

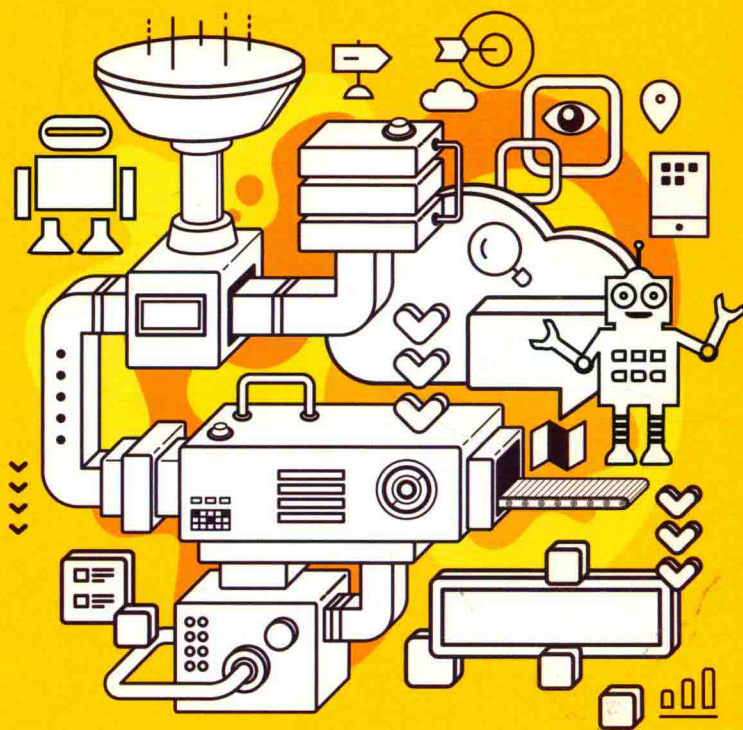
智能运维即未来 AIOps时代开启
腾讯|滴滴|美团|微博|清华等产学研界翘楚瞩目力荐

Broadview[®]
www.broadview.com.cn

智能运维

从0搭建大规模分布式AIOps系统

彭冬 朱伟 刘俊 等著



 中国工信出版集团

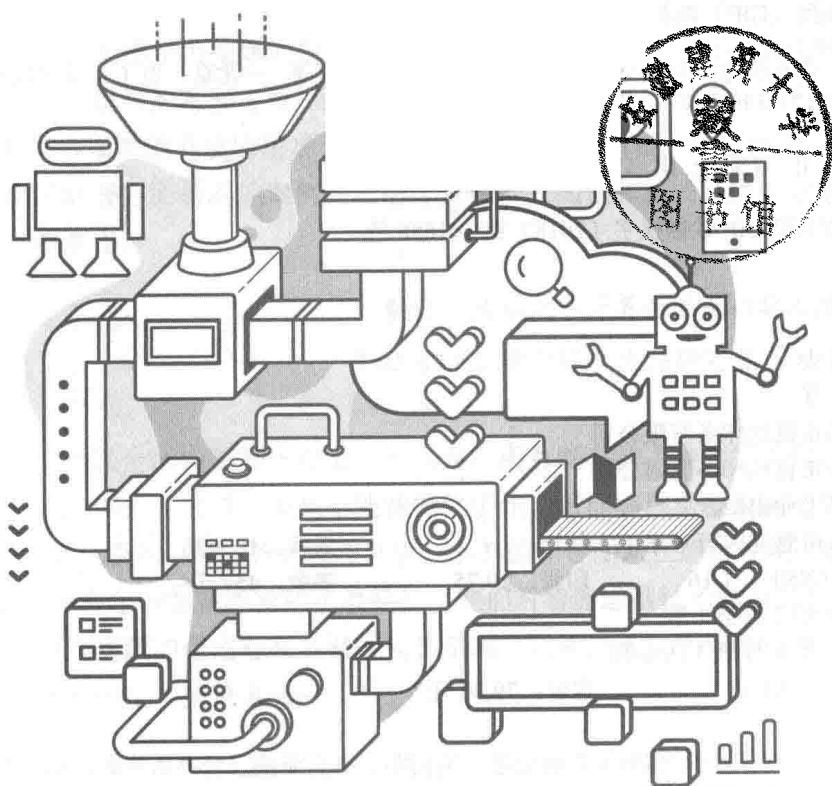
 电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn

