学校编码：10384　　分类号_____密级_____
学号：X2013232416　　UDC _____

厦门大学

工 程 硕 士 学 位 论 文

# 面向云计算的数据安全防护关键技术研究

# Research on Key Technologies of Data Security towards Cloud Computing

章 路

指 导 教 师： 王 备 战 教 授

专 业 名 称： 软 件 工 程

论文提交日期： 2015 年 09 月

论文答辩日期： 2015 年 11 月

学位授予日期： 2015 年 12 月

指 导 教 师：_____

答辩委员会主席：_____

2015 年 09 月

# 厦门大学学位论文原创性声明

本人呈交的学位论文是本人在导师指导下,独立完成的研究成果。本人在论文写作中参考其他个人或集体已经发表的研究成果，均在文中以适当方式明确标明，并符合法律规范和《厦门大学研究生学术活动规范（试行）》。

另外，该学位论文为（　　）课题（组）的研究成果，获得（　　）课题（组）经费或实验室的资助，在（　　）实验室完成。（请在以上括号内填写课题或课题组负责人或实验室名称，未有此项声明内容的，可以不作特别声明。）

声明人（签名）：

年　　月　　日

# 厦门大学学位论文著作权使用声明

本人同意厦门大学根据《中华人民共和国学位条例暂行实施办法》等规定保留和使用此学位论文，并向主管部门或其指定机构送交学位论文（包括纸质版和电子版），允许学位论文进入厦门大学图书馆及其数据库被查阅、借阅。本人同意厦门大学将学位论文加入全国博士、硕士学位论文共建单位数据库进行检索，将学位论文的标题和摘要汇编出版，采用影印、缩印或者其它方式合理复制学位论文。

本学位论文属于：

（　　）1.经厦门大学保密委员会审查核定的保密学位论文，于　　年　　月　　日解密，解密后适用上述授权。

（　√　）2.不保密，适用上述授权。

（请在以上相应括号内打“√”或填上相应内容。保密学位论文应是已经厦门大学保密委员会审定过的学位论文，未经厦门大学保密委员会审定的学位论文均为公开学位论文。此声明栏不填写的，默认为公开学位论文，均适用上述授权。）

声明人（签名）：

年　　月　　日

# 摘 要

近十年来，云计算成为当今世界信息领域的热门研究课题之一，也逐渐成为学术界、产业界最为关注的技术焦点。云计算提出的全新的服务方式变革和数据共享模式的变革，以其经济、高效、扩展性强的优点为互联网经济和全球的 IT 产业的发展注入了强大的动力，指引 IT 领域向专业化、规模化、集约化的趋势发展，是大型计算机从客户端-服务器端大转变之后的又一次巨变。

但从另一方面考虑,云计算所涉及到的信息安全问题也逐渐显现出来，成为其发展和普及过程中的“绊脚石”，如何为用户在云端存储的数据提供全方位的安全防护是目前亟待解决的问题。

本文针对目前的云计算的安全防护需要进行了详细分析，分别从访问控制机制、数据的完整性校验和数据的机密性保护三个方面进行了深入研究，并提出了相应的算法。主要的研究内容包括：

(1)一种基于统一身份认证的双向认证算法。以PGP安全信任模型为参考，设计了安全高效的服务器间认证算法及用户跨域身份认证算法，实现了用户单点登录的云平台安全访问机制；

(2)一种支持用户验证的可信云存储方案。基于秘密共享原理，对原始数据进行分块和多副本冗余存储，通过设计合理的挑战—响应机制，使得用户能够实时发起针对云端数据的完整性校验，并能对出错的数据块进行定位，最后，通过对云滴水印的产生、嵌入和监测，可以实施对数据被非法访问进行问责处理；

(3)一种改进的基于MDS构造矩阵的秘密信息分散存储方案。依靠MDS矩阵的“半可恢复性”，设计了秘密分散和秘密恢复算法，以及更为高效的MDS矩阵构造算法，利用多服务器的协作共同控制秘密信息的访问，使得用户仅在合法获取足够数据分块的前提下才能恢复秘密信息。

**关键词**：云计算；数据防护；身份认证

# Abstract

Over the past decade, cloud computing has become one of the hottest research topics in the field of information in today's world, and is also gradually becoming the technology focus of most concern in academia and industry. The completely new service mode and data sharing mode which are proposed by cloud computing, have injected a powerful force for the development of the Internet economy and the global IT industry, with the advantages of economic, high efficiency and strong expansibility, instructed the development of IT to the trend of specialization, scale and intensive, and become another great change.

But on the other hand, the information security issues involved in cloud computing have also gradually revealed, and become an obstacle in the process of its development and popularization. Therefore, how to provide users with a full range of security protection of data stored in the cloud is a serious problem to be solved at present.

This paper analyzes the security needs of the current cloud computing in detail, studies in-depth from the three aspects of access control mechanism, data integrity check and data confidentiality protection, and proposes the corresponding algorithm. The main research content include:

(1) A mutual authentication algorithm based on a unified identity authentication. Take PGP security trust model as a reference, design a secure and efficient authentication algorithm for servers and a cross domain identity authentication algorithm for users, with achiving a secure access mechanism for a user to log on to a single point.

