HZ BOOKS
华章科技

· 网络空间安全技术丛书 ·

CLOUD DATA
SECURITY

云数据安全

徐鹏 林璟锵 金海 王蔚 王琼霄 著

机械工业出版社
China Machine Press

云数据安全

CLOUD DATA
SECURITY

徐鹏 林璟锵 金海 王蔚 王琼霄 著

机械工业出版社
China Machine Press

图书在版编目（CIP）数据

云数据安全 / 徐鹏等著 . —北京：机械工业出版社，2018.8
（网络空间安全技术丛书）

ISBN 978-7-111-60509-6

I. 云… II. 徐… III. 计算机网络 – 网络安全 IV. TP393.08

中国版本图书馆 CIP 数据核字（2018）第 164431 号

本书主要介绍云数据安全的相关背景、几种重要的云数据安全的保护方法和针对具体应用场景的一种集成化云数据安全保护系统。全书包括 6 章：第 1 章介绍云计算的基本概念与发展历程、云数据面临的威胁、云数据安全的核心需求与技术思路；第 2 章介绍非共享数据在加密存储条件下实现密文去重的方法；第 3 章介绍不可信云平台条件下云数据的加密共享方法；第 4 章介绍不可信云平台条件下共享云数据的安全搜索方法；第 5 章介绍不可信云平台条件下共享云数据的安全实时协同编辑方法；第 6 章结合部分前沿密码技术，提出一套面向云邮件系统的云计算数据安全集成解决方案。

本书适合高校网络空间安全相关专业的学生、教师和云数据安全的研究者阅读。

出版发行：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：张梦玲　　责任校对：李秋荣
印　　刷：北京诚信伟业印刷有限公司　　版　　次：2018 年 8 月第 1 版第 1 次印刷
开　　本：186mm × 240mm　1/16　　印　　张：11.25
书　　号：ISBN 978-7-111-60509-6　　定　　价：59.00 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

华章科技
HZBOOKS | Science & Technology

前　言

云计算是一种以互联网为基础并具有动态延展能力的计算模式，能实现随时随地、按需、快速便捷地使用共享计算、存储、网络、应用程序等资源。云计算能通过网络将海量规模的数据计算处理程序自动拆分成若干个子程序，再由多台服务器组成的庞大网络集群进行分析和计算，之后将处理结果返回给个人用户或者企业，在几秒的时间内就能处理庞大的数据信息。借助基于云计算搭建的云平台，能降低企业运营成本，提高资源利用率，减少用户在基础设施建设、维护和软件管理等方面的开销。云计算这种创新模式已经成为新一代信息技术产业发展的根本动力，正在引领各行各业的深刻变革。

云计算在为用户提供高质量、快速便捷的服务的同时，也面临着一些严峻的挑战和问题。由于云平台中保存了大量的用户敏感与隐私数据，如何确保此类用户数据的安全成为非常重要的问题。因此，云数据的安全性已经成为近几年来云计算领域广泛关注的热点。

本书主要介绍云数据安全的相关背景、几种重要的云数据安全的保护方法和针对具体应用场景的一种集成化云数据安全保护系统。全书包括 6 章：第 1 章介绍云计算的基本概念与发展历程、云数据面临的威胁、云数据安全的核心需求与技术思路；第 2 章介绍非共享数据在加密存储条件下实现密文去重的方法；第 3 章介绍不可信云平台条件下云数据的加密共享方法；第 4 章介绍不可信云平台条件下共享云数据的安全搜索方法；第 5 章介绍不可信云平台条件下共享云数据的安全实时协同编辑方法；第 6 章结合部分前沿密码技术，提出一套面向云邮件系统的云计算数据安全集成解决方案。

目前国内外出版了很多关于云安全的书籍，它们大多数都是从宏观的理论角度阐述云计算安全的发展现状、面临的挑战与解决方法。本书从基本理论概念入手，通过“理论 + 实践”的模式由浅入深地分析云安全面临的威胁并介绍几种实现云数据安全的相关技术。本书不仅包括作者多年的科研和教学成果，还紧密结合当下前沿的学术研究理论，兼顾理论和实