非外借

智能运维

从0搭建大规模分布式AIOps系统

彭冬 朱伟 刘俊 王莉 陆松林 车亚强 著



电子工业出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

本书结合大企业的智能运维实践,全面完整地介绍智能运维的技术体系,让读者更加了解运维技术的现状和发展。同时,帮助运维工程师在一定程度上了解机器学习的常见算法模型,以及如何将它们应用到运维工作中。

全书共分4篇。第1篇运维发展史,重点阐述当前运维的发展现状及面临的技术挑战;第2篇智能运维基础设施,重点讲述大数据场景下的数据存储、大数据处理和分析的方法与经验,以及海量数据多维度多指标的处理分析技术;第3篇智能运维技术详解,重点关注在新时期大数据时代下智能化的运维技术,包括数据聚合与关联、数据异常点检测、故障诊断和分析、趋势预测算法;第4篇技术案例详解,为大家梳理了通过开源框架ELK快速构建智能监控系统的整体方案,还将分享微博平台和微博广告两个不同业务场景下智能监控系统的技术实践。

本书适合运维、开发、架构、DevOps工程师及广大互联网技术爱好者研读和借鉴。

未经许可,不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有,侵权必究。

图书在版编目(CIP)数据

智能运维:从0搭建大规模分布式AIOps系统 / 彭冬等著. —北京:电子工业出版社,2018.7
ISBN 978-7-121-34663-7

I. ①智… II. ①彭… III. ①软件维护 IV. ①TP311.53

中国版本图书馆CIP数据核字(2018)第141688号

策划编辑:张春雨

责任编辑:葛娜

印刷:三河市良远印务有限公司

装订:三河市良远印务有限公司

出版发行:电子工业出版社

北京市海淀区万寿路173信箱

邮编:100036

开本:787×980 1/16 印张:20.75

字数:457千字

版次:2018年7月第1版

印次:2018年8月第3次印刷

印数:5002~6501册 定价:79.00元

凡所购买电子工业出版社图书有缺损问题,请向购买书店调换。若书店售缺,请与本社发行部联系,联系及邮购电话:(010)88254888,88258888。

质量投诉请发邮件至 zlts@phei.com.cn, 盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式:010-51260888-819 faq@phei.com.cn。

推荐语

本书和国内外第一个《企业级 AIOps 实践建议白皮书》在很多内容上不谋而合，也是对 AIOps 白皮书的深度细化和技术补充。本书基于新浪微博平台技术部门及广告技术部门的生产实践，有理有据，有说服力。文字简练、表达清晰。相关实践可落地，让有志于实践 AIOps 的企业及运维同仁“开箱即用”。

——**萧田国** 高效运维社区发起人、AIOps 标准及白皮书发起人

伴随着大数据和机器学习技术的发展，智能运维（AIOps）也成为运维领域极大的热点。本书不仅介绍了智能运维的发展过程，还针对大数据处理和智能运维相关的底层技术进行了详细的分析，并且结合微博的具体应用场景提供了在实时监控、报警、异常检测、故障根源分析、趋势预测、数据关联等方面的实战案例，浓缩了大量开发知识和实践经验。这是一本非常有参考价值的智能运维著作。

——**裴丹** 清华大学计算机系长聘副教授、青年千人、
美国 AT&T 研究院前主任研究员、智能运维算法专家

当下，世界已经在全面向智能化迈进。可以说，所有既有的“传统”工作，都可以用智能化思路来改造。从反向的角度讲，如果不能智能化工作，将面临逐渐被淘汰的局面。当然，作为支撑整个 IT 世界地基的运维工作，也在其中。本书的作者团队在运维体系、系统架构、大数据及 AI 等方面具有多年的非常丰富的实战经验，他们对智能运维技术体系进行了全面的梳理，将从思路到工具再到实践的全过程在本书中完整呈现。做面向未来的智能运维，做不被机器淘汰的运维人。在此特向读者推荐本书。

