Application and Research of Content Management System based on Trusted Computing

X.-Q. Ma[1,2], Yi Huang[1,2*], Bo Lv[3] and Y.-Y. Liu[1]

[1)] Key Laboratory of Machine Vision and Intelligent Information System, Chongqing University of Arts and Sciences, Yongchuan, Chongqing, China (mxq345@sohu.com, lxq1982@163.com)

[2)] Guizhou Academy of Science, Guiyang, Guizhou, China (Corresponding author*: cqhy@21cn.com)

[3)] School of Economics and Management, Guizhou Normal University, Guiyang, Guizhou, China (nvbo@163.com)

***Abstract*—Trusted computing is a hot topic of information security technology research nowadays. It was developed based on computed system. The paper analyzed the present situation about trusted computing. In order to satisfy the requirements of little enterprise and government, we designed and implemented the content management system (CMS) based on trusted computing. It improves the security of the enterprise information. Its implementation takes a good reference for improving the application of CMS.**

***Keywords*—trusted computing, content management system, information security, access control**

基于可信计算的内容管理系统研究及应用

马新强[1,2]　黄羿[1,2]　吕波[3]　刘友缘[1]

[1)] 重庆文理学院机器视觉与智能信息系统重点实验室，永川，重庆，中国

[2)] 贵州科学院，贵阳，贵州，中国

[3)] 贵州师范大学经济与管理学院，贵阳，贵州，中国

摘　要　可信计算是目前信息安全技术研究的一个热点，它是在计算系统的基础上发展来的。分析了国内外可信计算的研究现状，针对中小企事业单位的需求，提出了基于可信计算的内容管理系统的设计思想，尤其在安全性上的改进，保障了企业各类信息的安全可靠。其实现思路对于提高内容管理系统在企业的应用提供了很好的参考价值。

关键词　可信计算，内容管理系统，信息安全，访问控制

1．引言

当前，在开放而又动态变化的网络、硬件的缺陷、软件的错误和安全的漏洞等客观因素使得病毒、木马、黑客、各种攻击等威胁着信息系统的安全，用户的无知、误操作、缺乏安全意识等主观因素使得信息安全的形势更为严峻。为了抵抗网络上的各种威胁，软件中还得集成许多安全功能，同时需要多种软件协同工作。于是，分析和监控复杂、开放计算环境的高可信计算研究受到高度的重视。

重庆市自然科学基金项目支持（资助号：cstc2011jjA40011、cstc2012jjA40068、cstc2013jcyjA40053），贵州省科学技术基金项目支持（资助号：20112205、20102104、20122185），重庆市教委科学技术研究项目支持（资助号：KJ121219、KJ131218），永川区自然科学基金（重点）项目（资助号：Ycstc,2013nb8001）。

2．可信计算及其关键技术

可信计算是当前国内外学术研究热点，而信息安全是可信计算的一个重要内容[1]。作为信息安全的核心，访问控制是计算机系统中的一种安全机制，它确保对系统资源的访问按照安全策略来进行[2]。访问控制是通过某种途径显式地准许或限制主体对客体访问能力及范围的一种方法。由于计算机技术及网络的发展，访问控制的研究发展很快，提出许多访问控制模型[3]。

2.1　可信计算

ISO / IEC15408 标准定义为：一个可信的组件、操作或过程的行为在任意操作条件下是可预测的，并能很好地抵抗应用程序软件、病毒以及一定物理干扰所造成的破坏

[4]。它强调行为的可预测性，能抵抗各种破坏，达到预期的目标。武汉大学张焕国教授等认为可信是指：计算机系统所提供的服务是可靠的、可用的、信息和行为上是安全的。相对应的可信计算平台是能够提供可信计算服务的计算机软硬件实体，它能够提供系统的可靠性、可用性、信息和行为的安全性[5]。

学术界把可信计算（Dependable Computing）定义为“系统提供可信赖的计算服务的能力，而这种可信赖性是可以验证的[6]”。即你必须用某种方法来验证你的系统是可信赖的。这就困难了。我们知道，法律对于人有所谓“无罪认定原则”，就是说，除非有证据证明某人有罪，否则他就是无罪的。而对于可信系统，我们执行的是“有错认定原则”。那就是说，用户可以对系统设计者和制造者说，除非你有足够的证据证明你的系统是可信的，否则我就认为你的系统是不可信的。例如，对于一个软件，如果开发者没有足够的理由说明它是正确的，用户就认为它是有错误的。这个要求对系统设计者和制造者来说，是个难题。需要可信计算技术来提供。而且，可信性必须成为可以衡量和验证的性能。我们简单的定义可信计算为：计算系统应计算准确、功能正确而完备、具有容错能力与可抵御外部干扰；可信≈可靠+安全。主要体现系统的“可信性”，计算系统的“可信”是一个目标，“可信计算”研究是一个系统工程[7]。

