

Architecture and Key Technologies
of Cloud OS

云操作系统架构 与关键技术

王继业◎著



中国工信出版集团



人民邮电出版社
POSTS & TELECOM PRESS

Architecture and Key Technologies
of Cloud OS

云操作系统架构 与关键技术

王继业◎著



人民邮电出版社
北京

图书在版编目（C I P）数据

云操作系统架构与关键技术 / 王继业著. -- 北京 :
人民邮电出版社, 2019.4
ISBN 978-7-115-50240-7

I. ①云… II. ①王… III. ①计算机网络—网络系统
—系统设计—教材 IV. ①TP393

中国版本图书馆CIP数据核字(2018)第266420号

内 容 提 要

本书从基础篇、设计篇、部署篇、应用篇 4 个方面,详细全面地论述了云操作系统的架构、各组件技术原理、部署优化与典型应用。全书深入浅出地阐述了云操作系统的基本概念和发展历程、总体技术架构、内核设计、基础服务、外部接口、部署及优化方法、典型应用实践等内容。本书强调理论与实践相结合,通过学习本书,读者能达到熟悉和掌握云操作系统关键技术、设计和应用部署的目的。

本书适合作为云操作系统产品技术架构师的技术工具书,同时也可作为决策企业 IT 建设和技术方向
的管理人员、高等院校师生的技术与架构参考用书。

-
- ◆ 著 王继业
责任编辑 李彩珊
责任印制 彭志环
 - ◆ 人民邮电出版社出版发行 北京市丰台区成寿寺路 11 号
邮编 100164 电子邮件 315@ptpress.com.cn
网址 <http://www.ptpress.com.cn>
北京印匠彩色印刷有限公司印刷
 - ◆ 开本: 787×1092 1/16
印张: 21.5 2019 年 4 月第 1 版
字数: 404 千字 2019 年 4 月北京第 1 次印刷
-

定价: 169.00 元

读者服务热线: (010)81055488 印装质量热线: (010)81055316
反盗版热线: (010)81055315
广告经营许可证: 京东工商广登字 20170147 号

序 1

从 2006 年谷歌、亚马逊等互联网公司提出云计算的概念以来，云计算在技术深度和应用广度上均有显著发展，云计算与移动互联网、工业互联网、大数据、人工智能等技术相辅相成，成为信息社会的主要支柱，推动信息社会进入大智移云的时代。

大型互联网公司利用云计算平台扩展业务范围，已延伸出各类数据分析和计算存储的平台，基于这些平台又出现了商务、物流、金融、安全等各类创新产品，催生了新业态，推动了生产方式和商业模式创新。实践证明，大数据、云计算等信息通信新技术与实体经济的融合已成为传统产业升级的新动能和新兴产业发展的新引擎。

但是我国大部分企业（特别是传统行业里的中小企业）的互联网化还处在起步阶段，自身建设云计算平台的能力不足，业务上云的比例还很低。大企业建设的大型云计算平台可以为中小企业提供云计算服务。大企业既有以云计算为主业的互联网企业，也包括传统产业的行业龙头企业，后者由于对自身业务需求具有透彻的了解，其提出来的云计算操作系统架构更适合本行业及相关行业的需要。

国家电网公司作为能源行业的龙头企业，其云计算研究起步早、发展快，在社会上有了广泛的影响，其全球能源互联网理念引领行业的发展，为产业的产业互联网化提供了成功的示范。为了推广国网云的建设经验，推动中小企业云计算应用的创新发展，国家电网公司的王继业同志撰写了《云操作系统架构与关键技术》一书。该书以云计算核心技术——云操作系统为对象，从基础技术、设计理念、部署方法 and 应用前景等方面对云计算做了全面而深入的研究和论述，介绍了云计算技术在企业 IT 重构中的作用，对各类云计算应用实践做了

清晰的梳理和说明，对应用云操作系统过程中面临的设计和技术难点做了深入解析。作者有丰富的建设和应用云计算的实践经验与亲身体会，从使用者所需的知识结构和关键技术的方式来撰写本书，系统性和实用性是本书的特色。相信建设云计算平台和使用云计算业务或者开发云计算应用的技术人员和工程师都能从本书中受益。

