基于数据中心的云计算审计系统设计探讨

Design of Cloud Computing Audit System Based on Data Center

张 伟(中国大唐集团科学技术研究院有限公司,北京 100040)
Zhang Wei(China Datang Corporation Science and Technology Researd Institude,Beijing 100040,China)

摘 要:
作为新IT模式的云计算正在逐渐进入实用化阶段。云计算以其基于互联网和按需获取等特征向用户承诺更低的IT运营成本。在降低IT运营成本的同时,云计算技术和分层服务等特点给网络安全带来新的挑战。基于数据中心云计算服务的信息安全审计系统研究分析,云计算审计系统在一定程度上降低数据泄露的安全风险,提高了云计算服务的可用性。安全可靠的云计算审计系统有利于提高云计算服务的安全保障能力,促进云计算产业快速健康地发展。

关键词:
云计算;审计系统;分布式
doi:10.12045/j.issn.1007-3043.2019.04.009
中图分类号:TN915.08
文献标识码:A
文章编号:1007-3043(2019)04-0040-05

Abstract:
Cloud computing as a new IT model is gradually entering the practical stage. Cloud computing promises users lower IT operating costs because of its Internet-based and on-demand access features. While reducing the cost of IT operation, the characteristics of cloud computing technology and layered services bring new challenges to network security. Research and analysis of information security audit system based on cloud computing service in data center shows that cloud computing audit system reduces the security risk of data leakage to a certain extent, and improves the availability of cloud computing service. Safe and reliable cloud computing audit system is conducive to improving the security of cloud computing services and promoting the rapid and healthy development of the cloud computing industry.

Keywords:
Cloud computing;Audit system;Distributed

引用格式:张伟. 基于数据中心的云计算审计系统设计探讨[J]. 邮电设计技术,2019(4):40-44.

0 前言

在我国,云计算应用市场有着巨大的发展潜力。一方面,我国拥有世界上数量最多的中小企业,对于这些处在成长期的中小企业而言,自己投资建立数据中心的投资回报率较低,并且很难与业务的快速成长匹配,而云计算的租用模式正好为这些中小企业提供了合适的解决方案;另一方面,众多的服务器、存储硬件厂商以及平台软件厂商都希望通过云计算平台将自己的产品推广到中小企业中,以便未来能获得更多的市场机会。数据中心的客户集聚了高科技中小企业,包括IT行业和金融行业的企业,这些企业在使用了数据中心提供的云计算基础服务后,必然对基于数据中心的云计算审计系统产生需求。信息审计应用于云计算服务能够提高网络安全防护水平,云计算审计系统所覆盖的范围能够发现大多数云计算服务可能存在的信息安全问题,从而减少信息风险。目前市场上一些云计算审计系统存在颗粒度相对较粗、审计信息不完全,事件发生后无法真正去溯源调查事件原因等问题。

收稿日期:2019-03-14

1 云计算审计系统体系结构

数据审计是以审计底层原始数据为切入点,在对信息系统内部控制测评的基础上,通过对数据的采集、

转换、整理、分析和验证，并且运用查询分析、数据挖掘等多种技术方法构建模型进行数据分析，发现趋势、异常和错误，从而收集审计证据，实现审计目标的审计方式。

数据审计在分布式云存储的环境下显得尤为重要。首先，在该环境中数据的完整性得不到保证，元数据服务器（Metadata Server）和数据块服务器（Chunk Server）相分离，之间的更新有延迟，造成元数据中描述的文件与实际存储在数据块中的文件不一致；其次，在多副本的条件下，对其中某一副本的更改也会造成数据不一致；而随着越来越多应用在分布式环境下的部署，各种应用的跟踪调试信息变得分散凌乱，特别是现在的一个应用程序可能需要多台后端服务器的配合，从而造成整个业务流程日志出现片段化，不利于排错和维护。当某一个节点出现问题后，如何能够快速地定位、纠错并恢复运行，这就是数据审计的职能。

此外，数据的隐私保护（文件是否被非法访问或更改）、部署在各个节点上的软硬件环境日志信息、安装的虚拟机部署状态信息，也是为了审计而需要采集的重要数据。

引入云审计系统之后，各个服务器节点的审计信息可以统一地通过一台或者多台审计服务器集中存放（根据策略），方便信息归总和查询，从而可以做到任何时候、任何地点的信息跟踪。

1.1 分布式文件系统

云审计系统针对海量级异构平台，构建在分布式文件系统之上。文件系统提供分布式的存储和读取，并确保文件内容的一致性和文件本身的安全性。另外分布式文件系统需提供一个高性能的网络连接设备及网络传输机制供各种应用程序调用。

1.2 审计服务器

审计服务器收集审计客户端的日志信息，提供一个高性能的框架，支持大并发量以及突发量的审计信息处理。支持多种网络协议，例如TCP、UDP等。审计服务器本身要完成对于审计内容的处理，例如时间戳的更新，内容完整性的校验。另外审计服务器还需要对接收的审计内容进行过滤，以方便维护和调试。同时审计服务器能够完成审计策略的更新，更新后会主动通知审计客户端，策略的更新需要进行可靠性保证，保证策略的一致性。

