USN 安全研究与设计*)

韩德志[1,2] 耿红琴[3] 谢长生[2] 胡玉平[2]
(暨南大学计算机系 广州510632)[1] (华中科技大学计算机学院 武汉430074)[2]
(河南驻马店师范专科学校 河南驻马店453000)

Research on the Security of the United Storage Network Converging the NAS and SAN

HAN DE-Zhi GENG Hong-Qin XIE Chang-Sheng HU Yu-Ping
(department of computer, Jinan University, Guangzhou 510632)[1]
(School of Computer Science, Huazhong University of Science and Technology, Wuhan 430074)[2]
(Zhumadian Teacher' College, Henan Zhumadian 453000)[3]

Abstract With the development of networked storage, the USN converging NAS and SAN appears, USN has many virtues: high performance, low lost and so on, but its security becomes more complex. Considering the situation, with making a deeply research on the performance, characteristic, and the architecture of the USN, the paper designs and implements a USN-based security algorithm, in which, the encryption and decryption are performed on the client, and the USN server only provides identity authentication for user and the integrity verification for data. The test result indicates the algorithm can prevent many kinds invalid attacks, and make few influences on performance of the USN.

Keywords NAS(Network Attached Storage), SAN(Storage Area Network), USN(United Storage Network), HMAC(hashed message authentication code)

1 引言

随着网络存储的发展,出现了两种非常重要的网络存储技术:NAS(Network Attached Storage)和 SAN(Storage Area Network)[1,2]。按照存储网络工业协会(SNIA)的定义:NAS 是可以直接联到网络上向用户提供文件级服务的存储设备,而 SAN 是一种利用 Fibre Channel 等互联协议连接起来的可以在服务器和存储系统之间直接传送数据的网络。NAS 是一种存储设备,有其自己简化的实时操作系统,它将硬件和软件有效地集成在一起,用以提供文件服务。而 SAN 是一种体系结构,它采用不同的技术(目前采用最多的是 FC(Fibre Channel)技术)构建的与企业原有网络不同的一个专用于存储的网络,存储设备和 SAN 中的应用服务器之间采用的是块 I/O 的方式进行数据交换。NAS 和 SAN 是两种不同的技术,它们应用在不同的场合下,企业由于不同的应用需求,可能同时拥有两种存储结构,但由于目前两者的建构技术、管理工具等都不同,企业要求构建和管理两种完全不同的网络,这大大增加了企业的总拥有成本,同时也不能很好地做到存储资源的整合等。在这种情况下出现了融合 NAS 和 SAN 的统一存储网络(United Storage Network,简称 USN),在建构 USN 时,又有两种建构技术:一种是基于 NAS 技术的 USN,一种是基于 SAN 技术的 USN。不论是基于 NAS 技术的 USN,还是基于 SAN 技术的 USN,其安全问题变得更为复杂,因为它们一方面要面对文件 I/O 请求的客户,另一方面要面对块 I/O 请求的客户。针对这种情况,本文通过对以 NAS 为主的 USN 安全技术进行深入研究的基础上,设计并实现了一种可靠、实用、代价低的安全算法。

