论基于云计算的数据存储安全策略

赵文娟
杭州华澜微电子股份有限公司，浙江 杭州 310000
DOI: 10.61369/TACS.2025050013

摘　　要： 随着基于云计算的大数据应用迅速发展，低成本人工智能（AI）大模型的普及，数据存储与管理方式正经历深刻变革。云计算为大数据存储提供了弹性、高可用和按需分配的资源支持，AI 的平民化大大提高了社会整体数据处理的效率，但数据的安全性问题也愈发突出，数据泄露、服务中断、完整性受损等问题都是难以承受的。如何在保证性能和成本优势的同时，确保数据在云环境下的安全，已成为业界关注的焦点。本文将从云计算与大数据存储的技术基础与关键问题着手，探究云计算环境下的大数据存储主要安全威胁，并针对性提出包括数据加密、受限访问与身份认证、数据隔离以及数据备份与灾难恢复在内的安全策略。

关 键 词： 云计算；大数据；存储安全；数据安全策略

On the Security Strategy of Data Storage Based on Cloud Computing

Zhao Wenjuan
Hangzhou Hualan Microelectronics Co., Ltd, Hangzhou, Zhejiang 310000

Abstract: With the rapid development of cloud based big data applications and the popularity of low-cost artificial intelligence (AI) models, data storage and management methods are undergoing profound changes. Cloud computing provides elastic, highly available, and on-demand resource support for big data storage. The popularization of AI has greatly improved the overall efficiency of data processing in society, but data security issues have become increasingly prominent. Issues such as data leakage, service interruption, and integrity damage are all unbearable. How to ensure data security in cloud environments while maintaining performance and cost advantages has become a focus of attention in the industry. This article will start with the technical foundations and key issues of cloud computing and big data storage, explore the main security threats of big data storage in cloud computing environments, and propose targeted security strategies including data encryption, restricted access and identity authentication, data isolation, and data backup and disaster recovery.

Keywords: cloud computing; big data; storage security; data security strategy

引言

过去十年间，数据量、数据类型及处理复杂性呈现爆炸式增长，传统的本地存储架构已难以满足需求。云计算作为一种通过网络提供按需计算资源的分布式计算模型（如服务器、存储池），实现了计算能力的弹性扩展和高效利用，以其灵活的资源调度能力和高扩展性，广泛应用于大数据的存储与处理领域。但由于数据集中化与虚拟化特征，云环境的安全威胁与风险也呈现多样化与隐蔽化趋势。例如，硬件损坏、算法缺陷、链路异常断开等都可能导致重大数据损失与业务中断。当前，我们在享受大数据存储带来的便利的同时，容易忽视存储机制的安全性设计，导致风险管理被动化。基于此，深入研究大数据存储安全策略，不仅具有技术意义，更关乎数据资产的完整与业务连续性，对政府、企业及科研机构均有现实价值。

一、云计算与数据存储技术基础

（一）云计算和低成本人工智能大模型的特征

云计算是大数据存储技术中的典型应用，它是一种以网络为基础的资源交付与管理模式，通过虚拟化技术将计算、存储、网络等资源进行集中管理和动态分配，用户无需关心底层硬件配置，即可按需获取所需服务。它的核心特征包括弹性扩展、资源池化、自助服务以及广泛的网络接入能力。云计算的分布式处理能力与集中化管理优势，可以有效解决海量数据存储与计算的基础需求。然而，云计算环境下资源的动态性和虚拟化特征，也使得数据安全、性能保障和服务可用性成为技术管理中的重点与难点。[1]

低成本人工智能（AI）大模型在传统大模型（如 GPT - 3、PaLM 等）的基础上，通过技术优化、架构创新、资源高效利用等手段，在训练成本、部署成本、使用成本等环节实现显著降低，同时保持或接近传统大模型的性能（如语言理解、生成能力）。

