



装备科技译著出版基金



大数据科技译丛

Cloud Computing for Data-Intensive Applications

面向数据密集型应用的云计算



李晓林 (Xiaolin Li) 邱小红 (Judy Qiu) 编著

郭建胜 郑明发 毛声 王晨 译



国防工业出版社

National Defense Industry Press



Springer



装备科技译著出版基金

大数据科技译丛

面向数据密集型应用的 云计算

Cloud Computing for
Data-Intensive Applications

李晓林(Xiaolin Li) 邱小红(Judy Qiu) 编著

郭建胜 郑明发 毛声 王晨 译

国防工业出版社

·北京·

著作权合同登记 图字:军-2016-130 号

图书在版编目(CIP)数据

面向数据密集型应用的云计算/李晓林,邱小红编著;
郭建胜等译. —北京:国防工业出版社,2019.4

(大数据科技译丛)

书名原文:Cloud Computing for Data-Intensive Applications

ISBN 978-7-118-11725-7

I. ①面… II. ①李… ②邱… ③郭… III. ①云
计算—研究 IV. ①TP393.027

中国版本图书馆CIP数据核字(2018)第296937号

Translation from the English language edition:
Cloud Computing for Data-Intensive Applications
edited by Xiaolin Li and Judy Qiu

Copyright © Springer Science+Business Media New York 2014

This Springer imprint is published by Springer Nature

The registered company is Springer Science+Business Media LLC

All Rights Reserved

本书简体中文由 Springer 独家授权国防工业出版社出版发行,版权所有,侵权必究。

※

国防工业出版社出版发行

(北京市海淀区紫竹院南路23号 邮政编码100048)

三河市腾飞印务有限公司印刷

新华书店经售

*

开本 710×1000 1/16 印张 25 字数 481 千字

2019年4月第1版第1次印刷 印数1—2000册 定价125.00元

(本书如有印装错误,我社负责调换)

国防书店:(010)88540777

发行邮购:(010)88540776

发行传真:(010)88540755

发行业务:(010)88540717

译者序

近十年大数据与云计算领域迎来了空前发展,云计算科学数据分析吸引了大量研究人员的关注。大数据的发展已不仅限于互联网领域,物联网大数据也通过云计算、机器学习或其他实验手段而为人所知。为充分利用这些数据造福社会,人们创建了海量云数据中心,并通过云应用提供在线服务。在这个过程中,大规模数据分析应用所需的主流软件工具在云中应运而生,人们所面临的挑战由物联网大数据的处理转向了云计算中密集数据的处理。

数据密集型应用不仅包括外源数据,还包括应用程序随着时间的推移累积起来的大量数据,如日志、性能指标数据、计算资源分配情况记录、磁盘数据存储信息等。对这些数据及资源进行处理、整合与调度,需要借助内存以外更庞大的辅助存储设备或第三方大数据系统。基于此,本书介绍了数据密集型应用的相关概念,支持数据密集型应用云环境配置的相关技术,数据密集型应用自身的计算效率、资源分配、数据存储以及将这类应用与其他模型整合的工作方案。本书涵盖了云计算和数据密集型应用课题研究的多个方面,算法理论与操作实践并举,不仅可以作为计算机相关专业学生的课外知识读本,帮助他们理解云计算与密集型大数据的相关领域,发展实践能力,也可作为相关研究人员的参考文献,为他们处理问题提供可借鉴的解决思路及方案。此外,本书总结了数据密集型应用及其云环境领域相关的方法策略、算法实验、实践经验等,涉及的案例类型丰富,不仅探讨了支持数据密集型应用的云环境的多个层次,介绍了数据密集型应用及云环境性能优化的方法,还展示了不同云平台上的相关实验结果,为数据密集型应用的扩展运用提供实践基础。本书从基础架构问题出发,延伸到复杂科学应用中数据分析的挑战,为国内了解数据密集型应用及云计算的前沿发展技术,追踪最新动态提供了有价值的参考。

参加本书翻译工作的是郭建胜、郑明发、毛声、王晨。毛声负责完成译者序、前言以及第1~8章的翻译,王晨负责完成第9~17章的翻译,由郭建胜和郑明发负

责校对与统稿工作。

最后感谢国防工业出版社的编辑们在翻译过程中给予的专业指导和有力帮助,也感谢其他同事对我们工作的支持。由于水平有限,翻译过程中难免出现一些错误和疏漏,欢迎读者批评指正。

