doi:10.3969/j.issn.1001-4616.2018.02.006

基于一次一密的可信存储网关

张　昕,谢宾铭,赵　云,高若寒

(南瑞集团有限公司,江苏 南京 211100)

[摘要]　本文分析了国内存储系统现状和分布式存储的发展趋势,基于当前国家对于数据安全的需求,提出利用虚拟化网关整合存储和可信两方面需求的思路,通过具体的软硬件产品组合对思路进行测试验证,证明方案技术的可行性.

[关键词]　存储网关,存储虚拟化,可信存储,CDP
[中图分类号]TP391　[文献标志码]A　[文章编号]1001-4616(2018)02-0026-07

Trusted Storage Gateway Based on One-time Pad Cipher

Zhang Xin,Xie Binming,Zhao Yun,Gao Ruohan

(NARI Group Corporation,Nanjing 211100,China)

Abstract:This paper analyzes the current situation of the domestic storage system and the development trend of the distributed storage,as well as the demand for data security,puts forward the idea of using the virutal gateway to integrate the storage virtualization and security requirements,and proves the feasibility of the scheme by the feasibility test of the specific hardware and software product combination.

Key words:storage gateway,storage virtualization,trusted storage,CDP

随着信息技术的发展,存储技术经历了磁带、磁盘、磁带库、磁盘阵列到存储网络几个发展阶段[1],存储容量也快速从 MB、GB 向 TB、PB 级别演进. 当前,由于分布式存储具有部署简单灵活、容量与性能管理方便、成本低廉等特性,在国内企业应用市场发展得如火如荼. 与此同时,国产化、数据安全的需求也日益高涨,如何整合企业中现存的传统存储与新型分布式可信存储资源,实现逐步升级迁移是当下亟需解决的任务. 本文针对这一情况,在分析存储虚拟化技术的基础上设计一套存储虚拟化方案,利用存储虚拟化网关技术整合传统 SAN 存储,同时加入可信机制[2],平滑过渡到分布式可信存储.

存储虚拟化网关[3]已有多年发展历史,其通过对存储进行统一管理,对硬件复杂性进行解耦抽象,提供给用户一个统一化的标准接口,使用户更多地关注业务系统,降低存储的管理成本. 可信存储、可信计算则是一门新兴技术方向,为信息行为安全而生,提出行为可预期的超前概念,实现数据的真实性、数据的机密性、数据保护以及代码的真实性、代码机密性和代码保护的目标[4]. 本文的贡献就是从实际应用的角度在存储虚拟化网关中引入动态密钥加密和防篡改机制,实现可信存储虚拟化网关,以此为基础整合传统存储和分布式存储资源,并解决用户对于数据安全方面的需求问题.

1　可信存储网关技术方案

可信存储网关兼容现有存储,提供透明存储池、可信存储池两种存储池. 透明存储池用于添加已有存储,通过转发 IO 请求,逐步将生产数据迁移到可信存储中,以便在后台数据迁移完毕后进行切换,不影响现有生产环境;可信存储池用于添加新增存储,实现存储虚拟化,统一管理的功能.

1.1　方案创新点

本文创新点分两部分内容,一部分是动态分组密钥,用来模拟一次一密;另一部分是数据微观多维度,

收稿日期:2018-02-13.
基金项目:国家电网公司“异构灾备技术研究及应用”科技资助项目.
通讯联系人:谢宾铭,硕士,工程师,研究方向:管理科学与工程. E-mail:xiebinming@ sgepri.sgcc.com.cn

用来实现短时间范围内数据可逆.

1.1.1　*动态密钥方式数据加密*

数据加密领域,理论上只有一次一密才能实现无条件安全,其他情况下,如果计算力足够,总能穷举. 一次一密[5]常用于流加密方式,但存储领域并不适合这种方式,因此本文通过使用分组加密算法如AES、DES、SM4,并采用动态密钥形式,模拟一次一密的特点,增加了破解难度的同时不影响加密性能.