——**王鹏云** 多盟联合创始人，蓝色光标技术创新孵化中心总经理，
曾任 139 移动互联研发总监、魔时网 CTO

伴随着云计算和 IoT 技术在各行各业的普及与发展，企业的运维规模不断扩大，我们可以十分确定地预见到，从自动化运维演进到智能化运维，是运维技术发展的必然趋势。本书是从运维大数据技术到 AI 运维技术的书籍，详细介绍了实施智能运维依赖的基础设施和架构技术，既包括通用的智能运维技术，又通过实战案例引出具体技术在运维领域的应用与效果，兼具参考性与实操性，是大家升级运维能力过程中不可错过的一本好书。如果你对运维领域感兴趣，本书将是你的不二之选。

——**梁定安** 腾讯运维技术总监、专家工程师

当大数据、人工智能等技术赋能传统行业、驱动变革时，运维+AI 则把 IT 行业中原来完全依靠人工经验、重复枯燥的工作，推向了自动化的智能运维，引领运维的下一步发展方向。智能运维极大减轻了运维工程师的负担，但也对他们提出了新的技术能力要求。本书主要从发展历史、技术体系以及实际应用等方面梳理了智能运维体系，使读者能够理解与掌握智能运维的知识与方法。本书可以作为运维工程师提升运维水平的重要参考，也可以作为运维工程师职业发展的参考。

——**钟华** 美团打车技术研发部负责人

智能运维已经成为运维领域公认的未来，所有人都想知道自己怎么做才能赶上智能运维的浪潮。本书介绍了作者在异常检测、根因分析、时序预测等智能运维领域的实践经验，更重要的是，书中整理和讲解了智能运维的两大基石：大数据和机器学习。相信读者阅读此书必有所获。

——**饶琛琳** 日志易产品总监，前新浪微博系统架构师

终于看到彭冬写的这本书！从运维平台的大数据处理到架构设计原理，再到 AIOps 的相关模型和算法，并将智能运维工程架构与算法实践相结合，本书非常具有参考价值。

——**陈晓峰** 火币集团副总裁

随着大数据时代的到来，在海量数据场景下，AIOps 开始成为下一代运维技术发展的热点。本书系统地介绍了大数据采集、存储、处理、计算以及策略应用的各个环节，并以微博监控为实际案例向读者展示了监控平台建设的实践经验。相信有志于 AIOps 方向的读者，读过此书后都能有自己的体会和收获。

——**陆沛** 滴滴打车 SRE 团队负责人、技术专家

推荐序 1：运维的时代化变迁

互联网刚兴起的时候，运维还只是一个简单的服务安装管理及监控工作，没人会想到人类在互联网上建立了如此庞大的业务生态。从衣食住行到教育金融，服务器的规模在急剧膨胀，从简单的人力可管控，逐渐进化到依赖自动化体系来管理，但是另一方面，仅依赖工具已经不能很好地解决运维场景的需求。

根据正式的定义，智能运维通常用 AIOps (Algorithmic IT Operations) 这个专有名词来表达，是指利用大数据分析、机器学习等人工智能技术来自动化管理运维事务。早期的大规模运维以服务管理及监控为目标，自动化运维工具可以满足绝大部分需求，大数据采集及分析主要应用在服务监控及分析方面。但随着服务规模的膨胀，通过人工来管理大规模服务已经力不从心，也很难达到服务可用性的要求。2017 年，Gartner 也发布了 AIOps 的市场定义，并预测到 2019 年，全球 25% 的企业将用 AIOps 系统替代传统运维管理系统。在 2017 年这个比例低于 5%，而到 2022 年将超过 40%。

微博作为国内典型的互联网服务之一，在智能运维的实践方面获得了很多经验。比如每年的春晚，有大量用户在微博上进行祝福与互动，微博的 Feed 项目、广告系统、搜索等服务的自动化扩缩容，通过数据标准化算法分析出 QPS 和慢速比，并根据实时压测反馈的数据生成水位线，结合两个指标和水位线的波动情况进行自动扩容和缩容。

在日常的运维工作中，智能运维也在多方面发挥重要作用，比如告警收敛，利用智能算法过滤，收敛大量无效、重复的告警信息，通过聚类算法将大量、多维度的告警聚合为少量事件，通过告警分类算法提高准确率和减少误报。另外，智能运维还被大量应用于故障定位及服务自动修复方面，基于服务运行的日志及告警数据，实现非人工干预的自动化处理，比如自动摘除、重启等操作。