（二）大数据存储模型与架构

大数据存储主要依托分布式架构实现高并发访问和高可用性，其常见模型包括分布式文件系统（如 HDFS）、对象存储和 NoSQL 数据库等。分布式文件系统以数据分块与多副本机制保证存储容量和可靠性；对象存储通过唯一标识符管理非结构化数据，适合图片、音视频等海量文件的存储；NoSQL 数据库则在处理结构化与半结构化数据时，具备高伸缩性和快速查询能力。在架构上，大数据存储通常分为数据采集、传输、存储和访问四个层次，各层通过网络和接口协议协同工作，确保数据在不同节点间的可用性和一致性。为适应云计算环境，这些架构需要具备灵活的横向扩展能力，同时支持跨地域、跨平台的数据协作与冗余存储，从而在面对业务波动时保持稳定运行。

（三）大数据存储技术的关键

尽管云计算为大数据存储提供了高弹性和低成本优势，但技术应用中仍存在多方面挑战。首先是数据一致性问题，在分布式多节点环境下，网络延迟或节点故障可能导致数据版本不同步，影响业务正确性。其次是存储性能与扩展性的平衡，过度追求扩展可能造成访问延迟增加，而性能优化又可能限制规模扩展。第三是安全性与隐私保护，数据集中化存储使得其更容易成为攻击目标，需要在加密、访问控制、审计等方面强化设计。此外，存储系统的运维复杂度较高，需在硬件容错、自动恢复、资源调度等方面具备完善的策略与工具。这些问题决定了大数据存储技术不仅要关注容量和速度，还必须在架构设计与管理机制中充分考虑可靠性与安全性。[2]

二、云计算环境下的数据存储安全威胁

（一）数据泄露与未授权访问

在云计算环境中，大数据往往集中存放在共享资源池中，虽然虚拟化和多租户技术提高了资源利用率，但也带来了潜在的安全隐患。数据泄露常由未授权访问、身份冒用或管理漏洞引发。例如，攻击者可能利用弱口令、系统漏洞或应用接口缺陷绕过认证，从而直接读取或篡改敏感信息。此外，云服务提供商与用户之间的安全责任边界不清晰，也可能导致内部管理人员误操作或越权访问。对于金融、医疗、政府等领域而言，这类泄露事件会引发严重的法律与信誉风险。在 AI 应用中涉及的用户敏感信息如个人的身份证号、生物特征，公司敏感信息如技术核心机密、财务报表，在设备、服务器或跨平台传输时若未进行有效的数据保护，用户隐私可能会暴露、公司机密可能被截获或篡改。为了防范，需建立严格的访问控制策略，结合多因素认证、最小权限原则以及实时访问日志审计，并在数据传输和存储过程中应用加密技术，从源头降低被截获和滥用的风险。

（二）数据完整性和一致性问题

在分布式大数据存储中，数据被切分并存储于多个节点，系统需通过复制与同步机制来维持一致性。然而，网络延迟、节点故障或版本冲突可能导致数据不一致，进而影响计算结果的正确性。数据完整性问题则指数据在传输、存储或处理过程中遭到未授权修改或意外损坏，导致原始内容丢失或篡改。这类问题不仅会影响业务判断，还可能被恶意利用制造数据混乱。例如，攻击者通过中间人攻击修改传输数据，或利用系统漏洞直接更改数据库记录。为此，需要在数据传输和存储环节使用哈希校验、数字签名及有效的存储加密技术，定期执行一致性检测和副本修复。同时，采用强一致性协议或可配置一致性模型，在性能与一致性之间寻找平衡，以满足不同业务场景需求。

（三）数据丢失及物理存储设备的损坏

虽然云计算平台通常具备多副本存储和容灾机制，但数据丢失仍可能因物理存储设备故障、人为误操作或恶意破坏而发生。设备老化、控制器损坏、电源异常等硬件问题，会在短时间内导致多个存储节点不可用。此外，管理疏忽或脚本执行错误也可能导致数据被误删，而攻击者则可能通过破坏存储介质或加密勒索的方式使数据无法恢复。防范措施包括采用跨地域备份、冷热数据分离、定期快照与冗余存储等策略，同时在运维管理中引入严格的变更控制和操作审核制度。针对核心业务数据，通过双活数据中心架构实现关键数据实时同步与低延迟备份，确保主存储系统发生故障时，能够依据预设的 RPO（恢复点目标）和 RTO（恢复时间目标）服务级别协议，在分钟级时间内完成业务切换与数据恢复，最大限度保障业务连续性。该方案需重点考虑网络带宽、数据一致性校验机制及异地灾备中心的运维管理能力，形成完整的业务连续性管理框架。

