基于内容安全的分布式网络存储策略

常呈武[1] 王 宇[2]

(西昌卫星发射中心指挥控制中心 西昌615621)[1] (装备指挥技术学院 北京101416)[2]

摘 要 与采用对等网络结构的存储系统相比,采用客户/服务器运行模式的传统网络存储系统有很多不足。从安全的角度来看,客户/服务器模式的网络存储系统往往采用基于连接的安全来保护数据,这种解决手段越来越不适合于保护日趋复杂的分布式网络应用。因此,为了保护用户数据,必须采用基于内容的安全技术,使安全与应用松散耦合,在降低系统的安全风险的同时,尽量保证系统的运行效率和可扩充性、可维护性。本文首先分析了两种不同网络结构下存储系统的特点,接着阐述了基于连接安全和基于内容安全的不同性质。最后,论述如何采用基于内容的安全技术实现分布式网络存储系统的安全。

关键词 基于内容的安全,分布式网络,存储系统,策略

Distributed Network Storage Measures Using Content-Based Security

CANG Cheng-Wu[1] WANG Yu[2]

(Mission Command & Control Center, XSLC, Xichang 615621)[1] (Academy of Equipment Command & Technology, Beijing 101416)[2]

Abstract Compared with peer-to-peer (P2P) mode storage system, traditional client/server mode network storage system has many deficiencies. From the view ofsecurity, the client/server mode network storage system often adopts connection-based security technology to protect data. This kind of resolution is becoming more and more unsuitable for protecting increasingly complex distributed network applications, such as distributed network storage system using P2P mode. So, in order to protect user data, content-based security technology should be used, which can make security mechanism loosely couples with applications, and can ensure running efficiency, expandability and maintainability of the system at the same time reducing security risk. This paper firstly analyzes different features of these two network mode, then introduces different characters between connection-based security and content-based security. At last, it describes how to implement the security measures of distributed network storage system using content-based security technology.

Keywords Content-based security, Distributed network, Storage system, Measures

1 客户/服务器存储与对等网络存储

传统的网络存储系统往往采用客户/服务器这种计算方式,即将用户的要存储的信息集中放置在某些大容量的存储服务器上。采用这种结构的网络存储系统有以下缺憾:

1)当并发访问量增加到一定程度时,存储服务器往往成为降低系统性能的瓶颈;

2)由于所有信息都是集中存放的,因此对存储服务器的安全性和可靠性要求极高。一旦存储服务器出现故障或者遭受攻击破坏,客户的数据会很容易丢失。

1999年,美国柏克莱大学开展了寻找外星生命的SETI@home研究计划,该计划首次采用对等网络[1](Peer to Peer:P2P)技术串联所有参与研究计划者的闲置电脑来执行庞大复杂的运算任务,然后再把结果传到SETI@home总部。这些电脑每天平均发挥的效能超过了全球造价最高的、运算最快的超级电脑。

和客户/服务器结构不同,在一个对等网络里,没有专门的服务,计算机之间也没有级别之分。所有的计算机都是平等的。每一台计算机都可作为它自己的服务器,由各个用户自己决定在网络上共享他或她计算机上的哪些数据和不共享哪些数据。采用对等网络结构的网络存储系统具有以下特点:

1)由于没有集中化的存储服务器,用户的数据可能存储在网络上的任何地方,甚至可以备份在多个地方,提高了存储系统的健壮性;

2)对等网络降低了网络的集中化程度,减轻了用户对集中式服务器的依赖程度,进一步确保了存储系统的可用性。

从以上分析可以看出,采用对等网络结构的分布式网络存储系统是必然的发展趋势。此外,文件共享、协同合作、分布式计算、智能代理等商业应用同样会倾向于使用对等网络。

2 基于连接的安全与基于内容的安全

基于连接的安全技术[2,3]是指通过保护面向连接的数据通信来保护网络应用的安全。很多传统的网络存储系统,使用了HTTPS、SSL(Secure Socket

常呈武 高级工程师,主研方向:信息安全,分布式处理;王 宇 副教授。

Layer)和 TLS(Transaction Layer Security)等基于连接的安全协议,借助于 VPN(Virtual Private Network)思想在客户终端与存储服务器之间构建出安全的通信隧道。这种做法存在以下问题:

1)客户与服务器之间必须建立直接的连接。如果受保护的网络应用要通过多个中间应用才能提供增值的服务,就要建立多个安全隧道连接(如多个 SSL 连接),这种做法不但增加了连接节点的安全风险,也极大地增加了认证管理的负担。

2)所有的通信内容都要被加密。在很多网络应用中,并不是所有传输数据都需要被加密,有的可能只需要采用完整性保护。不必要地加密所有内容会增加处理上的开销。

3)不易于实现非连接通信的安全。由于基于连接的安全解决的是连接通信的安全问题,因此非连接的网络通信,如 UDP 数据包,必须通过中间的协议转换网关,才能采用基于连接的安全技术实现安全通信。