AIOps 已经成为运维领域的发展趋势，但是目前可供参考的书籍较少，且大多偏理论及小规模服务，而针对真正大规模线上实操的书籍非常匮乏。微博技术团队的彭冬、刘俊、朱伟等同事，长期从事微博运维数据相关技术研发，对于如何将 AI 理论和技术应用于大规模服务管理

方面颇有心得。在本书中，系统地介绍了大数据运维基础理论知识，如数据采集、分布式消息队列、实时计算框架、时序数据库等，并结合大数据人工智能技术如 TensorFlow 机器学习框架、趋势预测算法等，介绍了微博平台、微博广告在 AIOps 上的具体实操。相信用心阅读的读者，可以从中深入了解到他们在这一领域的领悟。

目前 AIOps 在业界也刚刚开始，微博技术团队也乐于和广大开发人员分享微博在 AIOps 实践中的心得，欢迎大家关注@微博平台架构和@微博技术学院，了解后续相关的公开技术活动。

杨卫华 (Tim Yang)

微博研发副总经理

推荐序 2：运维的新视角

最近 20 年，互联网特别是移动互联网的普及让用户的获取突破了地理的限制，各个领域都出现了亿级乃至十亿级海量用户规模的高科技公司，如搜索领域的 Google、百度，社交领域的 Facebook、腾讯，出行领域的 Uber、滴滴，电商领域的 Amazon、阿里巴巴，以及中国最大的自媒体社交平台微博等。

服务海量用户需要超大规模的在线分布式系统和大数据处理系统。同时，在竞争日益激烈的市场环境下，高科技公司产品和技术快速迭代能力也至关重要。

在国外，运维团队是 infrastructure 很重要的组成部分，比如 Google 的 SRE (Site Reliability Engineer) 团队，就是从运维的角度为整个公司产品可用性服务的。在国内，近几年创业氛围非常好，2014 年出现了大量的 O2O 创业公司，2015 年直播和短视频企业大量崛起，2016 年共享单车等共享经济方向的创业公司非常红火，2017 年区块链几乎占领了大部分创业板块。对于这些创业企业而言，产品的快速迭代、极致的用户体验都是至关重要的，而在这个过程中运维团队就非常重要。

对于微博这样的成熟型企业，运维体系的完备性和先进性是必不可少的。从运维平台化到智能化，能让产品迭代更加高效，同时保障系统的稳定。微博作为拥有全球超过 4 亿月活跃用户的平台，其自身具有非常显著的特点：媒体属性和社交特性。微博传承了新浪的媒体属性，在媒体和新闻传播上有着非常重要的价值，而社交特性让这种信息的传播产生爆炸式的效应，可以看到，热门事件和重要的新闻资讯几乎第一时间都是在微博上进行传播的。同时，社交属性也体现在明星和大 V 效应上，一些行业的意见领袖拥有百万甚至千万级别的粉丝量，每天在微博上进行转发、评论等互动的行为非常频繁。正由于微博的这些特点，微博广告作为寄生于微博平台上的商业产品，也继承了相同的属性，同样也拥有相当复杂的业务形态。正是在这种复杂的业务场景下，大规模的传播效应、海量的数据、极其复杂的社交关系、多变的热门事件对整个广告系统的稳定性是一个极大的考验。那么如何在保证产品技术快速迭代能力的同时，保证系统的高可靠性和高可用性呢？彭冬领导的微博广告平台运维团队给出的答案是 AIOps。

通过 Big Data、AI/ML 等技术，AIOps 能更快更好地收集和分析超大规模系统的运行状态，并能大幅提升运维系统的自动化和智能化水平，从而提升运维工作的整体效率。

彭冬领导的团队搭建了一套这样的 AIOps 系统，并取得了明显的效果。这套智能运维系统在微博广告平台产品技术快速迭代的同时，提升了系统的可靠性和可用性，有力地支撑了微博广告收入的快速增长，智能运维技术在微博复杂的场景下得到了很好的实践。