2.2 可信平台体系结构

可信平台是以可信平台模块（Trusted Platform Module, TPM）为核心，把 CPU、操作系统、应用软件和网络基础设备融合为一体的完整体系结构。一个典型的 PC 平台上的体系结构主要可以分为三层：可信平台模块（Trusted Platform Module, TPM）、可信软件栈（Trusted Software Stack, TSS）和应用软件，如图 1 所示。TSS 是对可信计算平台提供支持的软件，它的设计目标是对使用 TPM 功能的应用程序提供一个唯一入口，并提供对 TPM 的同步访问管理。TSS 平台软件从结构上可以分为四层，自下至上分别为 TPM 驱动程序库(TPM Device Driver Library，TDDL)、可信软件栈核心服务（TSS Core Services, TCS）和可信服务提供者（Trusted Service Provider, TSP），全部运行于用户模式和运行于内核态的 TPM 驱动程序(TPM Device Driver，TDD)。

3．内容管理系统及其安全性需求分析

3.1 内容管理系统

企业内容管理(Enterprise Content Management System, ECMS)系统是继财务管理系统、ERP 系统之后企业管理系统、集成制造系统方面的又一热点领域，目前在国外非常热门，市场发展迅速[8]。目前，分析和监控复杂、开放计算环境下的高可信“内容管理系统（Content Management System，CMS）”研究受到高度的重视[9]。它是企业信息化建设和电子政务的新宠，也是一个相对较新的市场；是一种位于 WEB 前端（Web 服务器）和后端办公系统或流程（内容创作、编辑）之间的软件系统。内容管理解决方案重点解决各种非结构化或半结构化的数字资源的采集、管理、利用、传递和增值，并能有机集成到结构化数据的商业智能环境中，如 OA,CRM 等。内容的创作人员、编辑人员、发布人员使用内容管理系统来提交、修改、审批、发布内容。这里指的“内容”可能包括文件、表格、图片、数据库中的数据甚至视频等一切你想要发布到 Internet、Intranet 以及 Extranet 网站的信息。业界公认的国内比较权威的产品有 TurboCMS 和 TRS。国外从事内容管理软件研发的主要厂商包括 Vignette, Interwoven, Broad Vision, Openmarket，ATG，Allaire, Documentum, Hummingbird 等，这些公司 CM 产品和解决方案专业性很强，大多基于 J2EE 等平台，功能丰富，主要面向企业级用户，是 CM 市场的主要厂商。还有一些更窄的专业厂商提供内容管理某个阶段需要的功能，如 Verity 提供知识检索，Micromedia 提供内容创作平台，Akamai 和 Inkitomi 提供内容分发管理技术等。与此相反，Microsoft, IBM, Oracle 等公司提供通用平台性 CM 解决方案。但是目前 CM 市场仍有很多不完善的地方，尤其是安全性方面急需加强。从企事业单位信息化的观点来看，以下因素导致对内容管理软件的巨大需求：(1) 安全可靠的知识是企业的财富；(2) 信息的及时性和准确性；(3) 企业内外网统一的需求增长。

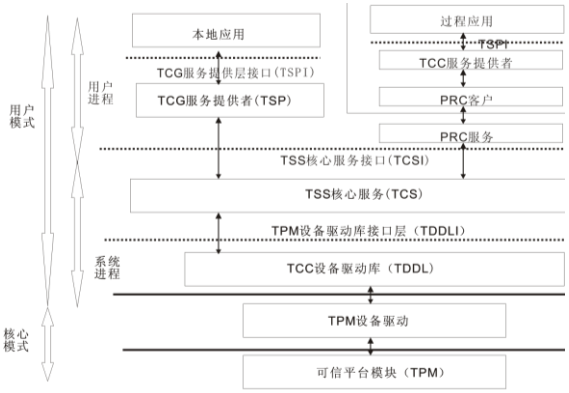

图 1 可信平台体系结构

3.2 安全性分析

国内外内容管理系统主要使用 Windows 2000 或 Unix 系统下的 SQL Server 或 Oracle 作为数据库服务器，投资

成本高昂。虽然 SQL Server 和 Oracle 都已经开发出 B1 级别的安全模块，但是该模块对我国是禁止出口的。同时在涉及国家安全的系统中使用国外产品，存在严重的安全隐患。从国家安全的角度来说，涉及到国家安全的数据库系统，不能使用国外数据库。这是因为长期以来，我国严重缺乏自主版权的数据库管理系统，其中真正能够实现产品化的数据库管理系统更是寥寥无几。