此方法可以防范物理设备丢失时的数据安全或者瞬间销毁数据,处理如物理设备失控、数据面临即将丢失、需要瞬间销毁、避免泄露等情况.

1.1.2　*数据防篡改*

数据防篡改即数据可逆,由于海量数据的量大价值小等特点以及从实用性角度出发,本方案采取只在微观尺度上维持多维度数据. 其基本原理及实现方式:由于基本存储单元的每次修改都会有日志记录,但最多保留 11 个维度的数据(可定义),最长时间长度为月(可定义). 通过在网关上建立虚拟存储单元 map 表的方法,实现在正常工作情况下,将数据映射到当前最新维度上,而当数据被篡改时,可以将数据映射到 11 个维度中任意一个进行回溯,恢复篡改前的数据.

1.2　系统整体架构

可信存储网关整体架构框架如图 1 所示.

(1)存储管理模块:对后端存储进行统一管理.

(2)镜像模块:将透明存储池里的卷数据映射到可信存储池里,便于生产数据迁移.

(3)日志模块:实现短信邮件报警功能.

(4)SCSI 驱动模块:兼容现有 SAN 存储设备,通过转发 IO 请求实现透明接管.

(5)透明存储池:通过直接映射原有存储,兼容客户现有存储架构. 在其上层使用 SCSI 过滤驱动模拟原 target 设备,然后转发 IO 请求.

(6)可信存储池:将新增存储添加到此存储池里,其原有数据会被抹除,加入到此存储池的所有存储设备原有数据会被重新初始化. 存储数据实现加密,若存储设备离开此环境,其内容将不可读取.

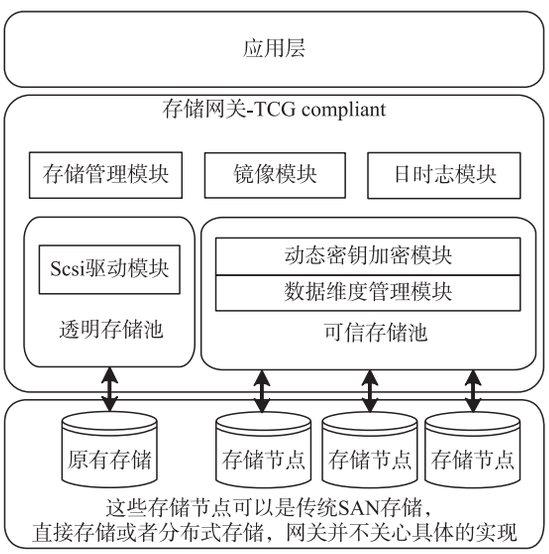


图 1　可信存储网关架构

Fig. 1　Architecture of the trusted storage gateway

(7)数据维度管理:对应可信存储里的 time stamping,由于性能及商业原因,其存在时间短暂,动态变化,因此这里称之为微观数据维度[6].

1.3　存储虚拟化技术

通过存储虚拟化[7]将后端连接的存储设备透明化,将其组成一个个存储池,然后在其之上分配逻辑卷提供给生产系统使用,对于生产系统来说,只需要关心逻辑卷. 因此存储虚拟化技术剥离了和硬件的耦合度,实现异构磁盘系统的有效整合及集中管理[8].

逻辑卷:LV,网关前端提供的存储设备,即生产系统中看到的存储设备,将其称之为逻辑卷.

物理卷:PV,网关后端连接的存储设备,将其称之为物理卷.

存储池:将一个或多个物理卷集合在一起形成的可用来分配逻辑卷的存储空间,逻辑卷只能从存储池里分配得到.

存储池分为:

(1)透明存储池:为企业内部已有的 SAN 存储,本网关仅对此存储池里的存储设备发起转发 IO 功能,用于无缝迁移接管数据.

