为基础，实践了面向数据体系的软件平台，其具体模式还应用到教育信息化、应急指挥等领域，促进了业界软件构架模式的革新，并产生了不错的影响，成了面向服务体系结构的重要补充和延伸。

作者睿智的观点和对数据的理解，来自其数年对软件相关理论的深刻认知和实践经验的总结，相信能成为业界的一本重要参考书，成为面向数据体系的软件框架模式领域的重要参考方案。

苗放等教授所著的《面向数据体系结构初探》一书，既有理论性阐述，又有具体的实践案例，是理解面向大数据软件工程的一本不可多得的书。该书描述了大数据软件工程的支撑基础、实践思想、技术要领和应用方法等重要内容，值得读者仔细研读和深思。

中国电信四川分公司总经理
四川省大数据产业联合会会长
成都市云计算产业联盟理事长

郝永刚

前 言

随着信息技术向数据技术变迁,计算机技术与网络技术发生了本质性变革,“互联网+”成为技术革命和产业变革的驱动力,促使软件构架模式面向模块化、面向层次化、面向服务架构(Service-Oriented Architecture, SOA)发展。软件系统也相应的发生更新,从最初面向命令集、面向过程、面向对象、面向构件、面向服务、面向云计算服务发展到目前以面向数据为导向,面向云计算为基础的多种模式并存的时代。特别是大数据时代的到来,软件系统中的信息表达发生了根本性变化,即在信息生命周期内表达呈多样性、实时性和价值性。因此,本书从面向数据维度,以数据为引导,初探面向数据架构(Data-Oriented Architecture, DOA)模式,重点从软件技术发展历程、软件开发方法,以及大数据软件等进行总结概述,并进行对比分析,从而给出DOA的提出背景。同时,为了更好地表达DOA,提出DOA构成、核心部件组成和面向互联网的体系结构,最后以在互联网下构建的智慧旅游服务平台为案例进行详细分析。

本书共有4章。

第1章,DOA提出背景。主要比较信息技术与“互联网+”、分析软件技术发展历程、软件模型与开发方法分析、面向服务的软件建模技术以及面向数据的软件描述、大数据软件总结对比。

第2章,DOA提出及基本构成。主要分析DOA支撑基础、DOA概览、DOA基本结构、DOA基本构成。

第3章,DOA体系结构。主要从行业分类、业务流向分类、体系结构分类和数据形态分类四个角度分析DOA分类;从云计算、大数据、服务/数据注册和可视化角度分析DOA体系结构。

第4章,DOA实践初探。主要在“互联网+”背景下,以DOA思维为基础,以一种智慧旅游服务公共平台为案例进行较为详细的探索。

本书在编写过程中要特别感谢成都信息工程大学马洪江教授给予的无私支持与帮助;感谢杨文晖老师的支持和给予的耐心,感谢博士解岩、李一平给予的帮助,感谢软件工程师陈亮、廖山清、张敏、沈少朋给予的全力协助,正是有了他们的支持和帮助,本书才得以出版。

本书研究结果是国家自然科学基金面上项目“客户端聚合服务的空间信息网络服务模式研究”(61071121)、“智慧旅游”四川省高校科研创新团队(15DT0039)、阿坝州科技成果转化项目“面向移动终端的阿坝州智慧旅游服务平台产业化开发”(CGZH2016007)、阿坝州科技计划项目“智慧旅游在阿坝州示范应用研究”、国家统计局统计信息技术与数据挖掘重点开放实验室开放项目“旅游大数据统计特征研究”(SDL201607)、四川省旅游

业青年专家 2016 年度培养计划项目“旅游大数据特征研究”(SCTYETY2016LE08)、西华大学省部级学科平台开放课题“云计算环境下一种基于位置服务的社会网络推荐方法研究”(SZJJ2015-060)、四川省哲学社会科学重点研究基地——四川旅游发展研究中心项目“互联网+下的四川智慧旅游电商发展机制研究”(LYC16-29)和广西省多源信息挖掘与安全重点实验室开放基础项目“云计算环境下一种基于 SLA 保障 QoS 的服务收益机制研究”(MIMS14-05)等科研项目阶段性成果。

本书第 1 章由苗放、周相兵撰写,第 2 章由苗放撰写,第 3 章由苗放、周相兵撰写,第 4 章由周相兵撰写,全书由苗放统稿。

本书在撰写过程中,虽参考了大量中外文献资料,但由于未能查阅更为广泛的文献和篇幅原因,可能相关文献未能一一列出、标出,在此对所有参考文献作者表示诚挚的谢意。

本书在研究和撰写过程中,虽几经努力,但限于作者能力与水平,难免有不足之处。因此,恳求各位专家、学者和广大读者批评指正。宝贵的建议或意见可通过邮箱 mf@cdut.edu.cn 反馈,谢谢!