2 基于 NAS 的统一存储网络

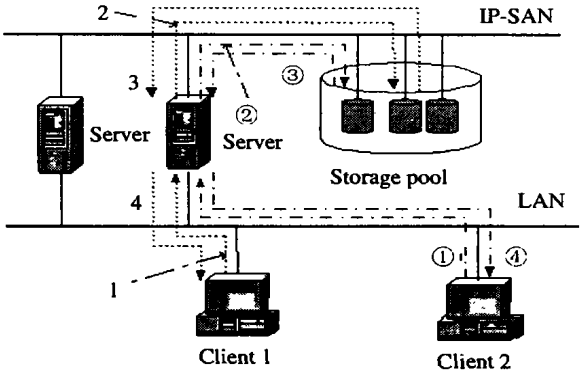
传统的 NAS 是可以直接附网的存储设备,其上含有操作系统(文件系统)和存储子系统,两者的关系是紧耦合的关系,存储子系统从属于单个的 NAS 设备,其扩展性是有限的。而基于 NAS 技术的统一存储网络(USN)融合了 NAS 和 SAN 的结构特点,首先以 NAS 的方式和外部网络相连,改变了传统的 NAS 设备中操作系统(文件系统)和存储子系统紧耦合的关系,将操作系统(文件系统)和存储子系统分开,其文件系统功能部分单独构成一个 USN server,用以保存元数据、完成文件系统的功能;其存储子系统采用外置式的 SAN 结构,并利用存储虚拟化技术将 SAN 中的各种存储设备虚拟、整合为一个统一的存储池,如图1所示。采用这种结构,系统可以同时提供 File I/O 和 Block I/O 两种类型的服务,全部的存储空间逻辑上也可以相应划分为提供文件服务的 File space 和提供数据块服务的 Block space 两部分。图1中,Client1、Client2分别为发出 Block I/O 请求和 File I/O 请求的客户端,server 为该存储网络的服务器,图2为整个系统大致的软件结构示意图,其中 server 上软件的开发是构建整个系统的关键。发出 Block I/O 的客户端 Client 1必须能够支持 iSCSI 功能,而发出 File I/O 请求的客户端 Client 2不用加装任何软件。

存储网络提供数据块服务时,采用的是 iSCSI 技术,此时 Client 上软件的层次结构如图2所示(Client 1)。具体的数据

*)本文受国家自然科学基金项目(60173043)资助。韩德杰、胡玉平 博士生,研究方向为基于网络的存储系统,基于 IP 的存储区域网。谢长生 教授,博士生导师。耿红琴 教授。

读写流程为(如图1):1)Client 1上的应用程序发出的块I/O命令(SCSI命令)经iSCSI设备驱动层和TCP/IP协议栈之后,封装成IP数据包,在IP网络上传输;2)封装后的SCSI命令达到server之后,经解封装,恢复成封装前的SCSI命令,server利用这些SCSI命令对存储设备发出块I/O读写请求;3)所需的数据块由存储设备返回server;4)数据块经server中的TCP/IP协议栈封装后在IP网络上传输返回Client 1,经其解封并由文件系统组合成文件。在这种工作方式中,文件系统的的各部分功能由Client和server分别完成,其中Client完成File I/O请求和Block I/O请求的转化,文件与数据块的映射等功能,而目录操作则必须由server完成,以维护整个存储池中Block space的一个全局、统一的目录视图,构成一个单一的存储系统映像,因此Block space上所有文件的元数据应该同时保存在Client 1和相应的serve上。

当存储网络提供文件服务时,其数据读写过程是:① Client 2向server发出文件读写请求(其工作方式和传统的NAS相同);② 文件读写请求经过server上的文件系统转化为Block I/O请求,从而直接操作SAN结构中的存储设备;③ 存储设备将相应的数据块返回给server;④ 数据块经过USN server上的文件系统组合成文件提供给Client 2。在这种工作方式中,由于采用的是以NAS的方式和网络相连,其工作方式和传统的NAS设备相同,因而具有了传统NAS的大多数优点:可以实现异构的文件访问和共享,又因为是通过IP网络访问,因而可以充分利用现有的网络环境,具有广泛的连通性并适应复杂的网络环境。由于其存储子系统部分是采用SAN结构(甚至采用FC等高速互联技术),因而又具有SAN结构的优点:高可扩展性、高可用性,以及存储资源的集中和统一管理等。