译者

2019年1月8日

前言

本书是第一本探究科学数据分析论中日益重要的云计算的论文集。几乎所有的科学领域都是由数据驱动的,所以这个话题的重要性可见一斑。数据来源范围广泛,包括大量物理实验、读取 DNA 的仪器设备、环境监测传感器、数字化图书馆、媒体流、个人健康监测器等。“物联网”正在经历爆炸性的发展,它深刻地改变着我们对周围世界的理解方式。针对大量数据,我们可以使用机器学习和分析论,发现复杂系统的一些属性;而这些属性在之前对于我们并不可见。这是真正的科学的第四范式的时代,其中数据分析结合了实验、理论和数值模拟,共同构成科学发现的支柱。例如,通过对大量人口进行全基因组关联研究,现在我们可以开始了解许多疾病的病因。城市环境中的传感器也正开始帮助我们了解如何改进城市现状,更好地规划未来。基于大数据的机器学习也正改变着我们的私人设备。使用大规模深度神经网络,对数百万小时的视频和其他数据进行分析,可以满足无人驾驶汽车所需的计算机视觉要求。基于相同的无监督机器学习技术,现在我们可以做到自然语言的语音实时翻译。

许多用于做大规模数据分析的应用所需的主流软件工具都是在云中诞生。人们已经创建了海量云数据中心为在线服务提供计算基础,现在在线服务已经成了数十亿人日常生活的一部分,主要包括互联网网络搜索、电子邮件、电子商务、社交网络、地理定位、地图服务、图片共享、自动化的自然语言翻译、文件准备和协作、媒体分布、会议和在线游戏等。这些应用程序均会累积大量数据,而要优化应用的性能,则要求对这些数据进行分析。例如,一个 Web 搜索查询要返回最佳结果集,需要通过更多的简单索引进行查找。先进的机器学习技术都会事先选出一些最好的链接发送给用户。将这些想法应用于解决所有领域的科学挑战并不会花太长时间。

面向可扩展的数据分析的云软件堆栈发展迅速,因此云架构模型也得以发展。伴随着涌现的新技术,早期的云数据中心正在重建,以更好地支持海量数据分析。数据中心网络正开始从超级计算机系统中借鉴思路,扩展带宽;而且现在由软件定义的网络也成为标准设计的一部分。有些云供应商正开始考虑服务器组合中的

GPUs。私有云和公共云正开始合力工作,云和超级计算机也正以多种创意形式得到运用。

这是涵盖云计算和数据密集型应用课题研究各方面内容的第一本书,涉及的内容从基础架构问题,延伸到复杂科学应用的数据分析中的挑战。本书对现代计算机领域的每一个学生而言都是“必读”的。我们很高兴李晓林教授和邱小红教授承担了这一项目。

Tony Hey

Dennis Gannon

微软研究中心,雷德蒙德,WA,USA

目录

第一部分 系统与应用

第 1 章 公共云上 LIGO 物理应用中的可扩展部署: workflow 引擎与资源配置技术	003
1 引言	004
2 相关工作	005
3 体系架构及设计	007
3.1 负载均衡	010
4 LIGO 数据分析和引力波搜索	012
4.1 应用需求	016
5 性能评估	017
5.1 平台可扩展性	018
5.2 工作者的动态配置	020
6 本章小结	020
参考文献	021
第 2 章 面向大数据的 FutureGrid 试验台	024
1 引言	024
2 FutureGrid 简介	025
2.1 硬件概述	025
3 面向大数据的服务和工具	028
3.1 测试平台即服务 (TestbedaaS)	029
3.2 传统的高性能计算即服务 (HPCaaS)	030
3.3 网格即服务 (GridaaS)	030
3.4 基础设施即服务 (IaaS)	030
3.5 云平台即服务 (PaaS)	032
4 FutureGrid 的运用	034

5	系统管理	040
5.1	系统开发团队的整合	040
5.2	DevOps	040
5.3	教育支持	042
6	Cloudmesh	043
6.1	功能	043
6.2	架构	044
6.3	云转移	047
6.4	图形用户界面	047
6.5	Shell 命令和命令行接口	049
7	本章小结	050
	参考文献	051
第3章 支持数据密集型应用的云网络		054
1	引言	054
2	构建云网络的模块和技术	055
2.1	数据处理中心网络	055
2.2	存储区域网	055
2.3	网络协议栈	056
2.4	局域网分区	056
2.5	VPN	057
2.6	虚拟网络和覆盖网	057
2.7	高性能主干网	058
2.8	软件界定的网络	058
3	云的内部网络	060
3.1	商业 IaaS 云	060
3.2	开源 IaaS 云	061
3.3	SDN 的网络虚拟化	063
4	内部云网络	064
4.1	CohesiveFT 的 VNS-Cubed 模型	065
4.2	Pertino 云网络引擎	065
4.3	ViNe 网络覆盖基础设施	066
5	案例分析	068
5.1	优化云内和云间的数据传输	068
5.2	虚拟机迁移	068