基于内容的安全技术采用了"安全边界最小化"的原则,它直接保护数据内容本身,并具有以下特点:

1)数据的安全性更高。由于敏感数据在生成时就被加密,并且只有在被使用时解密,因此基于内容的安全不但保护了传输中的数据,也保护了存储在本地或者服务器上的数据。

2)适合于保护对等网络结构下的应用。由于对等网络结构下的应用没有集中式的服务器,数据可能分布在网络的任何一台主机上,如果采用基于连接的安全技术,每台主机都必须受到保护,才能确保存储在其中的数据是安全可靠的。显然,保护数据内容本身比保护数据通信更简单、更安全、更可靠。

3)数据可根据其敏感程度采用不同强度的保护措施。

3 基于内容安全的分布式网络存储策略

分布式的对等网络存储系统是将用户的数据存储在网络上的不同位置,不需要集中的存储服务器。数据的存放位置可通过 LDAP 服务器或者 URL 给出。根据以上分析,基于内容的安全要比基于连接的安全更适合于保护分布式的网络存储系统,但它需要分布式的公开密钥基础设施和证书系统支持,如图1所示。

分布式网络存储系统是由分布在网络上的多个数据存储服务器组成的。每个存储服务器都能存储用户的数据,很可能用户自己的主机就是其他用户的数据存储服务器。LDAP 数据定位服务器维护有所有存储数据的 URL 位置信息。当用户想获取指定的数据时,首先要查询数据定位服务器,获得数据存储服务器的确切位置,然后才能得到所需的数据。下面从数据存储和数据访问两个方面探讨如何实现分布式网络存储系统的安全。

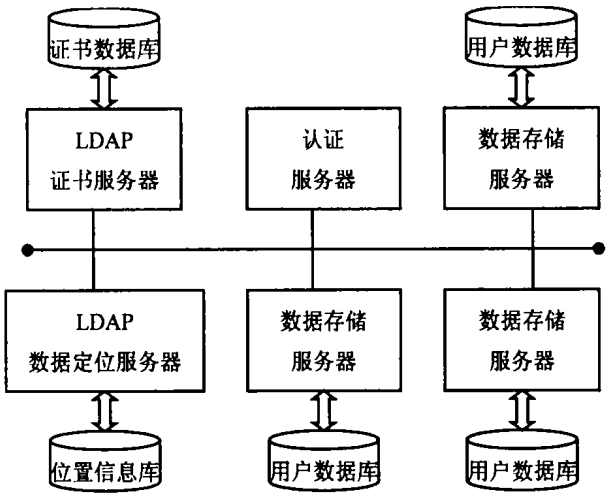

图1 分布式网络存储系统的逻辑示意图

3.1 数据存储的安全

为了保护数据的保密性和完整性,必须利用分布式证书系统为每个与数据打交道的用户和存储数据的服务器创建公开密钥证书和对应的私有密钥,并将所有证书存放在 LDAP 证书服务器上,便于使用者随时下载。每个用户和数据存储服务器都拥有自己的私有密钥。当用户生成敏感数据,如需要保密的文件时,立即采用随机产生的密钥加密(往往采用对称加密算法加密大量的数据)文件,并形成文件的数字签名。数字签名有两个,一个是文件未被加密时的数字签名,另一个是文件被加密后的数字签名,这两个签名都是采用用户自己的私有密钥生成的。

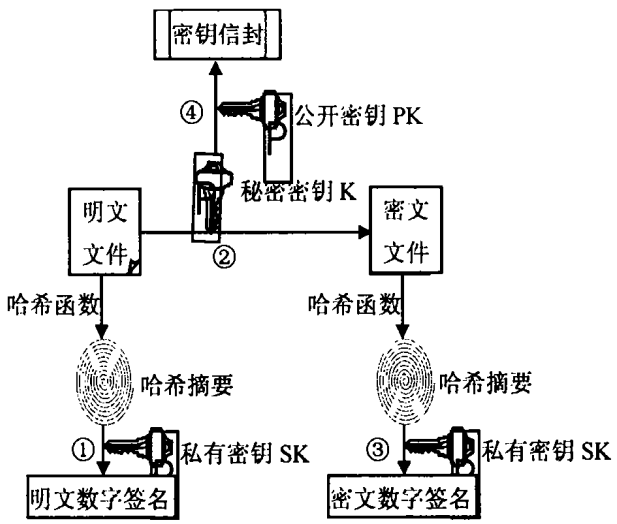

图2 实现安全的数据存储的过程

如果文件存放在本地,则需用用户自己的公开密钥(假设加密密钥与签名密钥一致,请注意,很多证书系统会提供专门的加密密钥)加密文件的秘密密钥;如果要将文件存储到指定的数据存储服务器