战，具备广度和深度，覆盖云数据安全的核心概念和知识点，第6章提出的面向云邮件系统的云计算数据安全集成解决方案具有较强的可操作性，便于读者学习和实战。

本书可作为高等院校信息安全和密码学专业本科生与研究生的教材，也可供信息安全从业人员、云计算安全研究人员参考。

本书的绝大部分内容来自作者的科研成果，同时也参考了信息安全领域的大量研究成果和相关技术资料。本书的编写得到国家重点基础研究发展计划（973项目）“云计算安全基础理论与方法研究”（No. 2014CB340600）的资助，特此表示感谢。

目　录

第1章

引　言

云计算的出现是传统IT和通信技术进步、需求推动和商业模式变化共同促进的结果，具有以网络为中心、以服务为提供方式、高扩展性、高可靠性和资源使用透明化等主要特征。业界认为云计算是继PC、互联网之后信息产业的第三次变革，将对社会信息化发展产生深远影响。本章将简单介绍云计算的概念与发展方向、云安全的发展现状与面临的主要挑战。

1.1　云计算与云安全

自2006年Google公司提出云计算概念以来，云计算迅速发展，并引起了产业界、学术界和政府部门的高度关注。云计算产业作为中国政府重点扶持的战略性新兴产业，已经逐步地从概念走向现实。2010年10月10日，中国政府将云计算产业列入国家重点培育和发展的战略性新兴产业。2012年，通信、互联网等行业“十二五”规划出台，物联网和云计算工程也被作为中国“十二五”发展的20项重点工程之一。目前，云计算已经在大多数企业中得到普及和应用。其中，伴随着云计算的快速发展，越来越多的用户将数据托管到云服务器。但是，由于引入虚拟化技术的云计算系统所特有的服务模式及其前所未有的开放性、复杂性和可扩展性等特点，传统的安全技术无法完全保证用户托管到云服务器的数据的安全和云计算平台自身的安全。因此，云计算为信息安全领域带来了新的挑战，也为信息系统引入了新的风险。现阶段云安全研究得到广泛的关注，成为云计算应用发展的重要课题之一。

1.1.1　云计算基本概念

云计算是以网络技术、虚拟化技术、分布式计算技术为基础，以按需分配为业务模式的新一代网络化商业计算模式，具有动态扩展、资源共享、宽带接入等特点。借助开放的网络环境，云计算为用户提供了强大的计算和存储能力，现已逐渐在产业界得到广泛应用[1-2]。

自2006年亚马逊推出弹性云之后，云计算从概念逐步走向了实践阶段。调查显示，2006年全球2900万项IT工作负载中，有98%是由传统IT技术完成的，而借助云计算的只有2%。但是，到了2016年，全球IT工作负载增加到了1.6亿项，其中在传统IT、公有云、私有云上完成的工作比例分别为73%、15%、12%，承载在云计算上的IT负载已经占到了整个IT工作的四分之一以上。根据中国信息通信研究院发布的《云计算关键行业应用报告（2017年）》[3]，全球云计算市场总体持续增长。2016年，典型云服务市场规模达到654.8亿美元，增速25.4%，预计2020年将达到1435.3亿美元，年复合增长率达21.7%。

与此同时，中国的云计算行业也蓬勃发展，保持高速增长趋势。2016年我国的云计算产业整体增速达到35.9%，市场规模达到514.9亿元。其中私有云规模达到344.8亿元，公有云服务市场整体规模约为170.1亿元，相对2015年增长66.0%，并且保持高速增长的速度，预计2020年公有云市场规模将达到603.6亿元。

云计算可以通过网络将庞大的计算处理程序自动分拆成无数个较小的子程序，再交给由多台服务器组成的庞大系统，经它搜寻、计算、分析之后将处理结果回传给用户。最简单的云计算技术在网络服务中已随处可见，例如搜索引擎、网络信箱等，使用者只要输入简单指令即可得到大量信息。未来如手机、GPS等移动装置都可以通过云计算技术，发展出更多的应用服务。进一步的云计算不仅只有资料搜寻、分析的功能，未来如分析DNA结构、基因图谱定序、解析癌症细胞等，都可以通过这项技术轻易完成。