注:
1、2、3、4是Client 1向系统提出Block I/O请求的数据流程
①、②、③、④是Client 2向系统提出File I/O请求的数据流程

图1 基于NAS的USN结构示意图

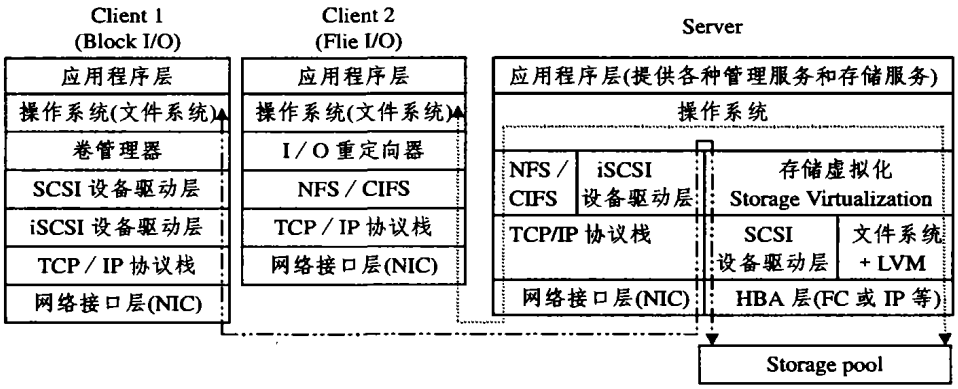


图2 基于NAS的USN系统软件结构示意图

3 USN安全系统的设计

在统一存储网络中,存储和传输的数据可根据用户的需要采用加密方式,也可采用明文方式。如果是加密数据,数据的加密和解密只在客户端进行,服务器不存在任何解密信息。用户在向USN存储系统写数据时,不论是加密数据还是明文数据,USN服务器都要对用户身份进行认证和对数据的完整性进行检查,这样既可以防止非法用户对USN的存储池进行写操作,又可以防止非法修改的数据写入USN存储池。在读数据时,USN服务器只把块空间的数据传给块请求的客户,把文件空间的数据传给文件请求的用户,数据的合法性和完整性由客户端认证。这样把加密解密数据的负担分给客户端,USN服务器只负责写数据的用户身份认证和所写数据完整性的检查,这给USN服务器增加很少的额外负担。

USN安全系统设计的关键是客户端的安全模块设计,客户端的安全模块要面对块I/O请求和文件I/O请求,对于块I/O请求的处理比较简单,对于文件I/O请求就要复杂得多。因为在写文件时,安全模块必须要对文件I/O请求客户端所写文件按要求分成一个个的文件块,然后对文件块进行加密和计算其HMAC[3],最后发送给USN存储池。

3.1 安全系统的设计目标

我们所设计的安全算法为安全系统提供几个重要的特性:

(1)既可为block I/O客户提供端到端的加密,也可为file I/O客户提供端到端的加密。

(2)USN服务器中不保存任何解密信息,数据的解密工作只在客户端进行。

(3)保证加密数据和非加密数据的完整性是安全系统的重要特性。用户从USN中读取数据时,必须能确信他所收到的信息是他自己所存放的,USN中非法修改的数据或非法添加的数据都可以被检测出来。

(4)安全系统提供像AFS[6]这样的访问控制清单,使每个安全数据文件能同时供多个的用户或至少一个Unix组用户共享,注意:多个用户或Unix组用户是共享同一私钥的。

(5)安全系统具有高性能和可扩展性。USN采用的是文件系统和存储池相分离,并且对I/O进行优化(零拷贝)[7]。

(6)我们通过使用用户口令来对用户的身份进行论证。为了防止时间过长,口令以及用户 HMAC 密钥被泄露,该安全算法通过一个时间戳来控制用户口令和 HMAC 的生命周期。

(7)客户端安全模块能对文件 I/O 请求客户端所要写入存储池的数据文件,按用户的要求进行分块,然后进行加密并计算 HMAC,或只计算其 HMAC,最后将文件块和对应的 HMAC 一起发给 USN 服务器。

3.2 基本的安全机制

基本的安全机制是用户可在客户端加密数据,并且给服务器足够的信息认证写数据者的身份和写数据的完整性,同时给读数据者足够的信息来认证端到端数据的完整性。

安全文件系统使用几个标准加密工具。客户端使用象 RC5[4]或 Blowfish[4]来加密数据,加密数据直到它在客户端解密后,它才是可读的。加密数据的密钥又被用户公钥加密,只有合适私钥的用户才能使用这些信息。