(2) A trusted cloud storage solution supporting user authentication. Based on the principle of secret sharing, the original data is divided into blocks and multiple copies of redundant storage. Through the design of a reasonable challenge - response mechanism, so that the users can initiate the integrity check of the cloud data in

real-time, and can also locate the error data block. Finally, based on the cloud watermark generation, embedding and monitoring, users can be processed accountability by unauthorized access to data.

(3) An improved secret information distributed storage scheme based on MDS structure matrix. According to the Semi-recoverability of the MDS matrix, we design the secret dispersion /recovery algorithms, and much more efficient MDS Matrix construction algorithms. Using multiple server cooperation to control access to secret information. In order to recover the secret information, the user must obtain the enough data legally.

**Keywords**: Cloud Computing; Data Protection; Identity Authentication

# 目 录

# Contents

# 第一章 绪 论

## 1.1 研究背景与意义

近十年来，云计算成为当今世界信息领域的热门研究课题之一，也逐渐成为学术界、产业界最为关注的技术焦点。云计算通过“网络即计算机”的工作模式将大量的存储、计算和软件资源集中云端，用户按需进行动态搜索虚拟化资源，大大提高了资源的利用率，有效节省了总的存储空间，为互联网应用的发展提出了一种全新的模式，是分布式计算、并行计算、网络存储、负载均衡、虚拟化、备份冗余、效用计算等多种计算机和网络技术发展的融合产物[1]。

在云计算环境中，服务商把软件应用、资源和数据统一看作为服务，以网络租用的方式提供给用户，用户可以从当前繁重的基础设施管理和维护工作中解放出来，从而能以更多精力专注于其核心战略性业务的开发，所以说，云计算提出的全新的服务方式变革和数据共享模式的变革，以其经济、高效、扩展性强的优点为互联网经济和全球的IT产业的发展注入了强大的动力，指引IT领域向专业化、规模化、集约化的趋势发展，是大型计算机到客户端-服务器端大转变之后的又一次巨变[2]。

目前，各大IT企业都在发展和部署其云计算架构。全球最大的搜索引擎公司于2007年10月宣布云计划，目前每年投入约16亿美元用于云数据中心建设，其实际产能相当于传统技术投入640亿美元，资本利用率提高40倍。IBM于2007年8月高调推出“蓝云”计算，大力发展应用服务器、存储及管理软件，同年，亚马逊也向开发者开放了“弹性计算机云”服务，软件公司可按需使用其数据中心的处理能力。2008年，VMware开始推出云操作系统、云服务目录构件及云资源审批管理模块等开放式云平台。各国政府也提出了云计算发展战略，美国提出“云优先”发展战略，联邦政府率先推出“一站式云战略”开放门户网站，为云计算应用提供应用平台；英国政府提出“政府云战略”，充分借助云平台模式提出政府工作效率；中国政府也将云计算列入了八大新型重点发展产业之一，已公布北京、上海、杭州、无锡、深圳为云计算发展试点城市，正在投资建设云中心。据权威机构Forrester预测，到2020年全球云计算市场规模将达2410亿美元[3]。

虽然云计算的发展非常迅猛，但从另一方面考虑，云计算所涉及到的信息安全问题也逐渐显现出来，成为其发展和普及过程中的“绊脚石”。近几年来，云计算有关的数据安全问题逐步暴露出来，已严重制约该行业的健康发展。2009年Gartner公司的调查报告指出，至少500位IT企业高管表示，因担心云计算平台在数据存储的安全性和隐私性受到威胁，目前仍倾向于使用其公司内部网络系统。在国内，2012年中国云安全调查结果显示，88%以上的受访者认为目前的云计算没有提出完整的安全防护架构。实际上，近来来不断爆出的云计算安全事故也证明了这种看法。2009年，Google公司发生用户文件外泄事件和Gmial邮箱用户数据丢失事件，造成严重损失，同年，亚马逊公司的“简单存储服务”发生服务器瘫痪。2011年，著名云服务商Dropbox遭受攻击，用户口令被清空和重置，阿里云公司因磁盘错误的维护，导致系统重启而丢失Team Cola公司的数据。2012年，盛大公司云服务也声称因磁盘存储问题故障而导致个别用户数据丢失。

所以，解决云计算面临的各种数据安全问题对于云计算的继续发展和推广具有十分重要的理论价值和现实意义。如果想要各大企业和用户安全使用云平台所提供的数据存储和计算服务，其首要前提是设计严格和全面的数据安全防护架构，分析并解决目前面临的账户劫持、数据非法访问、篡改攻击等一系列的安全问题。目前，军队的信息化建设也逐步通过建设云服务平台来提高其指挥、控制和资源利用效率，其对云计算的安全需求更高。