很高兴彭冬及其团队成员能抽出时间编写这本书，将他们在智能运维上的架构设计思路和实践经验分享给更多的互联网界朋友。这本书是他们多年积累的工作经验的总结，书中有大量案例和前瞻性研究，无论是工程架构还是模型算法，对读者都具有借鉴价值，也希望这本书能成为一座桥梁，促进更多的技术交流，从而让大家共同成长。

张志强

微博广告业务部总经理

推荐序 3：大数据时代的运维

在互联网时代，尤其是社交网络、电子商务与移动通信把人类社会带入一个 PB 级别以上单位的结构与非结构信息的大数据时代。数据量的爆发性增长，使企业 IT 架构不断扩展，服务器、存储设备的数量越来越多，网络也变得更加复杂。而大数据的 4V 特征，数据量大（Volume）、类型繁多（Variety）、价值密度低（Value）、时效高（Velocity）也使得传统的技术架构和路线，已经难以高效地处理如此海量的数据。可以说，大数据时代对企业的数据驾驭能力提出了新的挑战。尤其是大数据平台往往支撑着公司的搜索、推荐、广告等核心业务，为了保障良好的用户体验和业务效果，运维工作显得十分艰巨。相比于传统的运维方式，大数据时代的运维面临着集群规模更大、业务组件更多、监控可视化与智能化更为复杂等诸多难题。

我们知道，在互联网初期，大部分应用程序跑在少量的服务器上，网络带宽很小，存储量也很小，这个时候的运维更多的是解决类似于组网、操作系统等机房建设问题，应用的上线部署可以由开发工程师来完成，运维的工作职责没有那么明显。随后互联网进入高速发展期，数据规模从 GB 到 TB 再到 PB 级别，在存储量上超过千倍增长，在计算规模上可能也远远超过千倍增长，传统的通过单节点来存储和计算超过 PB 级别的数据已经比较困难，分布式集群的方式已经成为标准的解决方案。分布式系统在存储上解决了大规模数据单机无法承载的问题，同时在计算上解决了单机 CPU 或者内存等资源无法完全满足的问题，但是同时也带来了很多运维难题，诸如统一上线部署、大规模机器管理、降级、容灾、数据同步等。从数据规模到机器规模的扩大，传统的运维方式和方法已经不能满足产品快速迭代的要求，智能运维在这样的场景下应运而生。

智能运维是建立在运维基础上，通过一定策略和算法来进行智能化诊断决策，以更快、更准确、更高效地完成运维工作的技术体系。要实现智能运维的目标，需要有平台支撑，这也是 DevOps 很火的原因，很多运维工程师都掌握了开发工具和平台的本领，因此建立了高效的自动化运维平台。所以说智能运维是运维发展的高级阶段，也是互联网时代发展到一定阶段的产物。

智能运维的基础是建立在大规模数据分析和计算之上，当数据量很小时，我们甚至可以人

工判断和决策，一旦数据达到一定规模，大数据涉及的所有技术就都会成为智能运维所依赖的技术。一方面，可以说智能运维是一种新型技术，因为它从另一个视角去看待运维，对传统运维进行了创新和升华；另一方面，也可以说智能运维是一种经典技术，它是一系列成熟技术的结合体，它融入了运维技术、大数据、传统机器学习技术、机器学习、深度学习等方方面面的技术。

那么在大数据时代应该如何做好运维？我觉得有三个方面。

一是基础设施平台化，大数据的 4V 特性，相比于传统的系统运维，数据的处理框架变得更为多样化和复杂化，这要求我们必须夯实基础设施才能事半功倍。比如多源异构海量数据的分布式存储、离线批处理、高性能索引、大规模流数据处理，以及可视化监控与报警平台等。

二是集群管理自动化，降低运维复杂度。自动化能够提升稳定性，固化的操作交给机器去做，可以降低人为操作失误，提高线上的稳定性；自动化还能极大地提高效率，将运维人员从日常烦琐的操作中解放出来，把更多的时间投入到运维平台迭代优化上，从而更好地为业务运营服务。