客户端和 USN 服务器共享 HMAC(hashed message authentication code)密钥,客户端在发送加密数据或明文数据时,首先用 HMAC key 计算加密数据或明文数据的 HMAC,然后把 HMAC 同数据块一起发向 USN 服务器,USN 服务器用共享 HMAC key 计算加密数据块的 HMAC,然后进行比较,由此认证发送者的身份和发送数据的完整性。用户所写的数据和该数据的 HMAC 一起被存在 USN 的存储池中,当用户读数据时,要读的数据和对应的 HMAC 一起被发送到客户端,由用户进行完整性校验。

另外,为了防止重发攻击,用户必须提供口令,由 USN 服务器进行其身份论证,只有认证通过,用户才能对 USN 服务器进行读写操作。

3.3 安全系统的核心

我们在设计 iNAS(基于 iSCSI 的附网存储服务器)的安全系统的基础上[7],对其安全算法进行了改进,使其适应统一存储网络环境。虽然我们也采用与文[7]类似的安全算法,即使用四种基本的数据结构:数据体、文件体、密钥体和证书体。但我们在证书体中增加了用户口令项,即通过用户口令来认证用户的身份,这可以防止非法用户对 USN 的重发攻击。在密钥体中,我们用时间戳用来记录用户修改密钥体的时间,用来跟踪用户对密钥体的修改,这样减少了文[5]中数字签名的复杂性。在客户端安全模块中,我们也对文[7]文件分块数据结构进行了改进,使安全系统对整个客户端系统性能影响更小。

4 USN 安全算法

我们 USN 安全系统的设计目标:①对访问 USN 存储池的用户身份通过口令提供认证,只有认证通过的用户,才能对存储池进行读写操作;②USN 服务器要对用户写入存储池中的每一数据的完整性提供检验,只有检验通过数据才能写入存储池;③提供两种算法供用户选择,第一种算法是加密存储,第二种算法是明文存储。这两种算法安全性不一样,并对系统的速度等影响也不一样。

第一种算法安全性高,但会影响系统的速度。用户在向 USN 存池写数据时,用户首先要提供口令给 USN 服务器进行身份认证,身份认证通过后,用户用自己 RC5对称密钥对数据块或文件块加密,然后用自己的 HMAC 密钥对整个加密的数据块或文件块进行哈希运算得 HMAC,最后把加密数据块或文件块同其对应的 HMAC 一起发给 USN 服务器,USN 服务器对接到的加密数据块或文件块计算得其 HMAC,并将计算得到的 HMAC 与用户发送过来的 HMAC 进行比较,当二者一致时,USN 服务器才将用户的加密数据块或文件块和对应的 HMAC 一起存入 USN 存储池中。当用户读数据时,用户的身份认证通过后,USN 服务器将用户要读的加密数据块或文件块及对应的 HMAC 一起送给用户,由用户在客户端进行数据块或文件块的完整性检验及解密。

第二种算法安全性相对低一些,对系统的速度影响更小。在读写操作时必须用户通过口令进行身份认证,通过 HMAC 进行读写数据的完整性校验,数据块或文件块是以明文方式传送和存储的。

5 安全系统与文件系统的集成

安全系统很容易集成到 USN 服务器端和客户端的文件系统中。其客户端安全模块,有权限的用户可从 USN 服务器中下载并安装即可。用户可选择对写入 USN 中的数据加密传送和存储,也可选择数据以明文的形式传送和存储(保证所写数据的完整性)。块 I/O 请求客户的读写过程和文件 I/O 请求客户的读写过程,除了都要通过口令认证身份以外,有很大的区别。