希望本书的读者，在了解到云操作系统技术的同时，能够积极地投身到云计算产业实践中来，推动信息技术与实体经济深度融合，在应用中创新并加快企业数字化转型，为企业创造更大的价值。

中国互联网协会理事长
中国工程院院士



序 2

云计算作为新一代信息技术的重要组成部分，其技术和标准尚未被跨国公司所垄断。当前，OpenStack、Hadoop、KVM、Kubernetes 等开源云计算技术发展迅速，应用广泛，这有利于推广我国自主可控的云计算技术和解决方案。国产自主可控的云计算软硬件的性价比具有突出优势，而且目前基于开源技术体系的云计算解决方案已经较为成熟，缺少的是大规模的推广应用和实践经验。鉴于此，构建大部门、大企业自主可控的云计算，是摆脱跨国公司控制我国重要信息系统的切实可行的途径。我国已有一批企业有能力提供构建包括私有云、公共云、混合云等各类云计算的核心技术、软硬件装备和服务。国家电网公司作为目前国内最大的企业，世界五百强的第二名，在自主可控软硬件研发和应用方面取得了显著的成果，在云计算方向也有业界领先的产品和实践，应向业界分享自己的成果和经验，取得更大的效益。

2016 年我作为国家电网公司“电力云仿真实验室”项目验收专家委员会的组长，有机会到现场见证了项目建设的情况，该实验室的研究成果在国内各行业云计算应用中是相当超前的。这两年国家电网公司信息通信部组织开展一体化“国网云”的研发和建设，形成的一批研究成果在全国范围内产生了很大的影响。众所周知，关键核心技术是国之重器，对推动我国经济高质量发展、保障国家安全都具有十分重要的意义。我们高兴地看到，国家电网公司正在国产自主可控云计算核心技术的支撑下，大力推进网络安全和信息化建设。

国家电网公司的王继业同志从云计算的科学研究和大型企业的应用实践的角度出发，撰写了这本《云操作系统架构与关键技术》，它对云计算底层的最核心部分——云操作系统进行了全面细致的阐述，深度剖析了分布式架构的原理和设计方

法，同时紧密结合国家电网公司建设云计算的实践，对云操作系统的应用场景也进行了深刻详尽的探讨，实现了设计与应用的全面覆盖，是云计算领域一本难得的好书。我有幸先睹为快，并乐意推荐给相关学者参考切磋。

中国工程院院士 

前言

自电子计算机诞生以来，操作系统（Operating System，OS）一直是计算机的核心，负责控制和管理计算机的中央处理器（CPU）、内存、输入/输出设备、网络连接以及各种文件的组织存储、操作和保护，各种系统软件和应用软件的管理、安装、运行和进程作业管理，起着计算机“大脑”的作用。之后随着计算机科学与技术的发展，尽管各种个人计算机、大型计算机、计算机网络的规模、数量发生了巨大变化，但操作系统的基本功能、概念并没有发生根本变化。云计算也一样，云操作系统的基本概念、原理与早期的计算机操作系统是一致的。

但实际上随着电子数字计算机的广泛深入普及，操作系统概念的内涵和外延也在不断扩大，并出现了各种不同的形态和产品，比如嵌入式操作系统、移动终端操作系统、大型机操作系统、个人计算机操作系统、网络操作系统以及本书所提到的云计算分布式操作系统。操作系统有的规模很大，管理的资源也很多；有的规模很小，管理的资源也很少；有的时序性要求很高，有的则相对低一些。由于其门槛高，技术难度大，其产品也越来越集中，如桌面操作系统 Windows，手机操作系统 iPhone OS、Android，服务器操作系统 Linux 等，逐步在行业里起到中流砥柱的作用。