三是运维决策智能化，充分利用大数据分析技术提升预测、发现和自动检测的能力，预测分配资源，动态伸缩集群，实现智能预警，自动修复，最大化利用资源，减少开销。

本书作者是微博广告平台资深架构师，从 0 到 1 架构了微博广告大数据运维体系，历经纯人工操作、平台化、自动化、智能化多个发展阶段，可以说踩过许多坑，也总结出一套行之有效的方法论。微博广告五年来一直高速发展而系统始终保持高效稳定，这跟大数据运维体系的成功是分不开的。本书对于即将投入大数据运维行业的工程人员来说是一个非常好的参考，同时对于互联网企业的中高级从业人员也很有借鉴意义。

李东升

微博广告业务部总监

推荐序 4：智能运维也能“开箱即用”

可能鲜为人知的是，智能运维的概念最早并非 Gartner 提出的。

2015 年，国内运维社区热火朝天，也出现了四大运维“门派”。当时王津银同学提出精益运维，智锦同学提出白盒运维，本人萧田国提出高效运维，腾讯游戏的刘栖铜及同仁提出了智能运维。甚至，在上海举行的 2015 年首届 GOPS 全球运维大会，就是运维四大“门派”的集中会演。好不热闹。当年，自动化运维还并非广为人知，由此看来，腾讯游戏确实很有前瞻性。

2016 年，Gartner 提出了 AIOps，基于算法的 IT 运维（Algorithmic IT Operations）。这时也并非智能运维。只是到了 2017 年，Gartner 才修正为 Artificial Intelligence Operations，即目前被广泛接受的人工智能运维。

为什么要回顾这段刚刚发生的历史？这个其实也是我们国人很自豪的地方，也就是说，基于海量业务场景，国内的运维技术及能力可能不仅仅没落后于国外，而且甚至领先于国外。据我所知，截至目前，Facebook 聚焦在将 AI 应用于广告营收，但尚无 AIOps 相关应用。

AIOps，将 AI 这样“高大上”的名词和 Ops 结合到一起，很是美妙。这对于运维而言，可谓一次重生（甚至说，野百合也会有春天）。本来处于软件生产链最末端的、容易被“鄙视”、因为远离业务部门而老背锅的运维同仁，有机会借助 AIOps，从运维升格为技术运营，通过改善产品的用户体验、日活、DAU 及营收和利润，从而和业务部门建立广泛而又建设性的合作，彻底甩掉背锅的命运。

AIOps 可能更“招人待见”（虽然广义来说，AIOps 是 DevOps 在 Ops 侧的高阶实现）。DevOps，从名词上看，容易被人认为是 Dev 把 Ops 吃掉了，而且国内众多企业因为组织结构或监管的要求，难以全面实践 DevOps。但对于 AIOps，运维部门或数据中心可以在自己的能力域竖井式发展，是一种既有面子又有里子的好事情。

那么，互联网中小企业及广大传统企业如何践行 AIOps？毕竟 AIOps 容易让人感觉是阳春白雪，有些高山仰止。

为了集中解决这个问题，为了让广大企事业单位能充分享受到大企业的红利，在中国信息通信研究院指导下，由高效运维社区牵头，邀请到国内互联网、金融及通信行业顶尖企事业单位的 AIOps 负责人，一起制定了国内外首个 AIOps 白皮书《企业级 AIOps 实践建议白皮书》，并在云计算标准及开源推进委员会下正式成立 AIOps 标准工作组，以推进相关标准的制定工作。

AIOps 白皮书的发布在国内引起了很大的轰动，这是凝聚各大 AIOps 顶级专家智慧共识的框架及指南。但美中不足的是，没有具体的实践落地。

本书的出现，很好地解决了上述问题。

我们惊喜地发现，不谋而合的是，新浪微博出品的本书，在整体认知、内容编排及落地实践上，和白皮书基本一致，并可实际指导企业进行 AIOps 平台的建设，可谓相得益彰。

本书同样认为，AIOps 对运维而言是一个新机遇，详细解读了智能运维基础设施，包括开源数据采集技术、分布式消息队列、大数据存储技术、大规模数据离线计算及实时计算、时序数据分析框架等。这些内容和 AIOps 白皮书定义的 AIOps 能力框架如出一辙，但又可实践指导落地，甚至到了 Step By Step 的程度。