苗 放

2016 年 10 月

目 录

第 1 章 DOA 提出背景	1
1.1 IT 与“互联网+”简要比较	1
1.2 软件技术发展分析	3
1.3 软件模型与开发方法分析	7
1.4 面向服务的软件开发建模概述	10
1.5 面向数据的软件描述	12
1.6 面向大数据的软件发展分析	30
1.7 小结	38
参考文献	38
附录 1: 本章缩写词	40
附录 2: 本章涉及的相关大数据技术名称及网址	42
第 2 章 DOA 提出及基本构成	44
2.1 DOA 提出	44
2.2 DOA 概览	46
2.2.1 DOA 的基本原理和可能价值	47
2.2.2 DOA 可以解决的问题	47
2.2.3 不同视角理解 DOA	48
2.2.4 DOA 可能的方法	49
2.3 DOA 支撑基础描述	49
2.3.1 DOA 支撑基础之云计算描述	49
2.3.2 DOA 支撑基础之大数据描述	53
2.4 DOA 结构基本分析	55
2.5 DOA 基本分析及构成	59
2.5.1 面向数据的软件体系结构的机制	60
2.5.2 面向数据的软件体系结构的组成	62
2.5.3 数据注册中心 (DRC) 机制	63
2.5.4 数据权限中心 (DAC) 机制	64
2.5.5 DOA 安全机制	66
2.6 小结	68
参考文献	69
附录 1: 本章缩写词	73
第 3 章 DOA 体系结构	75
3.1 面向大数据 DOA 分类	75

3.1.1 按行业分类	76
3.1.2 按业务流向分类	78
3.1.3 按构架体系分类	80
3.1.4 按数据形态分类	110
3.2 DOA 体系结构	112
3.2.1 云计算基础环境	117
3.2.2 大数据环境与数据注册机制	120
3.2.3 数据注册应用与云计算服务组装	122
3.2.4 前端可视化体系	126
3.3 小结	128
参考文献	128
附录 1: 本章缩写词	132
附录 2: 本章涉及的相关可视化技术名称及网址 (URL)	133
第 4 章 DOA 实践初探	138
4.1 基本概述	138
4.1.1 智慧旅游	139
4.1.2 位置大数据	139
4.2 相关进展	141
4.2.1 智慧旅游与“互联网+”	141
4.2.2 面向数据的分析处理	142
4.2.3 基于位置的推荐技术	144
4.2.4 社会网络化	147
4.2.5 旅游数据与空间地理位置信息关联	151
4.3 “互联网+”下面向 DOA 的体系结构	152
4.4 旅游大数据构成及处理方法	154
4.4.1 旅游大数据特点	154
4.4.2 旅游大数据构成	156
4.4.3 旅游大数据分析技术	157
4.5 基于 DOA 构架的 STSP 实践	158
4.6 实现结果与分析	162
4.6.1 实验场景说明	163
4.6.2 不同浏览器的可视化性能	164
4.6.3 云平台运行性能	165
4.6.4 面向数据分析的云服务即时组合	167
4.6.5 面向数据分析结果的用户访问页面重构	168
4.6.6 基于位置的社会化网络推荐	170
4.7 小结	170
参考文献	172
附录 1: 本章缩写词	175
后记	176