1.3 审计客户端

审计客户端安装在需要被审计的节点（服务器或虚拟机），需要支持多实例运行（即可以在同一节点上分别审计不同类型的信息）。另外还需要支持多个级别的传输机制，根据审计内容的优先级程度自主切换。

客户端提供日志消息内容的缓冲，为服务器切换提供大容量缓存。客户端需要完成负载均衡任务，能够在有限的信息条件下完成最佳服务器的选择。客户端还需要完成策略文件接收的可靠性检查，保证与审计服务器的审计策略相一致。

2 审计服务器和客户端设计

2.1 审计服务器设计

审计服务器的系统结构如图1所示。服务器主要由3个组件构成：审计服务框架、审计数据操作集、第三方应用或统计分析模块。

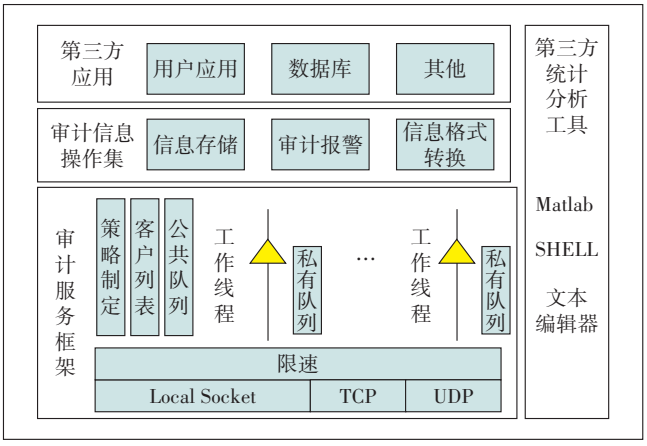

图1　审计服务器设计图

2.1.1 审计服务框架

审计服务框架的主要作用是正确接收各审计客户端发来的信息，同时维护一个客户端列表，若用户制定了新的审计策略，它将快速地对各客户端进行策略更新。审计服务框架主要由协议模块、策略模块、限速模块、审计报文队列、客户列表以及工作线程这几大部分构成。

目前协议模块只支持UDP和TCP 2种协议，审计报文内容可以基于任意一种协议之上，具体的选择由客户端决定。安全起见，策略更新只能通过TCP协议传输，以保证各审计客户端能够正确接收到最新的审计策略。

限速功能主要是审计服务器为了防止被攻击而采取的措施，可以针对特定地址或者审计报文总量进行速度限制，从而有效地维护服务器的高可用性。

下面开始描述一下审计信息接收的过程及各模块

间的交互。

审计服务框架在启动后，会根据用户要求，创建一个或者多个工作线程用来接收审计数据，而公共队列可以认为是接收的缓冲区。线程的数目是和CPU数量密切相关的，尽量做到均匀负载。可以将多台审计服务器组成集群，不同服务器上的工作线程采用同样的策略并行协同工作。

接收队列分为2种，一类是私有队列，每一个工作线程都有自己的私有队列空间，从socket接收到报文后会进行一些公共的处理，例如时间戳更新、完整性校验、公共策略检查等等。处理完毕之后就会把报文放入本线程的私有队列之中，缓存到一定数量再交由上层审计数据操作集批量处理。另一类队列是公共队列，它的容量要比私有队列大得多。当私有队列的报文数量达到一定阈值后，工作线程会有超负荷的危险，这时对新到达的报文不做处理，直接丢入公共队列中，当负载高峰过后，私有队列中的报文数量低于阈值时，再从公共队列中取出进行处理，可见，引入公共队列主要是为了避免工作线程因超负荷而丢弃报文（见图2）。

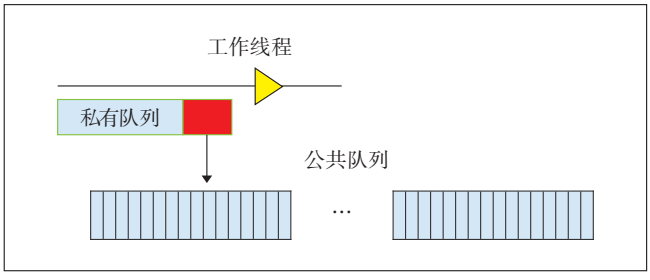

图2 私有队列与公共队列

对于审计报文消息，审计服务器不对UDP传输的报文提供任何形式的保证，但是只要这个报文能够被审计服务器接收，那么就会被认为是正确处理。在极端情况下，可能存在网络中断或者本地两级队列全满的情况，导致日审计报文丢失，对于这种情况，审计服务器不提供重传机制。对TCP传送的报文充分利用TCP本身的可靠性传输，但是这会带来很多额外的开销，除非极其重要报文（例如策略更新），否则不推荐采用TCP传输。