在加密情况下,块 I/O 请求客户在写时,其安全模块直接对块数据加密,然后由 iSCSI 和 TCP/IP 发给 USN 服务器,USN 服务器端的安全模块进行完整性的检查和存储;块 I/O 请求用户在读数据时,USN 服务器把用户要读的加密数据块和对应的 HMAC 一起发向用户;文件 I/O 请求客户在写数据时,首先用客户端的安全模块把要写的数据文件分成规定块大小的块文件加密,然后把加密的数据块文件和对应的 HMAC 及其它信息组成安全数据体文件,最后把安全体文件由 NFS/CIFS 和 TCP/IP 一起发送给 USN,USN 服务器端的安全模块进行完整性检查和存储。文件 I/O 请求客户在读数据时,USN 服务器的安全模块把客户端要读文件的安全数据体文件由 NFS/CIFS 和 TCP/IP 发给客户端,客户端的安全模块将各个加密数据块文件进行完整性检查,并解密后将各个文件块组成文件送到用户。

对于非加密的情况,客户端安全模块只负责生成 HMAC,USN 服务器端的安全模块只根据 HMAC 校验数据块的完整性。

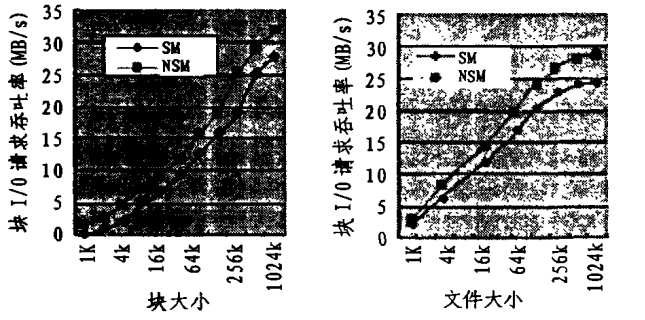
6 实验评估

我们主要是通过实验,评估安全系统对整个 USN 系统性能的影响,我们测试了两个主要参数:平均响应时间和吞吐率。平均响应时间是指,从客户端发 I/O 请求到 USN 完成 I/O 请求的平均时间;吞吐率是指,存储服务器每单位时间所完成的 I/O 请求数。平均响应时间和吞吐率可以从客户端测试,也可从服务器测试,在我们的试验中,我们是从存储服务器端测试的。

为了评估安全系统对整个系统性能的影响,我们精心设计了两组实验。一组是对 block I/O 客户的实验,另一组是对 file I/O 客户的实验。在 blockI/O 客户的实验中,客户端和 USN 服务器端的 iSCSI HBA,是利用普通 NIC,通过编译和运行 Intel 公司开发的 Linux iSCSI 数据包仿真的。

在实验中,我们用 IOmeter 测试程序,测试块客户请求和文件客户请求的平均响应时间及 I/O 吞吐率,共分四次进行测试。第一次测试 iSCSI 块客户和 USN 服务器不加安全模块

的情况，第二次测试 iSCSI 块客户和 USN 服务器加安全模块的情况，第三次测试文件客户和 USN 服务器不加安全模块的情况，第四次测试文件客户和 USN 服务器加安全模块的情况。其块 I/O 吞吐率和文件 I/O 吞吐率显示如图7所示。从显示的结果看，块 I/O 客户加载安全模块并选择加密存储时比不加载安全模块慢14%-22%；而文件 I/O 客户加载安全模块并选择加密存储时，比不加载安全模块慢15%-25%。在这两种情况下，如加载安全模块，但选择非加密存储时，对整个系统的性能影响更小。



(a) 块 I/O 请求吞吐率 (MB/s)　(b) 文件 I/O 请求吞吐率 (MB/s)

注：图中 SM 表示加安全模块，NSM 表示没有加安全模块。

图7　USN 顺序访问性能

结论　目前，在国际上，NAS 和 SAN 技术是两种比较成熟的技术，而寻求 NAS 技术和 SAN 技术融合构建 USN 则是一种全新的技术。我们在国家自然科学基金项目（统一存储网络的设计、建构和实验研究，编号：69873017）和高等学校骨干教师资助项目（存储区域网智能调度和管理技术研究）的资助下，用211工程提供的设备构造了一个简单的 USN。在此基础下，我们结合我们所设计的 iNAS 安全系统[7]，并基于 NAS 的 USN 的特点，进一步设计出了一套 USN 安全算法。通过实验验证，该安全算法对整个 USN 系统性能影响不大，与文[5]相比该系统通过用户口令认证，可防止拒绝服务攻击，并且既可面向块 I/O 请求，又可面向文件 I/O 请求。与文[7]相比，我们对客户端文件 I/O 请求安全模块进行了改进，使安全系统对文件 I/O 请求客户端的性能影响更小；在 USN 服务器端，我们通过增加 HMAC 认证和 I/O 写并行算法模块，使安全系统对 USN 服务器性能影响更小。