第1章 DOA 提出背景

信息技术 (Information Technology, IT) 影响且改变整个信息化世界, 促进了全球信息化, 实现了全球空间一体化的互通, 使得信息化成为世界各国重要的发展战略, 并为各国经济发展贡献了重要的基础性战略资源, 形成具有独有和整合性的信息产业, 并能带动其他产业与信息产业间关联发展; 最终致使信息产业的快速发展且提高了生产效率, 改善了生产力方式, 改变了社会形态, 使全世界联系在一起, 形成了满足信息化特征的数字世界或数字地球。但随着网络体系的演化、人们活动范围变迁、生产结构多维化、人们对信息化技术日益依赖, 以及网络速度提升、分布计算能力的增强、终端应用普及化; 同时伴随日益产生的丰富的数据量, 人们获取信息的多样化需要, 要求信息表达具有较好的体验性、实时性、主动性和完备性, 使得经典的 IT 结构、模式、方法和技术体系不能满足变化的信息化要求, 更不能表达新型信息化特征以及空间化和多样化变化要求, 如云计算升级了服务计算, 变革了计算模式, 改变了传统的分布式计算结构, 提高了网络计算效能, 降低了网络应用成本, 使人们及时获取有效的信息结果成为可能; 大数据改变了人们对数据的认知、管理模式, 使数据具有生命及无处不在的价值, 使数据成为哲学层面的战略资源, 而且也能在信息产业中体现无尽的价值, 大数据时代的到来, 成为本轮信息及产业革命的基石^[1]; 物联网的产生及发展, 并在互联网的支撑下使信息世界不但实现了互联, 还实现了互通, 使人与物、物与物可以实现通信, 并在云计算和大数据等的支撑下, 可以促使相关事物具备智能化功能, 达到自动认知识别、通信交流^[1-2]。

1.1 IT 与“互联网+”简要比较

自从 2015 年 3 月“互联网+”时代的到来, 目前正在变革的 IT 技术体系及产业结构, 正逐步替代传统的 IT 模式, 且重新定义了 IT, 使人们必须重新认知 IT 结构, 即从互联网化、数据化和终端应用可视化角度重新理解信息化, 经典的 IT 结构及技术即将成为历史, 成为信息化世界的基本组成单元。其中数据就成为新一轮互联网变革的原动力, 从而使一切互联网化活动和新型技术的应用都与数据相关, 使得数据成为互网络中的“石油”, 要求不同的应用场景需要根据各自活动来体现数据价值及预测即将发生的活动, 促进新一轮产业革命, 以构建可持续化循环、低碳节能型、创新型社会形态, 加快从传统工业化、粗放型向轻工型及现代服务型转型升级^[3-5]。

“互联网+”战略的实施与实践, 关键在于有新一代云计算、大数据、物联网、位置计算等新型技术作为支撑, 关键是能实现资源互联网化、数据化和终端应用可视化; 处于核心地位的就是大数据所表达的信息开放性、针对性、精准性和流动性且在不同网络应用领域所能体现的价值性; 价值流向的重要基础就是云计算给予了大数据分析处理的底层计

算。不过，目前数据应用水平受制于“互联网+”技术体系的成熟度和应用广度，数据活化受限于数据理论发展水平及技术处理能力，数据终端应用及价值体现来源于人们对“互联网+”的接受能力和效率；面向数据的软件工程还处于起步阶段，而且受限于数据价值体现及数据网络化程度。表 1-1 是“互联网+”模式与 IT 模式的区别。

表 1-1 “互联网+”与 IT 模式的区别

模式	特征	优势	挑战
IT	①帮助政府、企事业单位实现信息化	①有效实现信息共享，提高工作效率	①弹性与可伸缩性不足
	②关注需求和业务走向，以需求和业务变化影响 IT 结构变化	②降低了企业运营成本和管理成本，使信息流通更快捷	②难以预测信息走向及未来变化情形
	③偏重信息化管理能力，强调信息化直接应用	③改变了整个社会形态格局，使信息无处不在	③无法估量信息的价值
	④全面促进了网络技术与信息化产业发展	④增强企业经营能力和革新了交易模式	④产生信息鸿沟与信息不对称现象
	⑤引领了信息化新理论、新方法与新技术的出现	⑤快速改变了人们生活模式，使世界变得“更小”	⑤不能有效释放信息内在动力
	⑥IT 结构自主演化能力弱，模式较为固化	⑥允许根据需求逐步演化成信息化产品	⑥人们被动接收信息，缺少主动性
“互联网+”	①从战略高度构成一种新产业态	①引发社会协同分工形态变革，优化资源配置	①数据安全及隐私目前难以得到保障
	②将 IT 构架升级为 DT 构架，形成互联网技术构架	②对数据资源的释放价值和流向	②“互联网+”模式处于初期发展阶段
	③终端服务个性化应用	③模式更扁平网状化	③机遇与挑战并存
	④促进产业重构、重组、重分发及互联融合	④形成了多方融合的新网络基础体系结构	④对“互联网+”模式处于认知阶段
	⑤跨界融合创新、颠覆与尊重人性，发掘人的价值	⑤正哺育一批新型轻量的互联网产业	⑤不能很好地实现创新价值的跨界传递
	⑥是一种“开源、开放、共创、共享”的思维模式	⑥成就了互联网梦想，让有梦想的人可以任情发挥	⑥需要从不同视角认知“互联网+”
	⑦引领了产业变革和经济转型升级	⑦是一种具有战略与技术、产业结合的系统性工程	⑦上中下游产业链式互联网变革