三、云计算环境下数据存储安全策略分析

（一）数据加密及扰码

数据加密及扰码技术是保障云计算环境中数据存储安全的核心手段之一。其主要目标是在数据传输和存储过程中，确保敏感信息即使被非法获取也无法被解读。加密技术涵盖了对静态数据（静态加密）和传输数据（传输加密）的双重保护。静态加密通常通过对硬盘或数据库中的数据文件采用对称加密算法如国密、AES、TCG，确保数据即使存储介质被窃取也难以恢复原文；扰码技术的核心是伪随机序列的生成与应用，不同技术的差异体现在序列生成方式（LFSR、非线性反馈等）、同步机制（自同步、帧对齐等）和应用场景（通信、存储、加密等），常与编码（如 CRC、卷积码）、加密算法结合，共同保障数据的传输效率与安全性；传输加密则依赖于 TLS/SSL 协议保护数据在网络传输过程中免遭窃听或篡改。[3]

然而，加密技术的实施也带来计算资源和性能开销，尤其是在大数据环境下，如何在保证安全性的同时兼顾系统响应速度，是设计中的重要考量。云服务商需结合具体业务场景，合理配置硬件加速设备和优化算法实现，减少加密解密的性能损耗。数据加密技术通过保护数据的机密性，形成了数据安全防护的第一道屏障，是云计算大数据存储不可或缺的基础设施。同时，结合完善的密钥管理和算法创新，能够有效应对不断变化的安全威胁，为企业和用户的数据资产提供坚实保障。

（二）访问控制与身份认证

访问控制和身份认证机制是确保云计算环境中大数据安全访

问的关键环节，直接决定谁能访问哪些数据以及在何种权限下进行操作。身份认证是验证用户或系统身份的过程，常见方法包括用户名密码、多因素认证（MFA）、生物识别和基于证书的认证等。访问控制则基于身份认证结果，依照预先设定的策略授予或限制用户对数据资源的操作权限。主要模型有基于角色的访问控制（RBAC）、基于属性的访问控制（ABAC）和基于策略的访问控制（PBAC）等。细粒度权限分配和最小权限原则成为防止内部威胁和误操作的有效措施。实时监控和异常行为检测机制也不可忽视，它们能及时发现非正常访问，触发自动阻断或报警，保障数据安全。[4]

访问控制和身份认证机制中，软件层面的身份识别机制因依赖逻辑算法与代码实现，易受逆向工程、内存注入、钓鱼攻击等手段破解，其核心验证逻辑暴露于可篡改的软件环境，密钥或凭证存在被窃取、仿制的风险。相较之下，依托专用存储安全芯片构建可信执行环境将密钥存储、加密运算等核心操作固化于物理隔离的硬件单元，可抵御侧信道攻击、物理篡改等高级威胁，其根信任机制与不可克隆功能具备抗抵赖性与唯一性，能显著提升身份认证的底层安全性，实现更可靠的身份确权。

完善的身份认证与访问控制体系不仅是大数据存储安全的守门员，更是实现合规性要求的重要保障。随着云计算技术不断演进，结合人工智能和行为分析的智能认证与权限管理方案将成为未来发展的趋势，有助于构建更加安全、灵活的云上数据生态。

（三）数据隔离与安全

虚拟化技术作为云计算的基础支撑，其安全性直接影响大数据存储的整体防护效果。虚拟化通过在物理硬件上运行多个虚拟机（VM），实现资源的高效共享和隔离。然而，虚拟化层的漏洞、管理不善或者配置错误，都可能导致“逃逸攻击”，攻击者突破虚拟机边界，进而入侵宿主机或其他虚拟机，造成严重安全隐患。