对系统安全的研究无论是在理论上还是在相关实用产品种类较少，开发基本停留在大学等研究机构的实验室内，应用范围也较窄[10]。彭智勇教授等研究了国内外数据库安全研究现状，提出了可信数据库的基本概念[11]。东软集团开发的 Open BASE Secure 和华中科技大学的 DM5 都基本达到 B1 级安全数据库标准。多级安全保密功能是华中科技大学 DM5 数据库管理系统的最大特色。系统采用了三权分立的安全机制，数据库管理员 DBA、数据库安全员 SSO 和数据库审计员 AUDITOR 分别在三个层面负责自主存取控制、强制存取控制和审计，他们各自独立、相互制约，更好地实现了数据的保密性、完整性和可用性[12]，但未见在其上开发内容管理系统的报道。

3.3 安全性需求分析

目前国内外内容管理系统主要是针对大型企业或网站、媒体内容管理设计，价格昂贵、设计复杂、使用困难，不适用于我国中小企事业单位一般的管理用途。这些系统设计了繁复的多种功能，但真正得到广泛应用的功能主要是全文搜索、安全控制、网页内容编排几个功能。不用说国外高端软件昂贵的价格，即使是国产软件，如 TurboCMS，价格在 10 万到 25 万之间；TRS 价格在 15 万左右一个许可，全套 TRS 的总价格在 40 至 50 万元，也超过了我国中小企事业单位的承受能力。因而，研究和开发适用于各类中小企业及各级党政机关、事业单位的安全、可信内容管理信息系统，实现各种非结构化或半结构化的数字资源的采集、管理、存储、保护、搜索、利用、传递和增值，为上述单位的管理提供有效信息和决策支持，具有重要的实际意义。

4．基于可信计算的内容管理系统设计与实现

以可信计算及技术作为核心安全控制平台，采用已申报专利的个性化、智能化文档处理技术，利用具有自主知识产权的基于大型数据库 LogicSQL[13]的 B1 级别安全数据库和实方企业信息搜索工具为核心技术研制适用于各类中小企业的内容管理信息系统 TCCMS。系统开发环境采用 Windows + Apache tomcat + JavaEE + LogicSQL；其中系统抓取工具采用实方软件自主开发的分别针对网页、文件和数据库的抓取 crawler；还有实方授权下的自动推取程序 pusher；二次开发下的计算环境恢复程序中间件工具。 系统结构如图 2 所示。

利用以下五种改进策略，保证企业各类信息的安全可靠，有效地实现企业信息的自动提取、实时备份等。

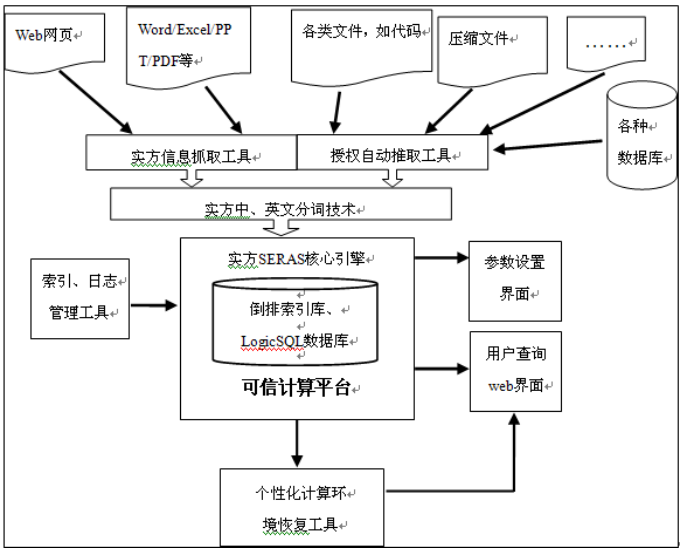

图 2 可信的内容管理系统结构

4.1 可信计算技术

以可信平台为基础，把信任根和信任链是作为内容管理可信平台的关键技术。一个可信计算机系统由可信根、可信硬件平台、可信操作系统、可信数据库系统和可信应用系统组成[1]。信任链是通过构建一个信任根，从信任根开始到硬件平台、到操作系统、数据库系统、再到应用，一级认证一级，一级信任一级，从而把这种信任扩展到整个计算机系统。

4.2 企业信息检索安全管理

Internet search 的信息是公开的，而 Enterprise search 所涉及的信息是企业内部的，只有具备相应的权限才能搜索和查看相关的信息，所以必须加强企业信息搜索的安全。所有数据来源包括：文件、电子邮件、各种文档、 网页及其它形式分散存在于个人计算机中的信息，都能被访问控制列表（ACLS）所保护。本系统采取不同的方式保证企业的信息安全：