云计算主要有3种服务：基础架构即服务（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）。与传统IT服务模式相比，云计算自诞生以来就具有显著的特点和巨大的优势，主要体现在建设成本、扩展性、可靠性、服务质量以及远程访问等方面[4]。

IaaS 把厂商的由多台服务器组成的“云服务器”基础设施作为计量服务提供给用户。它将内存、I/O 设备、存储和计算能力整合成一个虚拟的资源池，为整个业界提供所需要的存储资源和虚拟化服务器等服务。这是一种托管型硬件方式，用户付费使用厂商的硬件设施。例如 Amazon Web 服务、IBM 的 BlueCloud 等，均是将基础设施作为服务出租。IaaS 服务的优点是用户按需租用相应计算能力和存储能力，大大降低了硬件开销。

PaaS 把开发环境作为一种服务来提供。这是一种分布式平台服务，厂商提供开发环境、服务器平台、硬件资源等服务给用户，用户在其平台基础上定制开发自己的应用程序，并通过其服务器和互联网传递给其他用户。PaaS 服务能够给企业或个人提供研发的中间件平台，提供应用程序开发、数据库、应用服务器、试验、托管及应用服务。PaaS 可以理解为在 IaaS 提供的硬件服务之上，还额外搭建好了服务器环境、中间件、数据库等。开发者只需要将网络应用代码上传部署，应用就可以运行起来，这样既降低了 IT 运维成本，还省去了大量的开发与运维工作量。

SaaS 提供商将应用软件统一部署在自己的服务器上，用户根据需求通过互联网向厂商订购应用软件服务，服务提供商根据用户所订购软件的数量、时间的长短等因素收费，并且通过浏览器向用户提供软件的功能。

云计算被视为信息技术的第三次浪潮，是未来新一代信息技术变革、IT 应用方式变革的核心，将带来工作方式和商业模式的根本性改变，云计算的核心发展方向已经成为当前 IT 界乃至全社会关注的焦点和热点。

1.1.2　云计算核心发展方向

1. 混合云成为重要发展方向

云计算的部署模式可以分为公有云、私有云和混合云，其中混合云是将公有云、私有云进行混合和匹配，将私有云、公有云的优点融为一体的云计算服务架构。混合云可以同时克服公有云与私有云的不足，比如：公有云的安全性和可控性问题；私有云的性价比不高、弹性扩展不足的问题等。对于用户来讲，希望得到一体化解决方案，同时希望在公有云中构建私有云。当用户认为公有云不能够满足企业需求的时候，可以在公有云环境中构建私有云来实现混合云。基于混合云的优势，可以预测混合云是未来市场的发展趋势，也是目前企业越来越多采用的商业云解决方案，根据最新报告，2017 年的云计算市场中，混

合云是主旋律，近一半的大型企业在2017年年底部署了混合云。

2. 云计算是人工智能的底层建筑

人工智能、大数据和云计算，三者可以称为铁三角关系，是天然耦合、不可分割的。2017年中国“互联网+”数字经济峰会上，腾讯总裁马化腾表示：“大家在讨论，未来的趋势是云、人工智能，还是大数据。”他说，“都可以整合在一块儿，未来就是在云服务器上用人工智能处理大数据。”云计算提供对海量数据的强大计算、存储能力，人工智能需依托云计算的强大计算能力进行训练、推理和预测，大数据又为人工智能的训练、推理和预测提供海量的运算和测试数据集。

人工智能是一门研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的技术科学。机器学习是人工智能的核心要素，主要是研究计算机如何模拟或实现人类的学习行为，以获取新的知识或技能。与以往的算法相比，这一算法完全利用输入的数据自行模拟和构建相应的模型结构，具有灵活且可以根据不同的训练数据自优化的能力等优点，但这些优点带来的是显著增加的运算量。在云计算的高计算能力应用过程中，并行计算可以利用多个处理器解决一个大问题，提升了计算效率，能为人工智能提供近乎无限的运算能力。