“互联网+”应用、发展以及商业市场智慧化不仅受到顶层战略设计和发展的影响，还受到所支撑的商业运维服务平台健壮性、可用性和易用性的制约；同时也受到传统软件工程思想牵制、经典软件开发策略思维限制。因此，以 IT 为导向的软件工程正在逐步演化为以“互联网+”及技术体系为思想的新型软件工程，以实现真正意义上的软件自动生成、构架演变、自主演化和软件可重用等，从而延长软件生存周期，提升软件自适应能力；而互联网软件工程其核心是在互联网下以数据为中心进行软件构架及软件生成，最终形成以云计算服务与数据相结合的软件发展新模式和新方法，即以基础设施即服务

(Infrastructure as a Service, IaaS)、平台即服务 (Platform as a Service, PaaS)、软件即服务 (Software as a Service, SaaS) 和数据即服务 (Data as a Service, DaaS) 等多种服务方式呈现给用户。

1.2 软件技术发展分析

起初软件工程是用于解决大规模软件应用过程中所表现出的不确定性、不统一性、脆弱性、软件生存短暂性等问题,以及缺乏系统性和工程性所带来的软件开发周期长,易随需求变化及功能改变导致整个过程失效等不足。因此,软件工程的出现有效解决了所面临的危机(即软件危机),不过,随着网络环境变革、软件复杂度增加、人们对软件需求的变化及体验效果的改变,使得传统的软件工程指导思想日益不能满足“互联网+”要求下的软件发展需要,特别是大数据时代的到来,加剧了软件体系结构和组织结构的变化,革新了传统软件工程开发模型。针对这些事实,近年来,国内外研究人员相继提出了网构软件、网络化软件、云计算软件、大数据软件等,并给出了相应的软件构架方法,如针对网构软件所提出的 ABC (Architecture-Based Component Composition) 方法^[7],基于 ABC 的软件开发流程如图 1-1 所示^[8];网络式软件的角色—目标—流程—服务 (Role-Goal-Process-Service, RGPS) 需求元建模框架方法^[11],云计算软件的 IaaS、PaaS、SaaS 和 DaaS 等,大数据软件的面向数据架构 (Data Oriented Architecture, DOA)。其中,DOA 为本书重点讨论的内容。

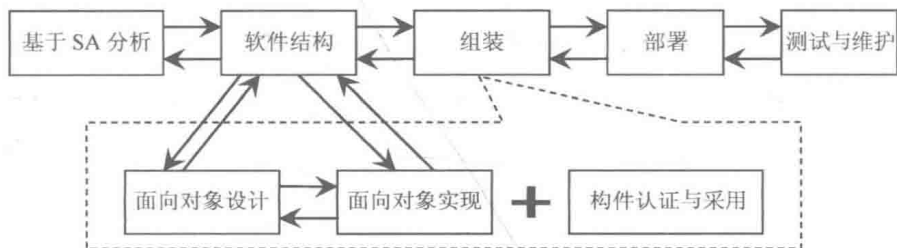


图 1-1 基于 ABC 的软件开发流程结构图

互联网时代开启了软件工程变革,促使经典软件工程成为互联网时代软件工程的重要组成部分,而且也从本质上更新了软件开发方法,如传统的软件开发方法主要采用自上而下、逐步求精迭代递进的方法推进软件开发,并根据需求特征对整体软件进行分解开发的策略推动软件一步步走向一定成熟度,使整个软件开发处于可控范围;但只要需求改变了,整个软件计划也要随之变化,需要重新对功能进行重定义和重拼接,而且软件也是随着软件结构、模型及平台的变化而变化,在一定程度上难以满足软件开发的自适应要求,更难以满足具有开放化、开源化、普适化、云计算化和数据化等互联网时代软件特征要求,如经典的基于构件的软件开发 (Component-Based Software Development, CBSD) 通过自适配器、构件组装方法和构件更新方法来改变软件结构,通过各构件间的层次及粒度关系来控制软件生成。