云计算是计算机行业发展的新阶段，它通过对信息基础设施（比如服务器、存储、网络和系统软件）的虚拟化和分布式处理，提高资源的利用效率和共享水平，减轻机构使用网络的总体成本。Gartner 和 IDC 均预测，未来全球 90%的企业将不同程度地使用云计算服务或平台，上云已经成为一种标配。云计算的广泛应用，使得计算机的存储和计算等资源重新放到后台（即云端），前端仅仅是一个人机接口界面，这与早期的“大（中、小）型机+终端”的集中式计算存储模式很相似，所不同的是新型终端事实上也具备了很强的信息处理能力，另外分布式处理弥补了单个主机能力不足的问题，可以根据用户需求有效扩展计算能力和存储，从而做到了系统资源的弹性扩展或收缩，系统做到了可用和好用。云计算是新一代的信息基础设施，云操作系统是新一代信息基础设施的核心，它直接管理了云中（可能是一些

计算机，可能是一个数据中心内部的所有资源，也可能是多个数据中心的所有资源）的计算机硬件、存储、网络以及软件资源，规模更大，内容更为复杂，理解更为费力。为了使广大计算机工作者和用户深入了解云操作系统，本书详细梳理了云操作系统的概念、架构、原理、设计方法以及部署方法、应用案例等，希望能对大家有所裨益。

2015 年，国家电网有限公司（以下简称国网公司）发布《信息通信新技术推动电网和公司创新发展行动计划》，正式启动“国网云”建设，2017 年 4 月，“国网云”正式发布。“国网云”始终秉承“开放创新”“自主可控”的建设原则，逐渐推动信息化建设和运行模式的深化转型，提高资源共享水平，促进“互联网+”创新发展。作为主要负责人，我亲历“国网云”建设历程，见证了“国网云”技术团队的快速成长。技术团队汲取着互联网企业宝贵的实践经验，开展了国网云操作系统 SG-COS 自主核心技术研发攻关，并在公司各单位实现了规模部署。

鉴于市场上的书籍普遍定位到某一个具体的产品，专注阐述其功能实现和部署方法，本书编撰的初衷在于，一方面，希望能够将国网公司对云操作系统在企业级生产环境下应用的心得、经验进行总结并对外分享；另一方面，希望能够从方法论的角度提炼云操作系统的设计理念和模型，改变市场上主流书籍“重部署、重产品，轻设计、轻原理”的局面，更好地从根本上提升中国在云计算产业链上的话语权。

衷心希望更多的企业 IT 管理者、决策者、架构师及立志成为架构师的技术人员能从这本书中获取有价值的信息，进而对自身职业发展和所在企业的业务发展有所帮助。对于有一定技术背景，希望对云操作系统架构有一个整体了解的 IT 教学工作者、高等院校的学生，本书也是一部不错的书籍。

特别感谢中国科学院周孝信院士主持本书的审校工作，中国工程院邬贺铨院士、倪光南院士先后对本书编撰给予悉心指导并亲自作序，以更为前瞻、开拓、深刻的视角多次修正和完善本书，亦加深我个人对云计算的认识和理解。同时也十分感谢国网公司信息通信部曾楠、王晋雄同志，国网冀北电力王东升、国网江苏电力张明明同志，国网信通产业集团李富生、孙德栋、张春光、王思宁、李云、孙磊、付兰梅、彭嫚、陈影、李天啸、童骁、贾翠玲等诸位同志协助完成书稿内容验证、统稿和校注等相关工作，正是他们对“国网云”的深刻理解，才促成了本书的成稿。

最后，借此书出版向所有参加和给予帮助“国网云”设计、研发、建设和运行的单位与个人表示由衷的感谢！



2018 年 12 月于北京

目录

绪论	001
----	-----

第一部分 基础篇

第1章 操作系统发展历程	007
--------------	-----

1.1 操作系统的基本类型	007
1.1.1 批处理操作系统	008
1.1.2 分时操作系统	010
1.1.3 实时操作系统	011
1.2 通用操作系统	013
1.3 单机操作系统	015
1.3.1 单用户单任务操作系统	015
1.3.2 单用户多任务操作系统	016
1.4 网络操作系统	016
1.4.1 网络操作系统的特征	017
1.4.2 常用网络操作系统	017
1.5 分布式操作系统	020

第2章 云操作系统概述	022
-------------	-----

2.1 云操作系统概述	022
2.2 云操作系统发展阶段	024
2.2.1 阶段一：分区虚拟化	024
2.2.2 阶段二：虚拟化+资源池管理	026
2.2.3 阶段三：云操作系统	031