针对云计算的安全问题，信息安全国际会议RSA2010已将其列为重点解决的课题之一，并专门成立针对云计算安全问题的研讨会。另外，许多企业、标准化组织、研究团体也开始致力于云数据安全的理论研究和产品设计。

## 1.2 云计算及其模型架构

### 1.2.1 云计算的定义及特点

到目前为至，关于云计算的确切定义，学术界并没有一个能被业务广泛接受的标准。中国云计算计算专家咨询委员会副主任刘鹏教授给出的云计算定义是：“云计算是通过网络提供可伸缩的廉价的分布式计算能力”。百度百科给出的定义是：云计算是基于互联网相关服务的增加、使用和交代模式，是通过互联网来

提供动态易扩展的虚拟化计算资源。美国国家标准与研究院给出的定义是：云计算是一种以可配置资源池来按需付费、提供便捷可用资源的网络访问模式，可用资源包括应用软件、服务、网络及存储空间[4]。

在IBM描述的云计算服务是一个可按需进行动态分配和配置的可伸缩服务平台，在物理上表现为存在海量的服务器硬件设备，主要由计算资源、存储网络、网络辅助设备和安全设备组成。其架构如下图 1-1所示：

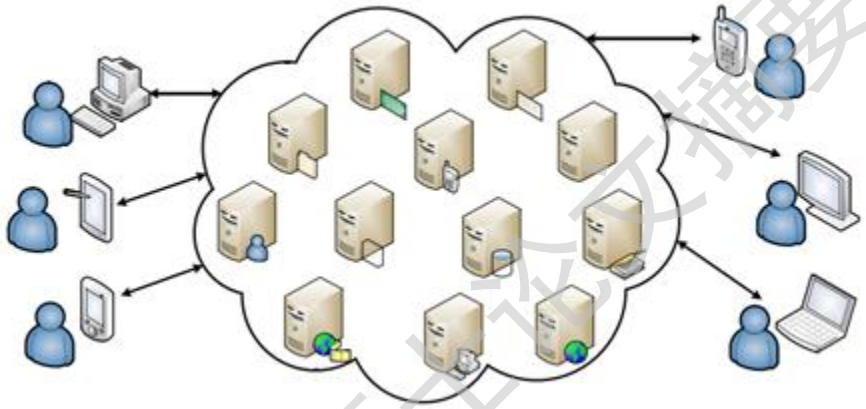

**图 1-1： 云计算服务架构**

美国国家技术与标准研究中心(national institute of standards and technology，NIST)提出云计算应具有的五大关键特征包括可度量服务(measured service)、虚拟资源池(resource pooling)、 高带宽网络(broad network access) 、自助按需服务(on-demand self-service)和高速弹性架构(rapid elasticity)，如下图 1-2所示：

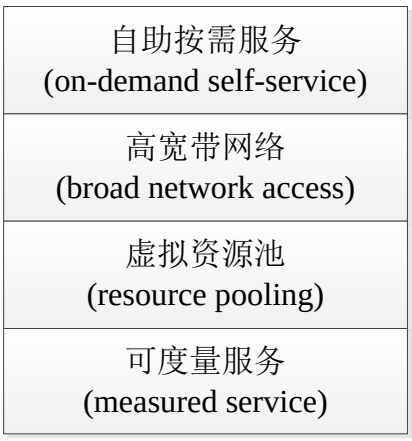


**图 1-2： 云计算关键特征**

自助按需服务是指用户和云计算服务商仅通过少量的交互即可获取所请求的计算和存储资源；高带宽网络接入保证了用户能够得到更高传输速度和更低价格的网络接入服务，并支持客户端平台的扩展性，为不同种类的终端接入提供便利。虚拟资源池是指云计算平台能为用户提供透明可靠的海量资源，用户可按需来申请或分配；高速弹性架构进一步规定了用户在使用和配置云计算资源时自主度和可扩展度；可度量服务为优化资源的使用、提供使用资源时的测量能力、实现自动化控制提供解决方案。

从目前的研究现状来看，云计算应具备以下特点[5]：

(1)超大规模：为保证“云”能够为所有用户提供足够的计算能力，Google公司的云计算已拥有100多万台服务器，AMAZON、IBM、Micorsoft、Yahoo等也已经拥有数十万台云服务器；

(2)虚拟化：云用户能在任意时候通过任意网络终端访问而获得云服务，资源在云中以虚拟方式存在，用户无需关注其具体存储形式；

(3)高可靠性：云计算平台应通过访问认证、计算节点同构、数据多副本容错等种机制保证用户数据的安全及可靠；

(4)通用性：规模的可伸缩以保证用户数量的增长，不针对特定的应用而提供通用性服务；

(5)按需服务：通过其庞大的资源池为用户提供如自来水、电、煤气式的流式资源服务和计费；

(6)价格低廉：通过特殊的容错机制、自动化管理、通用性资源复用、能源利用等为用户提供更为优越的性价比。

### 1.2.2 云计算模型架构及应用模式

云计算的架构根据其体系从底层到高层可分为物理层、核心层、资源架构层、开发平台层和应用层等五个层次[6]。如下图 1-3所示：