(2)可信存储池:用于新增数据,具体分为两类:

(a)静态存储池:预先分配好和物理卷的对应关系;

(b)动态存储池:用时分配的策略,由于逻辑卷和物理卷的对应关系并不预先固定,因此实际需要再进行动态分配.

存储虚拟化网关,其很大一部分功能主要在于平滑过渡现有 SAN 存储到可信、分布式存储[9]. 涉及到的流程如下:

(1)将用户原有 SAN 存储接入到网关的透明存储池;

(2)在静态存储池里分配一个同样大小的逻辑卷;

(3)设定该逻辑卷为原有 SAN 存储的镜像;

(4)等到卷状态为同步完毕时,建立该镜像卷与生产系统之间的通道,直接取代原有存储设备.

1.4　动态密钥加密

单纯的分组加密算法[10]如 AES、DES、SM4 等,理论上可以通过穷举方式来破解. 因此本文提出动态密钥加密方法,模拟一次一密这种理想算法.

其数学模型如下:

$$Key = Hash(lambda(rawkeys, dataid)).$$

其中函数 Hash 的输入参数为一组密钥种子,由函数 lambda 计算得到,函数 lambda 的入参有两个:

rawkeys:原始密钥池,一共 256 个密钥,每个密钥长度为 256 位,系统初始化时随机生成,用证书加密存储在 UKey 或者生产系统里.

dataid:表示待加密数据的标志号,采用数据所在的存储位置也就是逻辑扇区号作为标志号.

函数 lambda 用来生成密钥种子,首先计算 dataid 的 crc 值,生成一个 32 位数,依次取其中的 8 位(对 256 取模)作为索引号,从原始密钥池中取出密钥种子,作为输出传递给 hash 函数.

函数 hash 是一个高性能散列函数,将 lambda 输出的 4 组 rawkey 作为因子组成一组数据,然后进行散列得到一个 256 位长的密钥用于后续的 AES、SM4 等加密算法.

由此可见,数据加密密钥由 rawkey 唯一确定,而 rawkey 又由 dataid 唯一确定,dataid 采用了数据存储位置作为标志,但同一位置只可能存储一份数据,因而具有唯一性,这样最终生成的密钥没有周期性,近似等于一次一密;而且其加密算法性能损耗只是在常规 AES、DES、SM4 之上加了一个快速散列函数 hash 的损耗,相对于加密算法以及磁盘 IO 来说不在一个数量级,可以忽略.

基本流程如图 2 所示.

其中:

D1 = D[CRC&0XFF]

D2 = D[(CRC>>8)&0xFF]

D3 = D[(CRC>>16)&0xFF]

D4 = D[(CRC>>24)&0xFF]

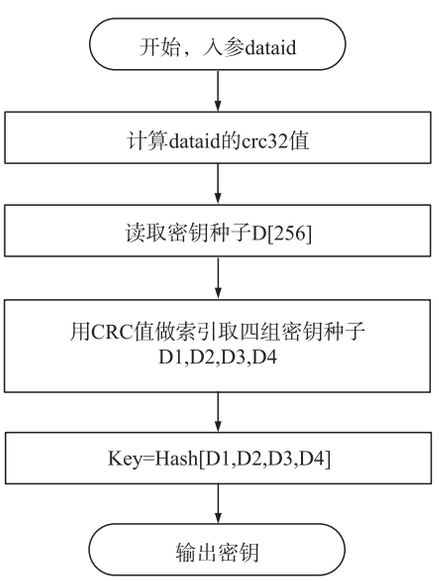


图 2　动态密钥基本流程

Fig. 2　Flow chart of the dynamic key generation

1.5　数据维度管理

数据维度的工作原理:在每次写数据时,首先通过数据备份构造一张链表,其中每个数据块对应一个表项,表项里记录了刚保存的数据所在的数据块位置,因此构成一个链表;由当前块开始遍历链表,可以找到历史时间的任意数据,每一次修改称为一个维度. 这个表称为维度表,里面包含下一维度的指针. 对于任意一块数据,维持预设定好的 N 维数据(短时间突发维度可以超过预定值的 11 倍),通过后台维护对早期维度数据的丢弃管理.