2.3 云操作系统主流产品031

2.3.1 亚马逊云操作系统031

2.3.2 阿里云操作系统034

2.3.3 微软云操作系统035

2.3.4 华为云操作系统036

第二部分 设计篇

第3章 云操作系统技术架构041

3.1 一个最简单的云操作系统041

3.2 云操作系统的主要功能043

3.3 云操作系统的逻辑架构045

3.4 典型服务的设计思想048

3.4.1 “管理者-生产者”模型048

3.4.2 关键组件实现模型050

第4章 硬件抽象与封装053

4.1 计算资源053

4.1.1 计算类硬件设备053

4.1.2 计算资源的抽象054

4.1.3 计算资源的封装060

4.2 存储资源063

4.2.1 存储类硬件设备063

4.2.2 存储资源的抽象065

4.2.3 存储资源的封装072

4.3 网络资源076

4.3.1 网络类硬件设备076

4.3.2 网络资源的抽象078

4.3.3 网络资源的封装079

第5章 资源调度087

5.1 计算资源的调度087

5.2 存储资源的调度089

5.2.1 块存储089

5.2.2	文件存储	090
5.2.3	对象存储	091
5.3	网络资源的调度	092
5.3.1	路由算法	092
5.3.2	负载均衡的算法	092
第 6 章	资源编排	094
6.1	通用化的编排方式	094
6.2	虚拟机的编排	094
6.3	容器的编排	096
6.3.1	Swarm	097
6.3.2	Kubernetes	098
6.3.3	Mesos&Marathon	100
第 7 章	管理类基础服务	102
7.1	用户界面	102
7.1.1	概览	102
7.1.2	组件介绍	102
7.1.3	使用示例	103
7.2	用户认证服务	103
7.2.1	概览	103
7.2.2	组件介绍	104
7.2.3	使用示例	105
7.3	数据库服务	107
7.3.1	概览	107
7.3.2	组件介绍	107
7.3.3	使用示例	109
7.4	镜像服务	109
7.4.1	概览	109
7.4.2	组件介绍	110
7.4.3	使用示例	110
7.5	通信服务	110
7.5.1	概览	111
7.5.2	组件介绍	111
7.5.3	使用示例	115