3. 云计算助力区块链

作为第一个用于商业和个人交易的去中心化、点对点全球平台，区块链的出现可以说是近年来最令人兴奋的技术突破之一。区块链是一个可信任的、由加密技术保障安全的分布式账本，是数字时代最安全的系统。

云计算服务具有资源弹性伸缩、快速调整、低成本、高可靠性的特质，能够帮助中小企业快速、低成本地进行区块链开发和部署。两项技术融合，将加速区块链技术的成熟，推动区块链从金融业向更多领域拓展。目前，已经有大量的企业开始选择基于云的区块链网络。这一趋势在2018年仍在延续。将区块链应用于全球供应链，预计每年可能会产生超过1000亿美元的效益。最佳的区块链系统将以API或者解决方案的形式存放在云服务器上，供企业大规模使用。

4. 容器服务

容器服务具有部署速度快、开发和测试更敏捷、系统利用率高、资源成本低等优势，

随着容器技术的成熟和接受度越来越高，容器技术将更加广泛地被用户采用。谷歌的 Container Engine、AWS 的 Elastic Container Service、微软的 Azure Container Service 等容器技术日臻成熟，容器集群管理平台也更加完善，以 Kubernetes 为代表的各类工具可帮助用户实现网络、安全与存储功能的容器化转型。在国内，各家公司积极实践，用户对于容器技术的接受度得到提升，根据调研机构数据，近 87% 的用户考虑使用容器技术。

5. 云安全需优先考虑

近年来，重大 IT 安全漏洞对商业、政府或大众产生不利影响的事件时有发生。用户对云服务持不信任的态度是正常现象，大多数用户对云计算安全、性能、可靠性等都持怀疑的态度。用户在选择云服务提供商的时候，首要关心的三个要素为稳定性、安全性和网络质量。虽然更多的企业在向云服务迁移，但是安全问题仍然是云服务提供商必须解决的一大障碍。云服务商如果想要企业和组织大规模地运用云计算技术和平台，就必须全面完善地分析云计算面临的各种安全威胁，着手解决这些安全问题并提出完备的云平台安全防护方案。云安全联盟 2016 年发布的“十二大云安全威胁”报告[5]中提到，云环境下安全威胁包括：数据泄露、身份验证和凭证被盗、界面和应用程序编程接口（API）安全、系统漏洞问题、账户劫持、内部恶意人员、高级持续性威胁、永久数据丢失、缺乏尽职调查、云服务滥用、拒绝服务攻击、共享技术和共享危险。

1.1.3 云安全重要性

云计算作为近些年信息通信领域发展最迅速的产业之一，对国民经济和社会发展的战略支撑与创新引领作用日益凸显。加快发展云计算产业，对加快经济增长，促进产业结构创新升级，推动与传统产业的融合发展具有重要意义。但是，在发展的同时难免会出现一些问题。自从云计算诞生以来它就一直经受着安全问题的考验，其中最大的问题就是系统漏洞和数据安全问题。人们利用云计算得到的一些数据会聚集在云中，他们时刻担心这些数据被云破坏或者传播出去，云安全正变得更加不容轻视。尤其在群雄逐鹿的中国云计算市场，用户信心正在建立过程中，安全性问题是云计算厂商首先要考虑的。

云安全是对云平台自身的安全保护，主要利用面向云架构环境的安全策略、技术产品解决云环境下的安全问题，提升云平台自身的安全性，保障云计算业务的可用性，数据机密性、完整性和隐私权的保护等。随着云服务器承载的重要信息数量增多，云计算与人工

智能逐渐融合，安全问题也凸显出来，业务威胁、网络攻击和数据泄露正在成为云计算上的“达摩克利斯之剑”。因此，保护云设施安全以及云服务器的数据、业务和应用安全成为发展基础和关键。云安全也随之迎来2.0时代，从安全架构、技术应用到技术生态、安全运营都将发生革命性变化。根据Gartner的数据，全球云安全服务市场保持强劲增长势头，在2017年达到59亿美元，相比2016年增长21%。云安全服务市场的整体增速高于信息安全总体市场。Gartner预计，云安全服务市场将在2020年接近90亿美元。