其原理示意图如图 3 所示.

当数据发生修改,可以将前一维度数据映射出来,然后进行恢复,如图 4 所示,映射维度 D2 的数据.

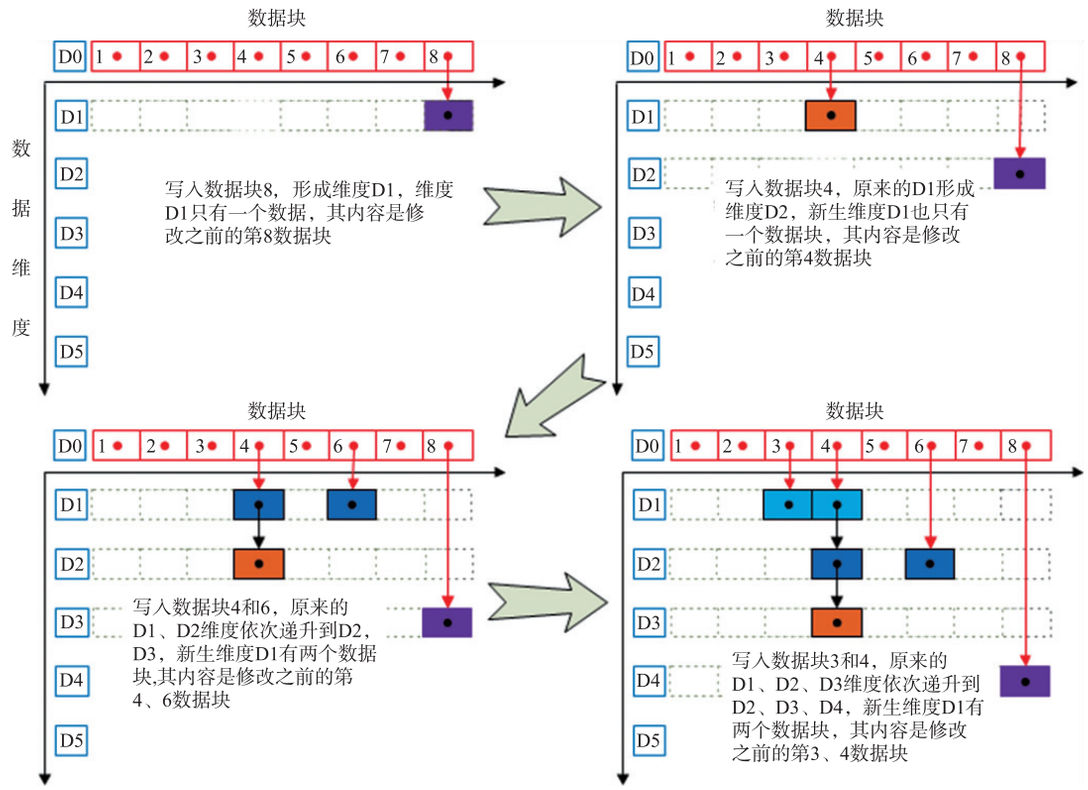
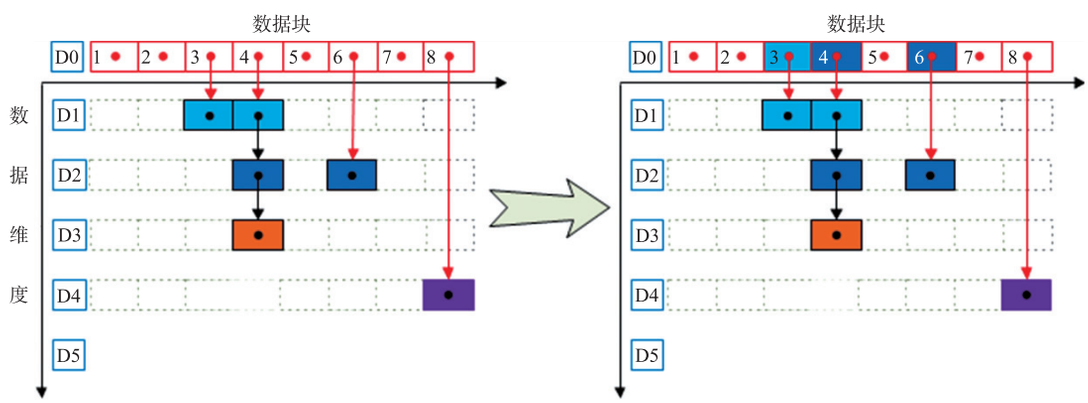