参考文献

1　IBM redbook：IP Storage Networking：IBM NAS & iSCSI Solutions

2　Brocade whitepaper. comparing the Storage Area Network and Network Attached Storage

3　Krawczyk H, Bellare M, Canetti R. HMAC: Keyed-Hashing for Message Authentication. IETF Network working Group RFC2104, Feb. 1997

4　Reid J. Plugging the Hole on Host-Based Authentication. Computers and Security, 1996. 661～671

5　Miller E, Long D, Freeman W, Reed B. Strong Security for Distributed File Systems. IEEE Micro, 2001, 20(1): 34～40

6　Thadani M, khalidi Y A. An Efficient Zero-Copy I/O, Framework for UNIX. SUN Microsystems Laboratories, Inc

7　韩德志，等. 基于 iSCSI 协议的附网存储安全系统的研究与设计. 小型微型机系统（已录用）

（上接第102页）

7　Santoro N, Widmayer P. Distributed Function Evaluation in the Presence of Transmission Faults. In: Proc. Intl. Symposium on Algorithm, SIGAL'90, Lecture Notes in Computer Science 450, Springer, Berlin, 1990. 358～369

8　Najjar W, Gaudiot J L. Network Resilience: A Measure of Network Fault Tolerance. IEEE Transactions on Computers, 1990, 39(2): 174～181

9　Chen M S, Shin K G. Depth-First Approach for Fault-Tolerant Routing in Hypercube Multicomputers. IEEE Transactions on parallel and Distributed Systems, 1990, 1(2): 152～159

10　王国军，陈建二，张祖平. 局部子立方体连通的超立方体网络容错路由算法和概率分析研究：[湖南省自然科学基金报告]. 2002

11　Zhang Junying, Xu Jin, Bao Zheng. Tolerantly Linear Separability of Boolean Functions and its Numbering. In: 1996 Intl. Conf. on Signal Processing Proc. 1996. 1433～1436

12　Ould-Khaoua M, Sarbazi-Azad H. An Analytical Model of Adaptive Wormhole Routing in Hypercubes in the Presence of Hot Spot Traffic. IEEE Transactions on Parallel and Distributed Systems, 2001, 12(3): 283～292

13　Sarbazi-Azad H, Ould-Khaoua M, Mackenzie L M. An Analytical Model of Fully-Adaptive Wormhole-Routed k-Ary n-Cubes in the Presence of Hot Spot Traffic. In: 14th Intl. Parallel and Distributed Processing Symposium (IPDPS'00), 2000. 605～610

14　dandamudi S P, Eager D L. Hot-Spot Contention in Binary Hypercube Networks. IEEE Transactions on Computers, 1992, 41(2): 239～244

15　Dally W J, Aoke H. Deadlock-free Adaptive Routing in Multicomputer networks Using Virtual Channel. IEEE Transactions on parallel and Distributed Systems, 1993, 4: 466～475

16　Abraham S, Padmanabham K. Performance of the direct Binary n-cube network for multiprocessors. IEEE Transactions on Computers, 1989 (7): 1000～1011

17　Ozguc B, Isler V, Aykanat C. Subdivision of 3D Space Based on the Graph Partitioning for Parallel Ray Tracing. In: Proc. of the 2nd Eurographics Workshop on Rendering, Barcelona, 1991

18　MacDonald J D, Booth K S. Heuristics for Ray Tracing Using Space Subdivision. The Visual Computer, 1990, 6(3): 153～166