从硬件维度实施存储隔离，通过物理或逻辑层面的硬件级边界划分，能构建更可靠的数据安全屏障。依托安全存储控制器芯片的硬件虚拟化技术如专用的安全分区，将敏感数据与普通数据的存储区域在硬件层面强制隔离，避免跨域访问。硬件隔离机制可规避软件层隔离的权限逃逸风险，其访问控制逻辑固化于硬件电路，能抵御恶意代码注入、内核级攻击等威胁。结合硬件安全模块的密钥管理，可实现数据存储的机密性与完整性双重保障，为核心数据提供底层级安全防护。

（四）数据备份与灾难恢复

数据备份与灾难恢复机制是确保云计算环境中大数据安全与业务连续性的基础保障。尽管云平台通常具备高可用架构，但设备故障、软件缺陷、人为误操作等都可能导致数据丢失或服务中断。完善的备份体系能够在关键时刻恢复数据，减少损失；灾难恢复则是通过预设的应急方案，实现系统的快速恢复与业务的持续运转。

在大数据存储中，备份策略需考虑数据量大、更新频繁和多样化的特点。常见方法包括全量备份、增量备份和差异备份，通过合理组合降低备份时间与存储成本。备份数据通常采用异地存储，避免单点故障对数据安全造成影响。同时，为防止备份数据被篡改或泄露，需结合加密技术确保备份内容的机密性和完整性。[5]

存储设备级别的硬件备份凭借物理层的数据冗余机制，其可靠性显著优于系统级备份。这类方案依托 RAID 阵列、硬盘硬 RAID 存储技术等硬件级冗余架构，通过专用控制器实时同步数据至冗余存储单元，实现毫秒级故障切换，存储设备的异常断电保护设计，可保证在系统供电突然断开的情况下保障数据完整性。相较之下，系统备份依赖软件逻辑与文件系统接口，易受操作系统漏洞、恶意软件篡改或备份进程中断影响。硬件备份则通过独立于宿主系统的专用芯片与总线通道完成数据复制，规避了上层软件栈的安全风险，坏道自动修复（SMART 技术）为数据完整性提供底层级保障。备份与恢复不仅关注技术层面，更注重流程规范和人员培训。多部门协同合作，形成完善的应急响应机制，提升整体抗风险能力。

四、结束语

云计算及 AI 为数据存储提供了全新的架构思路和经济模式，但其安全问题同样不容忽视。面对数据泄露、服务中断、完整性受损等多重威胁，单一防护手段难以构建稳固的安全体系。本文从技术基础出发，分析了云计算环境下大数据存储的安全威胁，并提出了包括加密技术、访问控制、数据隔离和备份恢复在内的多层次策略。这些策略相互配合，可显著提升存储系统的安全性与可靠性。随着区块链等新兴技术的引入，数据存储的安全保障有望更加智能化与自动化，但同时也需应对新形态的威胁。对管理者而言，应在技术部署与安全文化建设双向发力，才能在享受云计算带来便利的同时，稳固守护核心数据资产。

参考文献

[1] 戚小虎 . 一种基于云计算的计算机网络安全存储技术研究 [J]. 信息记录材料 ,2024,25(06):202-204+207.

[2] 荣华良 , 李春蕾 , 张燕平 , 等 . 基于云计算的大数据存储安全策略分析 [J]. 信息记录材料 ,2024,25(08):177-179+183.

[3] 吴刚 . 基于云计算的大数据存储安全策略 [J]. 中国宽带 ,2023,19(11):127-129.

[4] 李红霞 . 基于云计算的网络存储技术对数据安全的影响 [J]. 网络空间安全 ,2024,15(04):314-317.

[5] 王莉 , 王智 , 王丽珍 . 基于云计算的数据安全存储策略探析 [J]. 网络安全技术与应用 ,2021,(06):68-70.