2017年，Alert Logic公司公布的《2017年云安全报告》[6]分析了该公司产品2015年8月至2017年由3800多名客户发现的超过220万起安全事件。通过图1-1可看出，公有云服务经历的安全事件最少。平均来说，在18个月内，在公有云平台上运行应用程序的每个用户经历了405次安全事件；在自构建云系统上运行应用程序的每个用户经历了612次安全事件，与公有云相比增加了51%；在混合云上，每个用户经历了977次安全事件，与公有云相比增加了141%；在托管的私有云上，每个用户经历了684次安全事件，与公有云相比增加了69%。

这些云安全事件的爆发说明云服务存在着极大的安全威胁，重视和考虑云数据安全已经刻不容缓。不过，云安全也在发展进步，越来越多行之有效的安全和控制措施在不断投入实际应用中。

1.2　云数据面临的威胁

云计算需要以分布式计算方式处理海量数据，即很多数据计算是由多个独立的计算资源共同完成的，这样，大量的中间数据需要通过网络传递，这个过程往往缺乏保护，存在极大的安全隐患。不仅是处理中的数据，存储在数据中心里的数据也面临严重的安全威胁。另外，云计算缺乏对数据内容的辨识能力，往往在得到数据后直接计算，缺少检查和校验机制，因此会使一些无效数据或者伪造数据混在其中。这一方面可能影响计算的结果，另一方面也占用大量计算资源，影响计算效率。云计算安全联盟（CSA）近期发布的报告[5]总结了12种威胁云计算安全的“罪魁祸首”，其中数据泄露、数据丢失和数据劫持三种威胁排名靠前。针对云数据安全问题，本节将介绍几种常见的云数据安全威胁及其可能造成的危害。