5.3 云上的科学应用程序.....	069
6 本章小结	070
参考文献.....	071
第4章 IaaS 云基准测试:方法、挑战和经验	074
1 引言	074
2 计算机系统基准简介	076
2.1 为什么要基准测试.....	077
2.2 基准测试的要素.....	077
3 IaaS 云基准测试的通用架构	078
3.1 概述.....	078
3.2 显著的设计特征.....	079
4 IaaS 云基准测试中的开放性挑战	080
4.1 方法论.....	080
4.2 系统属性.....	081
4.3 工作负载.....	082
4.4 度量.....	083
5 IaaS 云基准测试的经验	083
5.1 方法:SPEC 云工作组.....	084
5.2 SkyMark:IaaS 云基准测试的框架	084
5.3 IaaS 云性能的真实评估.....	085
5.4 工作负载的统计模型.....	086
5.5 开放数据:几个有用的档案库	087
6 本章小结	087
参考文献.....	088
第5章 数据密集型应用的 GPU 加速云计算	092
1 引言	092
2 背景和相关工作	093
2.1 云计算.....	093
2.2 GPU	094
2.3 MapReduce	094
2.4 通用图形处理.....	096
3 基于 GPU 集群的 MapReduce	097
3.1 Mars 概述	097

3.2	MarsHadoop	099
3.3	Mars-MR-MPI	100
3.4	实验	101
4	GPU 和云上的图像处理	105
4.1	GPU 上的并行图像处理	105
4.2	云上的并行图像处理	108
5	本章小结	110
	参考文献	112
第 6 章	面向数据密集型科学应用的自适应工作负载分区与分配	117
1	引言	117
1.1	工作成果总结	118
1.2	组织结构	119
2	相关工作	119
3	问题描述	120
4	Apala	121
4.1	统一工作负载	121
4.2	分解统一的工作负载	122
4.3	分配统一的工作负载	124
5	评估	125
5.1	设置	125
5.2	分解模式	127
5.3	负载平衡的有效性	128
5.4	可移植性	129
5.5	分区的代价	131
6	本章小结	132
	参考文献	133
第 7 章	DRAW:面向具有局部兴趣的数据密集型应用的一种基于数据分 组感知的新型数据配置方案	135
1	引言	136
2	动机	137
3	基于数据分组感知的数据配置	138
3.1	历史数据访问图 (HDAG)	138
3.2	数据分组矩阵	139

3.3	最优数据配置算法	141
3.4	其他注意事项	142
4	分析	143
4.1	“随机=最优”的可能性	143
4.2	数据分布的最优度	145
4.3	随机分布的最优度	147
4.4	多副本机架	148
5	方法	148
5.1	测试平台	148
5.2	应用程序	149
5.3	实现	150
6	实验结果与分析	151
6.1	实验结果	151
6.2	MapReduce 程序的性能改进	152
6.3	敏感性研究:副本数目	154
6.4	DRAW 的代价	155
7	相关工作	156
8	本章小结	156
	参考文献	157

第二部分 资源管理

第 8 章	应用自适应组合拍卖的任务资源有效匹配	161
1	引言	161
1.1	贡献总结	163
1.2	组织结构	163
2	相关工作	163
3	问题定义	164
3.1	前言	165
3.2	任务资源匹配:从博弈论的角度解析	167
3.3	任务—资源匹配:示例场景	167
3.4	动态迭代拍卖简介	167
4	用于高效任务—资源匹配策略的基于拍卖的设计	169
4.1	同质性案例	169
4.2	K 类异质性案例	174

4.3 通用异质性案例	175
5 结果评估	175
5.1 任务调度的性能	175
5.2 激励兼容性的验证	177
5.3 统一价格拍卖的性能比较	178
6 本章小结	179
参考文献	179
第9章 具备自主能力的高级网络基础设施的整合	182
1 引言	182
2 发展现状	184
2.1 联合计算网格	184
2.2 云计算的整合	185
2.3 互操作性标准化活动	185
3 聚合分布式资源的联合模型	186
3.1 要求	186
3.2 联合架构	187
3.3 Comet 云	189
3.4 自主管理	189
3.5 启用自主能力	191
4 应用场景	191
4.1 CDS&E 应用	192
4.2 企业业务数据分析	196
5 经验教训	199
6 本章小结	200
参考文献	201

第三部分 编程模型

第10章 科学 workflow 管理系统从网络到云的迁移	209
1 引言	210
2 挑战及可用的解决方案	211
2.1 传统科学 workflow 的挑战	211
2.2 将 workflow 应用程序迁移到云	213
2.3 将 workflow 管理迁移到云中	214