(1) 搜索管理员能直接指定授权角色，通常这种情况是在企业搜索管理界面下通过简易目录访问协议 LDAP 用户和组进行授权列表中应用，只有具有权限的用户和角色才能允许访问和搜索所需的文档。

(2) 具有权限的用户和角色才能允许访问，采用自动“推取”的服务程序方式，避免在抓取方式下对于敏感、

私有信息的访问而产生的安全问题。

(3) 在信息自动实施备份的过程中，采用自动“推取”的方式，即把本地个人 PC 上的各种信息运用自动服务程序实时向服务器（实方数据库）传送备份。这里重点解决的难点是能够设计出自动传送服务程序，采取这种方式实时备份而不采取爬行方式抓取信息。

4.3 自动“推取”的服务程序方式

对服务器上的采取“抓取”方式继承与改进，采用自动“推取”的上传服务程序，自动“推取”的服务程序方式具有的特点：①安全性，保护企业网络中其它的计算机的安全，避免了抓取方式下对敏感信息产生的安全问题。②主动性，由计算机终端自动向上传送信息。③效率高，干扰少；避免抓取下的各种文件信息重复搜索遍历，节约时间，还避免其他抓取的干扰。④主动存储，有利于恢复备份。

4.4 高效的信息检索功能

系统采用先进的倒排索引技术和 TOP—K 搜索查询算法实现最快速的查询，并把最准确的结果返回到用户界面；尤其在全文查询搜索方面：提供搜索导航、关键字搜索、高级搜索、分类搜索；用户可以自定义常用搜索条目，自动保存最近的 10 条搜索历史。

4.5 个性化的计算环境备份与恢复

在实方搜索引擎的基础上，采用个性化、智能化文档处理技术，对不同用户的计算机产生的计算坏境进行二次移植和开发。根据不同用户的分类能够进行不同计算环境的备份和恢复。

5．结束语

可信计算机与普通计算机相比，安全性大大提高，但可信计算机也不是百分之百安全。目前可信程度的具体评估目前也没有统一的标准，尤其是对软件系统的可信性，还存在较多难点。对可信内容管理系统构建与分析，设计一个基于可信计算的内容管理系统综合试验平台，形成“基于可信计算的内容管理系统”应用范例，为进一步研究可信计算环境提供参考。通过在上海实方软件公司、吉林省消防队和贵州省公安厅的典型应用，再次证明了其在一些中小型企事业单位应用的实际价值。

参考文献(References)

[1] H.-G. Zhang and Y.-X. He, Research and development of our trusted computing. Communication of the CCF, 2010,6(2):8-19.

[2] Y.-F. Li and C.-X. Shen, A new operating system security model. Chinese Science (E Volume), Information Science, 2009, 36(4): 347-356.

[3] X.-Q. Ma and Y. Huang, Trusted computing model based on lattice, Journal of Communications, 2010, 31(8A):105-110.

[4] Pearson S, Trusted Computing Platform, the Next Security Solution. Bristol UK: HP Laboratories, 2002.

[5] X.-Q. Ma and D.-N. Li, Study on the multiple security models of content management system, Journal of Micro Computer Information, 2010,26(4-3):102-103(111).

[6] C.-X. Shen, H.-G. Zhang, and D.-G. Feng, et al, Overview of Information Security Chinese Science (E Volume), Information Science, 2007, 37(1):129-150.

[7] X.-Q. Ma, Y. Huang, D.-N. Li, Study on development of Trusted Computing, Journal of Computer Application, 2009, 29(4): 920-923.

[8] Content Management System, http://baike.baidu.com/view/857578.htm, 2014-1-31.

[9] Avizienis A, Laprie J C, Randell B, et a1. Basic Concepts and Taxonomy of Dependable and Secure Computing, IEEE Transaction on Dependable and Secure Computing, 2004, 1(1):11-33.

[10] M. Zhang, Z. XU, and D.-G. Feng, Database security, 2009.8, Beijing: Science in press.

[11] Z.-Y. Peng, Trusted database development concept, and challenges, Journal of Computer Application, 2008, 28(11): 2741-2744.

[12] Y. Huang, X.-Q. Ma, and X.-Y. Zhou, An access control model based on Trusted Computing, Journal of Chongqing University of Arts and Sciences, 2010, 29(3):54-57.

[13] L.-Y. Yuan. The Documentation of LogicSQL, Alberta University, Canada. 2013.12.