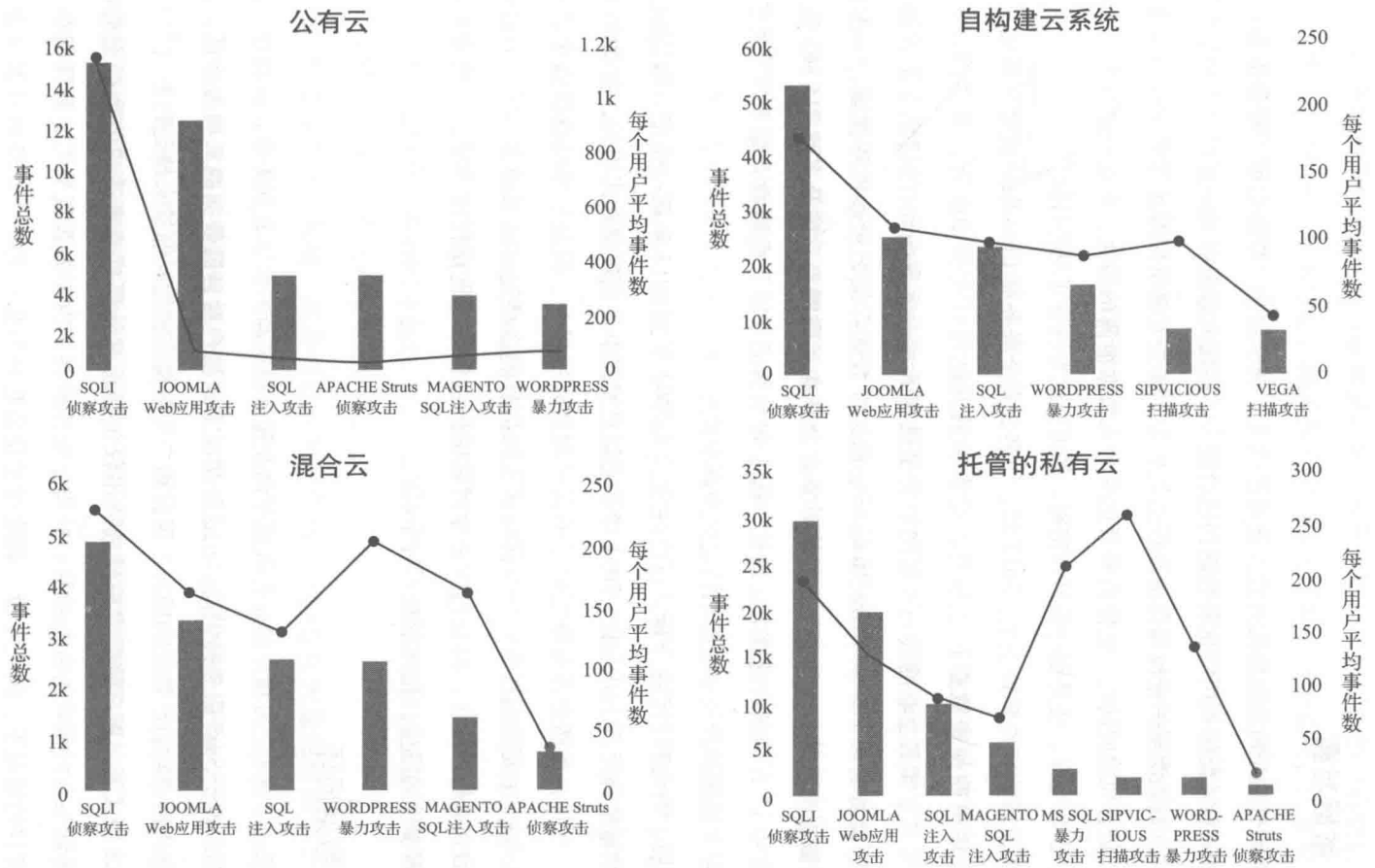

图1-1 《2017年云安全报告》

1.2.1　数据泄露

云计算时代也是海量数据的时代。在此基础上，网络搜索、数据挖掘、商务智能等技术的发展令相关行业能够利用信息和数据创造价值[7]。伴随大数据而来的是层出不穷的数据泄露事件，并且数据泄露的规模和范围也在迅速扩大。这些大规模数据泄露事件在为企业带来财产损失、信誉风险的同时，也使消费者饱受个人数据泄露的困扰。在这些事件中，无论是信用卡号、医疗记录，还是银行账号和密码，都可能成为网络罪犯的目标[8]。

云数据泄露的事件频频发生，2017 年，知名云安全服务商 Cloudflare 被曝泄露用户 HTTPS 网络会话中的加密数据长达数月，受影响的网站预计至少 200 万，其中涉及 Uber、1Password 等多家知名互联网公司。据网络安全机构 Kromtech Security 的信息专家披露，疑似来自医疗设备公司 Patient Home Monitoring 的医疗数据存储记录遭破解泄露，一份包含 47GB 医疗数据文件的 Amazon S3 云存储对象被公开在互联网上，涉及多达 315 363 份 PDF 档案，包含近 15 万名患者的姓名、地址、医生和病例记录以及血液检查结果等隐私信息，这又是一起大规模的涉及公众隐私信息的泄露事件。

对收集的案例进行汇总分析，可以发现，从 2002 年到 2017 年第一季度，敏感信息泄露事件的数量整体呈现上升态势，2011 年敏感信息泄露事件出现爆发式增长，在 2016 年达到了峰值。近年，虽然企业对敏感信息的保护程度有所提升，但是敏感信息泄露事件仍然呈现上升趋势，主要原因在于：一方面黑客获取信息的途径变得越来越多，另一方面存储敏感信息的企业越来越多，但是很多企业对敏感信息保护的重视程度不足，导致越来越多的信息泄露事件的发生，整体形势不容乐观。

1.2.2　数据损坏

SaaS 模式下成千上万租户的业务数据存储在服务提供商的共享数据库中，在物理上，软件与服务都部署于 SaaS 服务提供商，在这种情况下，租户对自己数据的掌控力降低，很难觉察到恶意攻击者对自己数据的破坏，因此租户的数据完整性问题越来越受到重视[9-10]。

在 SaaS 模式下，租户数据完整性威胁不仅包括恶意外部攻击者造成的数据完整性破坏，同时还面临着内部攻击者的威胁：首先，服务提供商内部的恶意员工，有可能在没有得到租户授权的情况下，随意篡改、删除或者伪造租户数据；其次，在经历外部攻击、服