上，则需通过 LDAP 证书服务器获取指定存储服务器的公开密钥证书，并通过认证服务器验证该证书的有效性。如果证书是有效的，就采用该证书中的公开密钥加密文件的秘密密钥，形成“密钥信封”。最后，采用 XML[4]格式将文件的数字签名、密钥信封和对应的公开密钥证书组合在一起，与加密后的文件一起保存在本地或者指定的数据存储服务器上，再将文件的位置信息保存到数据定位服务器上。整个操作流程如图2所示，XML 属性文件的格式如图3所示。

```
<?xml version="1.0" encoding="UTF-8" standalone="yes"?>
<DataAttribution version="0.8">
    <Position>文件的 URL 位置</Position>
    <Author>作者的信息</Author>
    <Abstract>内容摘要</Abstract>
    <Certificate>作者的公开密钥证书</Certificate>
    <KeyEnvelop>密钥信封</KeyEnvelop>
    <PlainSignature>明文数字签名</PlainSignature>
    <CipherSignature>密文数字签名</CipherSignature>
    ... ...
</DataAttribution>
```

图3 与存储数据对应的 XML 属性文件

3.2 数据访问的安全

当用户想访问某个数据文件的内容时，首先要通过定位服务器查询它的存储位置，然后向对应的存储服务器提出访问请求。存储服务器收到请求后，采用自己的私有密钥对密钥信封解密，获得加密文件的秘密密钥；接着，它从证书服务器那里获得请求者的公开密钥证书，验证证书的有效性，然后采用请求者的公开密钥加密文件的秘密密钥，形成新的密钥信封，并修改对应的 XML 属性文件的“KeyEnvelop”节点；最后，存储服务器把用户所需的加密文件和新的 XML 属性文件返回给用户。用户收到返回结果后，验证文件作者的公开密钥证书是否有效，然后采用作者的公开密钥验证密文数字签名，以确保文件的完整性未遭受破坏；采用自己的私有密钥对密钥信封解密，获得加密文件的秘密密钥，解密文件，就能阅读文件的内容。整个操作流程如图4所示。

之所以采用两种数字签名，是因为有的文件内容不需要加密，只要使用文件的明文数字签名来保证其完整性和不可否认性；相反，对于加密文件，采用密文数字签名验证文件的完整性要比采用明文数字签名验证文件的完整性更有效、更安全，因为它的验证过程不需要对文件解密。

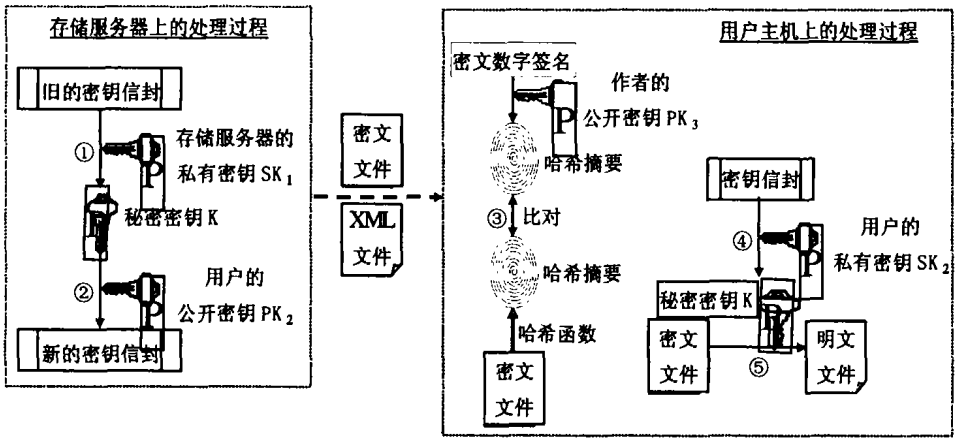

图4 实现安全的数据访问的过程

结论 采用基于内容的安全技术实现分布式网络存储系统的安全，不但解决了数据通信的安全问题，也解决了数据存储的安全问题，而且还具有其他诸多优点，适合于保护对等网络环境下的其他应用。基于内容的安全使安全措施与受保护的应用松散耦合，不需要对通信链路实施保护，不需要像 VPN 网关、安全代理服务器那样对数据进行集中的加解密服务，对用户进行集中的访问控制，对安全进行集中的管理，在保证系统原有效率的同时，体现了分布式的思想，因此特别适合于解决对等网络环境下的大规模、分布式应用系统的安全问题。本文只是从安全存储的角度来分析如何实现基于内容的安全，相信随着 XML 技术和.NET Web Service 的发展，这种安全解决方案会得到进一步推广。

参考文献

1 (美)Ford J L. 实用 Windows 对等网络连接教程[M]. 北京：清华大学出版社，2001

2 张红旗. 信息网络安全[M]. 北京：清华大学出版社，2002

3 冯登国，卿斯汉. 信息安全——核心理论与实践[M]. 北京：国防工业出版社，2000

4 (美)Ray E T. XML 入门[M]. 北京：中国电力出版社译，2001