3	Swift 和 OpenNebula 的集成	215
3.1	服务框架	215
3.2	集成架构和实现方法	216
4	性能评估	221
4.1	MODIS 图像处理工作流程	221
4.2	实验配置	221
4.3	实验结果	222
5	相关工作	225
6	本章小结	227
	参考文献	228
第 11 章 PaaS 云上执行的风暴潮集成		232
1	引言	232
2	架构/系统概述	234
2.1	Windows Azure	234
2.2	Sigiri 中间件	235
2.3	MapReduce:Twister4Azure	236
3	SLOSH 集合	236
4	联合部署和运行	239
4.1	Sigiri 中间件	239
4.2	MapReduce 运行过程	240
5	输出感知任务的部署	242
5.1	负载分区	242
5.2	Sigiri 中的工作-窃取算法	242
6	运行时间的建模	243
7	绩效评估	243
7.1	不同的工作负载	244
7.2	不同的并行度	245
7.3	模型评估	246
8	相关工作	247
9	本章小结	248
	参考文献	248
第 12 章 MapReduce 中的交叉相位优化		250
1	引言	251

1.1	MapReduce 在高度分布式环境中的性能	253
1.2	交叉相位优化	255
2	Oracle:模型驱动优化	256
2.1	模型及其优化	256
2.2	观点	257
3	映射—感知推送	258
3.1	并行推送和映射以隐藏延迟	259
3.2	并行推送和映射以改进调度方案	259
3.3	映射-感知推送调度	259
3.4	在 Hadoop 中的实现	260
3.5	实验结果	261
4	置乱感知映射	263
4.1	置乱感知映射调度	264
4.2	在 Hadoop 中的实现	265
4.3	实验结果	265
5	综合比较	266
5.1	Amazon EC2 环境下的测试	266
5.2	PlanetLab 环境下的测试	268
6	相关工作	268
7	本章小结	270
	参考文献	271
第 13 章	面向大规模迭代计算的异步计算模型	273
1	异步迭代	273
2	基于增量累积的迭代计算	276
2.1	DAIC 简介	276
2.2	异步 DAIC	278
2.3	收敛性	278
2.4	有效性	279
2.5	异步 DAIC 中的调度	280
3	编写异步 DAIC 算法	280
3.1	指导方案	281
3.2	案例	281
4	Maiter:一个支持异步 DAIC 的框架	283
4.1	系统设计	284

4.2 Maiter API	287
4.3 Maiter 程序示例	288
5 性能	290
6 相关工作	293
7 本章小结	293
参考文献	294

第四部分 云计算

第 14 章 Azure 云上的大数据存储与处理:大规模实验与经验教训	299
1 引言	299
2 面向数据密集型应用的云存储挑战	300
3 案例研究:Azure 云数据存储与管理	301
3.1 存储非结构化数据:Azure Blobs	302
3.2 存储 NoSQL 结构化数据:Azure 表	303
3.3 并发数据处理的同步过程:Azure 队列	303
3.4 其他	303
4 进一步优化:在产生大量并发任务时处理存储延迟	304
4.1 聚合虚拟磁盘进行高效通信存储	305
4.2 使用专用的计算节点进行可扩展的数据管理	306
4.3 面向 Azure 云中归约密集型处理过程的低延迟存储方法	309
5 在云端执行大规模大数据实验	311
5.1 脑-基因相关性研究	311
5.2 计算和存储问题	311
5.3 实验设置	312
5.4 结果	312
5.5 云中的生物大数据处理:经验教训	315
6 本章小结	318
参考文献	319
第 15 章 FRIEDA 云环境中的存储与数据生命周期管理	320
1 引言	320
2 相关工作	321
3 研究背景	323
3.1 存储模型	323

3.2	云环境中的 I/O 性能	324
3.3	资源模型	324
3.4	应用执行模式	325
3.5	科学应用数据类型	325
4	FRIEDA 生命周期	326
4.1	数据分区	327
4.2	数据管理	328
5	存储规划	328
6	存储配置与准备	332
7	数据配置	333
8	FRIEDA 执行过程	333
8.1	通信协议	334
8.2	执行阶段	336
8.3	数据分类	336
9	本章小结	337
	参考文献	337
第 16 章	作为云服务的文件传输管理	340
1	引言	340
2	数据不溶性的问题	341
2.1	解决方案	342
3	文件传输管理	343
4	StorkCloud	346
4.1	StorkCloud 调度程序	346
4.2	目录列表服务	347
4.3	Web API 和小型客户端 GUI	347
4.4	传输模块接口	348
4.5	优化模块	348
5	传输层吞吐量优化	349
5.1	优化技术	349
5.2	动态优化	351
5.3	MFT 系统中的示例	351
6	调度优化和预约	352
6.1	协同调度算法	352
6.2	历史性能数据的估计	353