软件计算范式角度：截至目前，软件计算经过了多种计算模式及其模式的扩展，共经历了单机版的计算模式、面向基础网络技术的客户端/服务端（Client/Server, C/S）模式、面向基础互联网的浏览器/服务器（Browser/Server, B/S）模式、传统意义上的分布式计算模式、网格计算模式、普适计算模式、服务计算模式、语义网（Semantic Web）及万维网（World Wide Web, WWW）计算模式和“互联网+”计算模式等^[6-9]。

网格计算的目的是有效整合互联网上的资源，为用户提供一种资源集约化的服务。但由于网格计算无法有效表达资源的弹性和可伸缩性，也难以解决资源聚合、分配和管理等问题，使其难以在互联网下发挥强大的计算能力和资源调配效率，相应地对软件的支撑能力也无法发挥出来，但能在一定程度上为 B/S、C/S 提供软件运行支撑。

普适计算的目的是为用户提供一种无处不在、随时随地、具有普遍性的服务计算模式，努力将空间一体与人们生活的物理空间关联在一起，形成具有整体性的人与物的网络世界，其实现结构就是从现有及发展中的网络为应用环境，以各种终端、嵌入式设备等表现方式为载体；但普适计算是一种宏大的计算模式，具有全域性、空天一体性、战略性等特征，使其在实践中更多的表现为一种理念、一种期望和一种指导思想，难以实现商业活动所需要的计算要求，而且具有连续、离散与聚合等属性的物理世界与现实融合还需要更多的演化和递进过程。不过普适计算可以为软件运行提供顶层及战略层面的指导。

服务计算的目的是以服务（Web Service）为基本单元实现跨平台、跨语言、跨系统、跨结构等的集成，并以服务组合方式快速、松散、灵活以及便捷地生成增值服务和 service 应用系统，其服务组合效率直接影响软件质量、可用性和成熟度，且以服务水平协议来契约服务组合过程中的规则，以服务质量衡量服务组合结果；从此使人们不在单一关注软件功能本身，而是通过服务的方式为用户提供所需要的功能。目前，互联网中存在大量的服务单元，缩短了软件开发及生成周期，但服务本身的可信性、服务快速发现、服务应用可扩展性、服务自动识别认知、服务协同等问题还需进一步研究。不过，服务计算已改变了传统意义上的软件开发模型，而且有效弥补了网格计算的不足，丰富了普适计算内涵，也促进了云计算发展，使其成为云计算的重要组成部分。

语义网及万维网计算的目的是通过对 Web 语义化处理后能在 WWW 中具有自适应性和自主性，提高网络资源共享和定位能力，以便于机器理解识别、交换与自动处理，最终使互联网能有机统一融合，实现 Web 语义智能化。但语义化后的智能程度还面临语义描述方法、元数据构建、结构构造以及日益增长的数据资源真实性、可信性描述等问题，怎样让机器理解无歧义语义信息还面临巨大挑战。而 WWW 作为一种描述网络演化进程的模式，通过分析用户上网行为来达到理解网络发展趋向，强调将社会学、心理行为学、物理学和管理学等学科相融合来研究 WWW 的发展模式和虚拟社会的结构，即 WWW 的研究将有助于发展和理解普适计算，使普适计算向纵深发展，以全面支持互联网发展，形成多种网络结构，如从 Web 1.0 到未来的 Web 5.0（表 1-2 是 Web 1.0 至 Web 5.0 的比

较)，以及社交网络（社交网络服务，Social Network Software/Social Network Site/Social Network Service/Social Network System，SNS）。