第 8 章 云操作系统的接口 116

8.1 HTTP 116

8.2 用户管理 118

 8.2.1 用户管理 118

 8.2.2 租户管理 118

8.3 权限管理 119

8.4 运行管理 119

 8.4.1 监控日志 119

 8.4.2 计量监控-资源 122

8.5 云主机管理 122

 8.5.1 云主机管理 122

 8.5.2 云主机操作 122

8.6 存储管理 123

 8.6.1 云硬盘管理 123

 8.6.2 云硬盘操作 123

 8.6.3 云硬盘备份 124

 8.6.4 云硬盘转移 124

 8.6.5 云硬盘快照 124

 8.6.6 对象存储目录管理 125

 8.6.7 对象存储对象管理 125

 8.6.8 共享文件系统管理 126

 8.6.9 共享文件系统操作 126

8.7 网络管理 126

 8.7.1 网络管理 126

 8.7.2 网络端口管理 127

 8.7.3 子网管理 127

 8.7.4 路由器管理 127

 8.7.5 浮动 IP 地址管理 128

8.8 镜像管理 128

 8.8.1 镜像管理 128

 8.8.2 镜像数据管理 129

8.9 编排管理 129

 8.9.1 栈资源管理 129

 8.9.2 栈管理 130

8.9.3 栈操作管理	130
8.10 容器管理	130
8.11 主机虚拟化驱动	132
8.12 网络虚拟化驱动	137
8.13 分布式存储驱动	138

第三部分 部署篇

第9章 典型云操作系统的部署	145
9.1 系统部署	145
9.1.1 按规模部署	145
9.1.2 主流的自动化部署工具	146
9.2 部署架构设计	147
9.3 部署资源准备	147
9.4 网络规划与连线	147
9.5 部署准备	149
9.5.1 配置服务器	149
9.5.2 部署服务器	150
9.6 手动部署云操作系统	157
9.6.1 加载物理机	157
9.6.2 配置网络环境	158
9.6.3 配置物理磁盘	161
9.6.4 测试验证网络	162
9.6.5 部署云操作系统	162
9.7 自动部署云操作系统	180
9.7.1 加载物理机	180
9.7.2 配置网络环境	182
9.7.3 配置物理磁盘	184
9.7.4 测试验证网络	185
9.7.5 创建部署环境	186
9.7.6 部署实施云操作系统	187
第10章 云操作系统的高可用	189
10.1 HA 介绍	189

- 10.1.1 冗余和故障切换189
 - 10.1.2 无状态和有状态服务190
 - 10.1.3 主从和主主集群190
- 10.2 基本环境的配置190
 - 10.2.1 硬件要求190
 - 10.2.2 Memcached.....191
- 10.3 配置共享服务191
 - 10.3.1 数据库191
 - 10.3.2 消息队列191
- 10.4 配置控制器192
 - 10.4.1 Pacemaker 架构.....192
 - 10.4.2 Pacemaker 集群.....193
 - 10.4.3 配置 VIP.....194
 - 10.4.4 HAProxy.....194
 - 10.4.5 Memcached.....194
 - 10.4.6 Keystone194
 - 10.4.7 Telemetry194
- 10.5 配置网络服务195
 - 10.5.1 网络的 DHCP Agent195
 - 10.5.2 网络的 L3 Agent196
- 10.6 配置存储197
 - 10.6.1 Glance API.....197
 - 10.6.2 Cinder API198
 - 10.6.3 共享文件系统199
 - 10.6.4 存储后端200
- 10.7 配置计算节点201

第 11 章 大规模集群设计与优化.....203

- 11.1 私有云待优化问题203
- 11.2 资源监控系统205
 - 11.2.1 监控系统简介205
 - 11.2.2 监控系统效率207
- 11.3 通信性能优化211
- 11.4 键函数优化213
- 11.5 模块依赖性能优化213

11.6 任务调配优化	214
11.6.1 队列资源调度算法	215
11.6.2 作业结构设计	216

第四部分 应用篇

第 12 章 云操作系统的应用场景	221
12.1 应用场景综述	221
12.1.1 公有云	221
12.1.2 私有云	224
12.1.3 混合云	227
12.2 典型场景一：异构资源的统一封装	229
12.2.1 场景描述	229
12.2.2 实现流程	229
12.2.3 应用案例	237
12.3 典型场景二：基础资源弹性伸缩	239
12.3.1 场景描述	239
12.3.2 实现流程	240
12.3.3 应用案例	244
12.4 典型场景三：多租户隔离	247
12.4.1 场景描述	247
12.4.2 实现流程	248
12.4.3 应用案例	251
12.5 典型场景四：一键部署	255
12.5.1 场景描述	255
12.5.2 实现流程	257
12.5.3 应用案例	261
12.6 典型场景五：故障自愈	265
12.6.1 场景描述	265
12.6.2 实现流程	267
12.6.3 应用案例	267
12.7 典型场景六：持续集成	274
12.7.1 场景描述	274
12.7.2 实现流程	276

12.7.3 应用案例276

12.8 典型场景七：灰度发布284

12.8.1 场景描述284

12.8.2 实现流程285

12.8.3 应用案例286

第 13 章 一个典型的云操作系统应用291

13.1 国网云操作系统 SG-COS 综述291

13.1.1 系统架构291

13.1.2 部署情况294

13.2 “国网云”大数据平台概述295

13.3 “国网云”云操作系统与大数据平台的融合297

13.3.1 融合方案概述297

13.3.2 物理机纳管方案298

13.3.3 半虚半实融合方案299

13.3.4 全虚拟化融合方案303

13.4 “国网云”企业级应用迁移上云的典型方式304

第 14 章 总结与展望306

附录一 术语及定义309

附录二 国内外主流云操作系统简介311

亚马逊 AWS311

微软 Cloud OS317

阿里飞天云平台320

华为 FusionSphere322

国网云操作系统324

参考文献328