本书是关于机器学习、数据聚会与关联技术、数据异常点检测技术、故障诊断和分析策略及趋势预测算法的，更是对 AIOps 白皮书关键技术章节的深度细化和补充。

只有书生才会纸上谈兵。本书更多的是介绍基于新浪微博平台技术部门及广告技术部门的生产实践，最后也很大方地给出了微博广告及平台的智能监控系统实现。

本书有理有据，有说服力。文字历练、表达清晰。相关实践可落地，让有志于实践 AIOps 的企业“开箱即用”，实乃国内运维同仁的又一幸事。故此推荐。

萧田国

高效运维社区发起人、AIOps 标准及白皮书发起人

前言

为什么要写这本书

中国互联网发展非常迅速，一方面得益于互联网基础设施的不断完善；另一方面得益于中国巨大的用户人群和消费市场。网络从 PC 到移动互联网时代过渡非常快，2017 年微博用户中移动端占比已经达到 92%，移动互联网的兴起带来了前所未有的新格局，围绕手机移动端的应用生态逐渐形成，大量 PC 时代的公司已经将产品的主战场转移到移动端。

据中国互联网络信息中心（CNNIC）发布的第 41 次《中国互联网络发展状况统计报告》，截至 2017 年 12 月，中国网民规模达到 7.72 亿，手机网民占比 97.5%，手机支付用户规模增长迅速，达到 5.27 亿，网络直播用户规模达到 4.22 亿，中国拥有人工智能企业 592 家，占全球总数的 23.3%。

互联网尤其是移动互联网的发展，也给企业带来了极大的技术挑战，如何保障线上产品各个服务和系统的稳定性、如何快速高效地诊断问题和定位问题等成为企业所面临的核心问题，而这些问题通过传统的运维方式已经无法有效得到解决。尤其是在大数据复杂场景下，对运维有了新的期待，也对运维有了更高的要求。

首先，系统产生的数据在很大程度上反映了系统状态和产品逻辑，监控系统需要具备快速搜集和处理数据的能力，能经过复杂的数据清洗并从大规模数据中抽取监控需要的指标，尤其是能对时序数据进行 ETL 及存储分析，将异构数据转化成监控系统能够理解的结构化数据。在这个阶段，如何保证数据的一致性和准确性、如何降低时延提高数据吞吐、如何降低监控系统对业务资源的影响等，都是在大数据环境下要解决的问题。

其次，监控系统要与报警系统融合，报警系统承载的是系统风险提示，对准确率要求很高，然而在现实环境下，大部分报警系统都存在报警项繁多、报警次数频繁、报警不准确等问题。

再次，智能化故障诊断、异常点检测、根因分析等是智能运维要解决的核心问题，然而智

能化系统要建立在高效的平台化基础上，目前极大一部分公司还没有真正进入平台化阶段，这就为智能运维带来了极大的阻力。

最后，智能运维需要运维工程师具备一定的人工智能、机器学习及深度学习等算法和建模能力，然而就目前来讲，运维工程师在这方面的技能是比较欠缺的。

国内在智能运维技术上，百度、阿里巴巴、腾讯和微博都有相当程度的经验积累，也得益于这些企业的运维平台化的体系逐渐成熟，在跟这些团队交流的时候，大家都有一个共识，就是希望能将智能运维技术应用到运维的各个维度，也相信智能运维将彻底改变运维的现有体系，并将极大丰富和完善传统运维，提高运维效率。

目前市面上与运维相关的书籍更多的是介绍某个单一技术的运维方法，少有对智能运维进行全面介绍的书籍，因此，我们觉得非常有必要梳理编写一本大数据场景下的智能运维技术的书籍，全面完整地介绍智能运维的技术体系，以及在大企业的运维实践经验，让读者更加了解运维的技术方向，在实践中能够有所借鉴。同时，也能帮助运维工程师在一定程度上了解机器学习的常见算法模型，以及如何将它们应用到运维工作中。

读者对象

本书面向的读者主要包括：

- 运维工程师
- 运维开发工程师
- 运维架构师
- 大数据工程师
- 对运维和大数据，以及 AIOps 感兴趣的工程师

如何阅读这本书

本书主要分 4 篇：第 1 篇运维发展史，重点阐述当前运维的发展现状及面临的技术挑战。第 2 篇智能运维基础设施，重点讲述大数据场景下的数据存储、大数据处理和分析的方法与经验，以及海量数据多维度多指标的处理分析技术。第 3 篇智能运维技术详解，重点关注在新时