图 3　数据变化维度图

Fig. 3　Process of data upating



映射维度 D2,所有小于等于 D2 的数据都同时映射到当前维度上. 此例就是将数据块 3、4、6 映射到维度 D0 上

图 4　数据恢复示意图

Fig. 4　Process of data recovery

2　可信存储网关测试方式及测试内容

为了验证方案的可行性,通过功能测试和性能测试,确认可信存储网关是否可以正确实施并具有实际使用价值.

2.1　功能测试

测试目的:验证存储网关的基本功能、数据迁移、可信存储以及数据维度是否可以正常工作,是否存在使用价值.

测试步骤如下:

(1)Initiator 端:财务管控数据库服务器(Oracle 数据库),使用 8G 光纤卡连接到光纤交换机.

(2)存储:

(a)SAN 存储为宏杉 MS3300 存储,直接通过 8G 光纤卡连接到光纤交换机,在没有添加网关之前,Initiator 直接使用该存储作为数据盘,已经存有数据.

(b)分布式文件系统采用 5 节点的 ceph 集群[11],操作系统为 ubuntu-14.04,ceph 版本为 0.94,每个节点上的 OSD 配置为 10 块 4T SATA 盘,journal 文件大小为 80GB,副本数设置为 3,pg_num = pgp_num = 2 048. Ceph 及操作系统参数配置采用软件默认设置. 块设备采用 Ceph 内核方式,通过 FC 通信网络映射给 Initiator 端使用.

整体线路连接示意图如图 5 所示.

(3)划分 Zone

生产系统主机端口和网关端口 1 划分到一个 Zone,网关端口 2 和 ceph 为一个 Zone.

Zone 划分遵循以下原则:

(a)主机端口和网关的 target 端为一个 Zone.

(b)分布式存储、SAN 存储和网关的 Initiator 端为一个 Zone.

表 1 为 FC Switch 的 ZONE 端口划分表.

(4)网关存储池

(a)透明存储池里添加 SAN 存储宏杉 MS3300,提供 1TB 的 LUN 给生产系统作为数据盘使用,存储 Oracle 数据库文件;

(b)可信存储池里添加 ceph 分布式文件系统提供的同样大小的块设备,为其创建单独的存储池,使用所有空间创建一个卷,设置其为 SAN 存储的镜像卷[12].

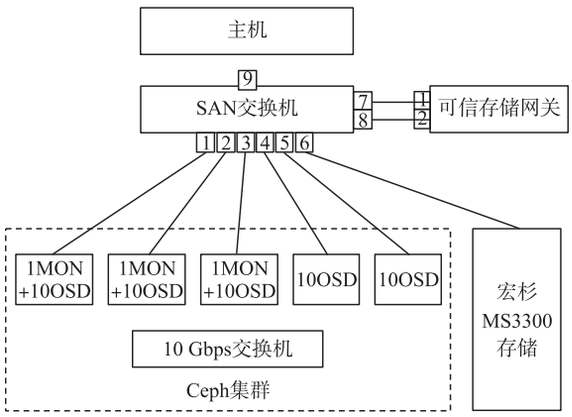


图 5　测试环境拓扑图

Fig. 5　Architecture of test environment

表 1　Zone 端口划分表

Table 1　Zone port configuration

存储端口	分区	网关端口
存储 1 节点端口	Zone1	1,7
存储 2 节点端口	Zone2	2,7
存储 3 节点端口	Zone3	3,7
存储 4 节点端口	Zone4	4,7
存储 5 节点端口	Zone5	5,7
SAN 存储 A 端口	Zone6	6,7
主机端口	Zone7	8,9

2.1.1　存储网关基本功能及可信测试

测试目的:验证存储网关的基本功能、数据迁移及可信存储是否可以正常工作,是否存在使用价值.