表 1-2 Web 1.0 至 Web 5.0 的区别

互联网技术	Web 1.0 (1990~2000 年)	Web 2.0 (2000~2010 年)	Web 3.0 (2005~2020 年)	Web 4.0 (2012~2030 年)	Web 5.0 (2020~2050 年)
特征	基本 Web、连接知识、具备基本信息交互能力，以信息编辑为主	Web 较为深度应用，形成社会网，加强了用户与网络之间的互动，实现了双向交流，提升了交互能力；完全基于 Web，并通过浏览器来完成工作，并可以拥有自己的数据	信息间可以交换和互动，并实现信息整合和分类，实现信息有效聚合，丰富普适计算，提倡个性化定制；突出数据的重要性	实现分配、预测，挖掘潜在价值，实现主动积极的服务，真正意义实现普适计算，且全面终端智能化	实现网络自由化和智能化，充分体现价值性和体验效果精准性，体现语用性
关键词	基本信息共享	信息共享、共建	知识识别	知识认知	语用性（自主智能化）
关键技术	ASP、JSP、LAMP、Perl 等	J2EE、.NET、Web Service、PHP、开源软件技术 SSH 框架等	Web Service、语义技术、推荐技术、云计算、大数据、移动互联网、物联网、人工智能等	大数据相关技术、云计算、物联网技术、高速互联网技术、人工智能	大数据及新型互联网技术体系、高级人工智能技术
软件模型/方法	面向过程	面向对象程序设计（如 UML、MDA）、组件、构件	SOA、DOA 等	SOA、DOA 等	DOA、POA 等
标记语言	HTML 等	HTML、XML 等	HTML、XML、OWL-S、HTML5 等	HTML5、HTML 等	HTML5、语用描述等
传输协议	HTTP、FTP 等	HTTP、HTTPS、P2P 等	HTTP、HTTPS 等	HTTP、HTTPS 等	HTTP、HTTPS 或其他协议
网络环境	简单 B/S	分布式 B/S	分布式 B/S 和 C/S（智能终端）	分布式 B/S 和 C/S、G/S（通用浏览器/服务云）	分布式 B/S 和 C/S、G/S（智能终端）或空间一体化模式

续表

互联网技术	Web 1.0 (1990~2000 年)	Web 2.0 (2000~2010 年)	Web 3.0 (2005~2020 年)	Web 4.0 (2012~2030 年)	Web 5.0 (2020~2050 年)
主要表达方式	直接搜索引擎、网站、数据库、信息管理系统、文件服务器、邮件服务、即时通信等	社区服务、RSS、维客、社会化网络、面向服务的应用系统、博客等	基于本体的语义化、人工智能、智能代理、知识结点、语义知识管理等,如微信、微博、大型电商平台等	体现数据价值,实现智能化应用,提供更为便捷的服务	对兴趣、爱好、行为、语言、主题、职业、专业进行聚集和管理,实现价值体现标淮性

“互联网+”计算模式不但描述了普适计算、服务计算、SNS、语义网与 WWW 等特征,还产生了一批重要的新型技术体系(主要含高速网络技术、云计算、大数据、物联网、空间信息技术等基础性技术体系)来支持互联网全域发展,更重要的还是一种商业计算模型,能全面指导当前产业结构的优化调整和转型升级,实现创新驱动、节能减排、尊重人性关怀、开放互联互通,最终实现可持续化与协同发展,实现全域互联互通,形成智慧化世界。同时“互联网+”还描述了互联网技术体系战术层面的实现,如可以在不同领域构建不同的智慧化服务平台,改造不同领域落后的发展模式,升级相关领域的产业结构等,以实现不同的新型产业形态。因此,“互联网+”的发展和实践,将彻底改变经典的软件工程思维,也将形成一种更加开放、开源且自由的软件形态,促进软件技术发展,促使新型软件开发技术形成。

上述几种主流软件技术计算范式,都是以网络为基础展开的,其主要集中在网络资源需要集约化应用管理、应用要求体验效果需要更为主动、获取网络服务需要更为便捷等,以最终形成一体化、融合化的新型计算结构与体系;并能推动产业变革,使信息产业向创新型发展,为现代服务业提供更加有效的支撑。如网格计算是从网络资源共享和集约化管理两个层次探讨未来的网络发展,不过就目前来看,网格计算已逐步被云计算所替代,这是因为云计算比网格计算更为简捷。服务计算更多是改变了软件功能的表达方式,即在网络中,以基于 XML (Extensible Markup Language) 的服务方式来描述 (Web Services Description Language, WSDL) 软件功能,并按可信度暴露在网络中供不同的软件应用,大大提高了软件在网络中的生命周期,为软件自动生成提供了重要的模式和基础。普适计算则从人机交互、直接体验效果和应用便捷性等方面探讨未来网络的发展及应用模式,并从战略层次指出了未来网络的发展方向。语义网是从人工智能角度描述自然语言及机器学习在网络中所能体现的资源共享和智能定位资源模式;而 WWW 则是从网络及网络用户关系、未来网络发展模式角度思考资源关系管理、用户行为关系分析、资源发现定位和应用模式,以支持用户快速且有效地应用网络中的资源。“互联网+”则是一种新型分布网络体系及结构,既包含了网络模式发展本身,也包含了网络商业结构及模