期大数据时代下智能化的运维技术，包括数据聚合与关联、数据异常点检测、故障诊断和分析、趋势预测算法；第4篇技术案例详解，为大家梳理了通过开源框架 ELK 快速构建智能监控系统的整体方案，还将分享微博平台和微博广告两个不同业务场景下智能监控系统的技术实践。

具体而言，每篇和章节的主要内容如下：

第1篇 开门见山：运维发展史

- 第1章运维现状，主要介绍运维职责、传统运维、运维分类等，我们将从运维发展的四个阶段：人工、工具和自动化、平台化、智能化来介绍运维的发展现状。
- 第2章智能运维，主要介绍运维面临的挑战，传统运维在海量数据存储、分析、处理，多维度多指标及复杂业务等方面都有一定的局限性，随着运维新时代的到来，智能运维（AIOps）将为运维带来新的机会。

第2篇 站在巨人肩上：智能运维基础设施

- 第3章开源数据采集技术，重点介绍 Filebeat、Logstash 等开源数据采集工具。
- 第4章分布式消息队列，重点介绍以 Kafka 为代表的分布式消息队列及相关技术。
- 第5章大数据存储技术，重点介绍大数据的存储相关技术，这也是大数据场景下的智能运维基础。
- 第6章大规模数据离线计算分析，通过大数据 ETL 技术、Hadoop 技术生态讨论在大数据场景下如何进行离线计算和分析。
- 第7章实时计算框架，重点介绍在实时流计算方面的相关技术和框架，将探讨如何在监控系统中选择和使用实时计算框架。
- 第8章时序数据分析框架，实时监控系統处理的是时序数据，本章介绍常见的时序数据分析框架及使用方法。
- 第9章机器学习框架，智能化是运维、大数据和 AI 的结合，本章简单介绍机器学习框架，并以 TensorFlow 为例介绍如何进行模型训练和实践。

第3篇 运维新时代：智能运维技术详解

- 第10章数据聚合与关联技术，在数据聚合方面讨论聚合方法、多维度数据聚合技巧，以及如何降低维度；在数据关联方面介绍如何在实时流场景下进行时序数据的关联。

- 第 11 章数据异常点检测技术，本章结合运维面临的如异常点检测、动态阈值等常见问题，共同讨论解决这些问题的一些相关模型和算法。
- 第 12 章故障诊断和分析策略，故障诊断是智能运维的一个很重要的研究方向，本章讨论智能运维在故障诊断、决策树模型、关联分析等方面的策略和模型。
- 第 13 章趋势预测算法，主要介绍走势/趋势预测方面的常见模型和方法，包括 ARIMA 及基于机器学习的 LSTM 预测技术。

第 4 篇 智能运维架构实践：技术案例详解

- 第 14 章快速构建日志监控系统，以 ELK 为例介绍如何使用开源框架快速搭建日志监控系统。
- 第 15 章微博广告智能监控系统，全面介绍微博广告智能监控系统架构和设计原理。
- 第 16 章微博平台通用监控系统，以微博平台监控系统为例，全面介绍通用监控系统的设计思路和具体架构。

需要注意的是，本文提到的智能运维即指 AIOps，后续篇章将不再进行说明和区分。

勘误和支持

由于笔者的水平有限，编写时间仓促，同时本书在创作过程中参考了大量的国内外技术，并结合实践经验进行了系统性总结。由于技术的发展非常迅速，书中难免会出现一些错误或者不准确的地方，恳请读者批评指正。

智能运维技术在发生着翻天覆地的变化，我们希望更多的人能参与到这个过程中，共同推动智能运维技术的普及，欢迎通过微信或者邮件与我们进行讨论。你可以通过微信 justAStriver、微博@AndrewPD 或者电子邮件 contact@andrewpd.com 联系到我们，期待能够得到你们的真挚反馈，在技术之路上互勉共进。

特别致谢

我们花费了大量的时间总结智能运维方面的技术并整理成书，在此非常感谢微博广告基础架构团队的各位同事，尤其感谢车亚强、刘俊、陆松林、王莉、朱伟（按姓氏拼音排序）等人