测试步骤如下:

(1)检查镜像实时复制是否完成,如果已经完成,则将 ceph 提供的存储映射到生产系统,原有 SAN 存储取消映射.

(2)切换回去,将 ceph 提供的块设备直接在 ceph 节点中 mount 出来,检查其内容.

测试结果如下:

(1)生产系统不受影响,成功完成数据平滑迁移.

(2)数据为加密状态,无法加载文件系统,证明存储设备的可信状态.

2.1.2　数据维度测试

测试目的:检测维度管理在占用有限存储空间的情况下是否能有效保护数据,是否有实际使用价值.

测试环境如下:

业务系统:财务管控数据库系统

操作系统:Windows server 2008 R2

存储大小:600GB

运行一个星期后,数据变化如表 2、表 3 所示.

表 2　数据变化量

Table 2　Amount of data changing

	数据量
初始数据	434GB
一周后数据	751GB
最大维度	63 921

表 3　数据变化统计

Table 3　Statistics of data changing

维度数	平面空间	比例	维度总空间=平面空间×维度数	比例
>2 000	66.78MB	<0.1%	300GB	39.95%
100~2 000	25.86MB	<0.1%	5GB	0.67%
2~100	899.66MB	0.21%	13GB	1.73%
1	433GB	99.76%	433GB	57.65%

数据维度分布如图 6 所示.

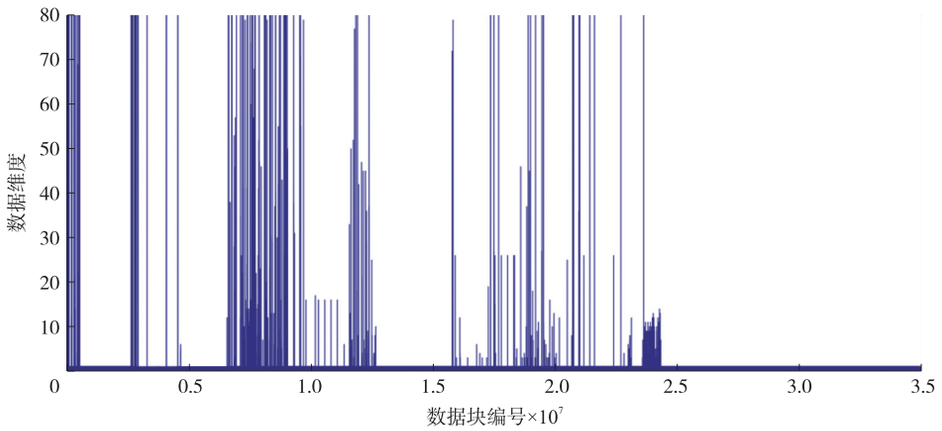


图 6　数据维度分布

Fig. 6　Data dimensional evaluation

从以上日常使用情况可知,99%的数据很少改动,频繁更新的数据只有 1%,比例很低,所以完全是可以通过限制维度上限,循环使用有限空间进行有效保护数据,存储资源消耗很低.

2.2　性能测试

测试目的:对比测试传统 SAN 存储和使用本网关后的虚拟化存储之间的性能差异,检查性能损耗是否在可接受范围.