式，目前国内正在驱动技术更新和产业转型。

1.3 软件模型与开发方法分析

软件模型描述了软件实体和实体之间的交互连接关系^[7]（表 1-3），而软件技术是用来实现软件实体的方法及工具，软件结构和软件描述语言是用来描述软件的体系结构，软件程序是用来表达软件功能。自从软件诞生以来，经历了二进制编码、命令集合的初级阶段，后逐渐发展到语句表达简单且针对性的软件功能；最后发展到以程序语言来简化且替代了以编码、命令和语句为主的方式，即通过函数/过程，如 C 语言/Pascal 等；模块/模板，实现方式表现有执行程序/动态链接库等；接口/抽象数据类型/对象，如 C #/Java/ C ++ 等，相关框架如 Spring、Struts 等；组件/构件，如 OMG（Object Management Group）组织的 CORBA/CMM（Common Object Request Broker Architecture/Capability Maturity Model for Software）、上海普元构件、Windows 下提供的开发组件 COM/COM+（Component Object Model, COM）等；Web Service，如 CXF、AXIS 等；网构软件，即具有适应开发、动态、多变的互联网要求的个性化、持续成长等特征的新型软件形态^[7]；IaaS/PaaS/SaaS/DaaS，如服务器端实现语言 Python 等，涉及的云平台如 OpenStack、CloudStack 等。具体比较见表 1-4。

表 1-3 不同软件开发模型的基本区别

类型	类型重用	组件重用	互操作性	可升级性	复杂性	时间周期	可靠性	质量
瀑布模型	不能实现	不能实现	没有	不可以	高度复杂	长周期	低	低
原型模型	不存在	没有工件	未规定	不可以	高度复杂	不独立用	低	无
RAD 模型	不能实现	不存在	不支持	不可以	高度复杂	快速开发	弱	低
演化模型	可实现	难实现	不互通	可能	高度复杂	长周期	弱	中
OO 模型	可实现	没特性	允许	可以	一般	长周期	高	中
RUP 模型	可实现	不包含	高规定	可以	一般	大量时间	中	中
敏捷模型	支持	无属性	限制	可以	一般	快速开发	低	低
CBD 模型	高度支持	核心功能	高度支持	可以	不复杂	快速开发	高	高
MDA	支持	支持	支持	不确定	一般	较快速	中偏上	中偏上
SOA 模型	高度支持	全支持	高度支持	完全可以	不复杂	可快速开发	较高	较高
云计算 服务模型	高度支持	全支持	高度支持	完全可以	不复杂	可快速开发	高	较高
DOA 模型	高度支持	全支持	高度支持	完全可以	不复杂	快速开发且 能自动生成	高	高