测试步骤:后端 SAN 存储以及网关均采用 8G 光纤卡连接到光纤交换机,分 3 种情形连接进行对比测试:

(1)SAN 存储直接连接到生产系统;

(2)SAN 存储先连接到网关透明存储池,然后再连接到生产系统,数据进行加密处理;

(3)SAN 存储先连接到网关可信存储池,再接到生产系统,对数据进行持续保护并加密.

对比测试结果如图 7 所示,其中读盘性能损耗在 5%以内,写盘损耗在 10%左右;如果不做 CDP 持续保护,则性能损耗可以忽略.

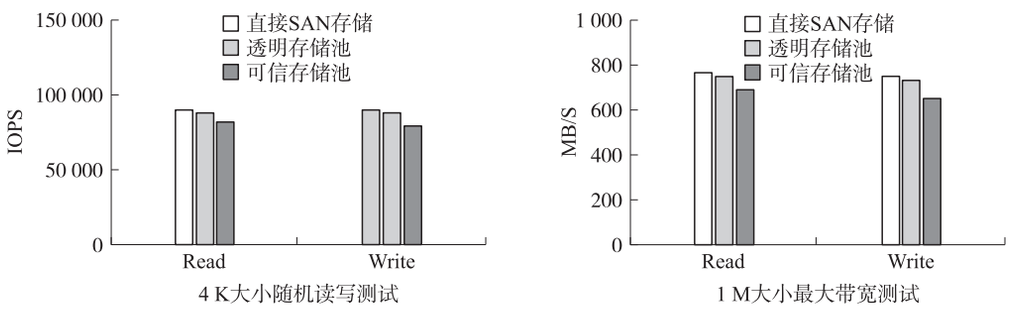


图 7　性能测试比较

Fig. 7　Performance evaluation

3　结束语

本文探讨了利用存储虚拟化技术整合传统 SAN 存储和新型分布式存储两种资源,通过存储虚拟化剥离硬件耦合度,实现存储资源的统一整合与集中管理,对分布式存储的应用普及具有良好的指导作用. 另一方面,本文在存储网关中引入可信机制实现多维度数据管理,通过模拟一次一密的特点,在不影响加密性能的同时增加了破解难度,防范物理设备丢失时的数据安全或瞬间销毁数据等业务场景,实现数据可逆并有效保护数据减少数据存储空间.

[参考文献]

[1]　杨宗博,郭玉东. 提高存储资源利用率的存储虚拟化技术研究[J]. 计算机工程与设计,2008(12):3224-3226.

[2] CHENGMING L,KOJI O. Trusted Storage Virtualization in Cloud Computing[C]//Proceedings of the Asia-Pacific Advanced Network(APAN). Bandung,2014,38:128-132.
[3] 常潘. 华东师范大学存储虚拟化改造[J]. 中国教育网络,2010(5):23-27.
[4] Sean W S. 可信计算平台:设计与应用[M]. 冯登国,徐霞,张立武,译. 北京:清华大学出版社,2015.
[5] DIRK R. Secure Communications with the One Time Pad Cipher Paper(English)[EB/OL]. http://www.snia.org,2008.
[6] TAKESHI T,ATSUO O. Technologies of ETERNUS VS900 Storage Virtualization Switch[J]. Fujitsu scientific and technical journal,2006,42(1):17-18.
[7] 李维林,朱志安,胡显涛. SVC存储虚拟化在容灾系统中的应用研究[J]. 数字技术与应用,2013(3):126-127.
[8] 王楠,蒋金虎. 存储虚拟化技术研究与比较[J]. 洛阳师范学院学报,2007(2):73-76.
[9] SHANNON C. Communication theory of secrecy systems[J]. Bell system technical journal,1949,28(4):656-715.
[10] SAGE A W. Ceph:Reliable,Scalable,and High-Performance Distributed Storage[D]. Santa Cruz:University of California,2007.
[11] Swidish Institute of Computer Science Secure Systems Group. Trusted Computing and Secure Virtualization in Cloud Computing[R]. Paladi Nicolae,2012:5-8.

[责任编辑:陆炳新]