表 1-4 常见软件模型与技术

	描述	代表语言或工具	特点
函数	在软件实体中用来做一件事的程序，具备在函数之间调用，形成更大的子程序、程序块	C 语言，通常可用 GNU 编译器套件来完成 C 程序编写及编译	基本数据类型、构造数据类型、指针类型和空类型四大类
过程	结构化语法严谨，过程层次分明，程序易写，并具备很强可读性	Pascal 语言，常用的编译器有 Turbo Pascal 或 Free Pascal	简单类型、构造类型和指针类型三大类
模块	通过函数或其他方式来实现一个特定功能	C 语言、Pascal、Visual Basic、Visual FoxPro 等	构成一组组具有重用性质的功能程序
模板	构建函数或过程或语言结构所构成的具有通用特点的特定功能	C 语言、Pascal、Visual Basic、Visual FoxPro、Visual C++、C++、Java、PHP、JSP、ASP、C# 等	可以根据功能的变化，在一空范围内允许对所形成的模板进行修改
对象	目前主流的程序开发方法，已扩展到多个程序领域，并已超越了程序开发本身和范围；同时也改变了软件开发的结构和形态，使软件开发变得更为容易与有效	Visual C++、C++、Java、C#、Visual Basic、Visual JS、Python 等	构成面向对象程序设计，每个对象都能够接收来自其他对象的信息、处理数据和向其他对象发送信息，具备重用性、灵活性和扩展性等多种特征
组件	表达一个软件实体的通用性功能，具有相对的独立功能，接口由契约指定，并与语境有明显依赖关系，可由第三方使用且能独立部署	通常有报表组件、图表组件、表格组件、数据处理组件等，也有 CORBA 等，还有诸如微软 COM/COM+ 组件等	是对常用数据和方法的简单封装，以减轻程序开发的复杂性
构件	包含了组件的属性，且符合一套特定领域接口标准并能实现的一组接口，软件实体中的功能就通过这些接口实现组装	CORBA、COM/DCOM 以及一些其他常用构件、不同软件平台拥有的构件	具有自描述性、集成性、定制性以及独有的连接机制
框架	是实现某一领域的软件实体，供同一领域应用程序部署及二次开发的软件	Spring、Struts、DWR、Hibernate 等	简化了软件实体开发，增强了软件的轻量性、重用性，降低了软件的复杂度，使软件开发更容易，并且通常是开源的
接口	软件实体中不同功能层之间的通信规则，而且必须由类提供的成员或实现其他接口	Java 程序接口、PHP 接口、.NET 程序接口等	在接口中将方法、属性、索引器和事件作为成员

续表

	描述	代表语言或工具	特点
Web 服务	是一种基于 XML 应用程序功能的描述方法，具有独立性、低耦合性、自包含性以及基于可编程的 Web 的应用程序	涉及的语言主要有 WSDL、UDDI、SOAP、OWS-L 等，开发语言主有 CXF、Axis 等，支撑环境主有 IBM WebSphere、JBuilder、Microsoft.NET、Eclipse 等	具备跨平台、跨语言、跨系统等多种特征，增强了软件健壮性，改善了传统软件的开发模式
软件服务	由一组 Web service 或 REST 组成的软件实体，且这些软件可以根据商业计划、企业、政务等变化，自动生成软件的能力	通常有 Alibaba、Google、Amazon 等云服务，云服务平台主要有 OpenStack、VMWare 等	在 Internet 下具备主动性、演化性、自动性、高可复用性等特性
数据服务	是一种以数据分析处理为结果的软件服务，企业相关的关键数据实体的操作；通常由事件驱动、按需应变、批处理等实现服务，并要求逻辑数据模型有保持一致的能力	Hadoop 系列数据分析处理软件、MongoDB、Stratosphere、Amazon 数据服务等	具备预测性、价值性、多样性等特性，能为软件生成提供

传统软件实体所具备的自包含性、功能独立性、可复用性和健壮扩展性已难以满足大数据、“互联网+”时代的要求，即从结构和表达角度要求软件实体从更加松散低耦合到能即时体现价值、具有更好的体验效果、具备普适服务等特点；从软件成长及自动生成角度要求软件实体具备自主适应化、协同合作性、自动感知认识性、实时随机反应性等新型互联网形态的软件特征；从软件实体之间的连接及交互能力角度，传统软件实体主要通过函数间传递、调用、继承等方式来实现软件功能之间拼接和通信，使得具有很高的耦合性，不易使功能扩展，也易造成功能逻辑复杂和关联关系紧凑。对面向服务的软件实体来讲，从表达方式上有效改变了传统的功能通信方式，即通过将软件功能发布成 Web Service 接口，然后通过接口来实现软件拼接，并组成新的软件实体，加之在互联网下有云计算和大数据的支撑，使整个软件实体具备了更多的灵活性和开放性，而且软件实体之间的连接也具备了多样性，如面向服务的 BPEL4WS (Business Process Execution Language for Web Services)、WS-BEPL (WS- * 系列标准和规范)，面向构件的刻画机制和协同机制，面向网构软件的连接子等^[7]；但增加了软件实体本身的可信性、安全性和自由性，使得难以有效把控软件实体的规模，在生成新的软件时也难以预测其可用性和健壮性。

传统的软件是为满足特定领域的需求而研发的，实体运行在相对静止、封闭、固定的计算机环境中，运行效果可以预测，软件规模可以预知，软件开发模型可以选择，对运行环境的依赖性相对较为简单，实体间的协作关系、同步异步关系可控，也较为单一，流程关系也是可操作的，而且对网络依赖程度也不大，这就直接造成了软件实体生命周期较