审计服务框架里面还有一个比较重要的策略模块，策略是审计客户端与相应审计服务器能够彼此连接的重要依据，服务器上每个工作线程采用一个策略，并根据策略标识来辨别需要接收各种审计信息。可想而知，只要它里面存在多个工作线程，一个服务器上可以制定任意多个策略。此外，服务框架可以对一个工作进程使用的策略进行动态更新，并同时将策略发送给对应的审计客户端，以支持动态批量审计。

策略的描述存在于策略文本文件中，工作线程通过该文件的地址与其中策略进行绑定。管理员可以通过各种编辑器进行编辑。策略的制定包含几个步骤。

a）更新策略版本号，只有接收到更新的策略版本，审计客户端才会进行策略更新。

b）设置客户端IP地址及掩码，例如192.168.2.10/24，IP地址之间需用逗号隔开；同样也可以用网段范围的形式出现，例如：192.168.2.5/24~192.168.2.15/24。

c）设置策略标识，即需要审计客户端发送什么样的信息，包括客户端软硬件运行状态、数据完整性、一致性状态、隐私数据访问记录、虚拟机运行状态等。客户端发出的报文中也含有这个标识，工作线程只接收包含相应标识的报文。

d）设置策略过期时间：设置一个时间窗口，当策略使用时间超过该窗口后，审计信息的发送和接收工作都将自动停止。此时需要替换为新的策略。

e）报警条件：当收到异常的审计信息，服务器端要给出报警提示，关于异常的定义要在策略中说明。

策略制定完毕以后，服务器按照策略和远程客户端连接，将当前在线的客户端组成客户列表并加以维护。成功以后开始发送策略版本号，考虑到会出现策略更新的情况，只有版本号比客户端的版本新才继续策略的传输，否则客户端直接拒绝。在客户端开始发送以后，通过确认审计报文的序列号来保证数据的正确传输。

动态更新审计服务框架维持的客户列表。一个新的客户端上线，会向服务框架发送请求接入的报文，若该客户端与策略中指定的IP地址相匹配，服务框架就把客户端信息加入这个列表，运行过程中不会删除客户信息，直到新的策略更新，服务框架会尝试连接所有该策略下的客户端，若尝试一定次数失败以后，就会把客户信息从链表删除。

2.1.2 审计信息操作集

审计信息操作集处于服务框架和第三方程序之间，起到了承上启下的作用，它将接收到的审计信息存储在服务器上，并根据上层的应用对信息进行格式上的转换。

a）审计信息存储：从工作线程的私有队列上取出审计报文，并以用户指定的格式保存在服务器本地的

审计日志文本文件中，用户可通过编辑工具对其进行各种操作。

b）审计信息报警：当接收到异常审计信息后发出报警。

c）审计信息格式转换：根据上层应用（数据库应用、数据挖掘等）的要求，进行相应的格式转换。

2.1.3 第三方应用或统计分析模块

审计服务器本身不做审计文件的检索、数据挖掘、压缩等操作，这些工作都由第三方程序（Matlab、Shell、文本编辑器等）完成，但是审计服务器提供接口，允许第三方程序对文件进行压缩等操作。

此外，审计服务器本身不考虑审计文件备份的问题，这方面的工作由云存储分布式文件系统来实现，服务器仅仅是共同打开同一份文件进行读写，背后逻辑借助文件系统来保证原子性、一致性以及安全性。

2.2 审计客户端设计

审计客户端安装在需要被审计的节点上，并根据审计策略收集相关信息，并发送给审计服务器。这些信息包括隐私数据访问记录、数据完整性验证信息、虚拟机日志、软硬件环境日志或其他用户指定的信息。设计图见图3。

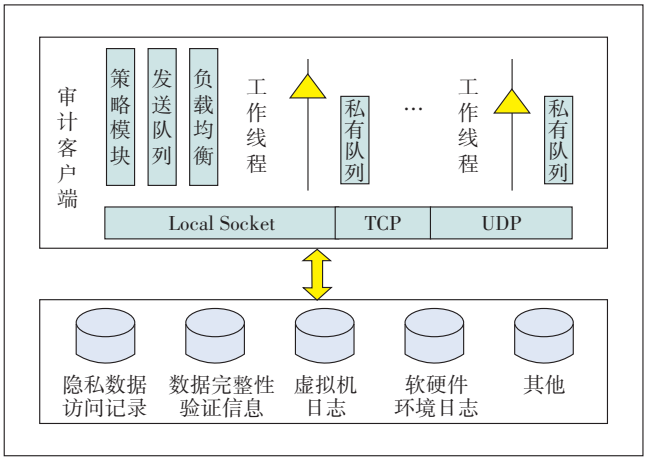

图3　审计客户端设计图

2.2.1 通信协议

客户端目前只支持TCP和UDP 2种通信方式，系统启动时候默认全部创建完毕。对于审计信息报文，2种协议都可以支撑，但推荐使用UDP协议，出于安全的考虑策略更新的消息只能通过TCP传输。

2.2.2 策略模块

策略模块主要完成策略文件的可靠性更新。该模块被加载的时候，首先创建监听socket用于策略更新，并将策略版本号设为0，表示此时尚无策略。当审计服务器的工作线程加载新的策略文件后，会主动触发更新机制，这时客户端会收到更新请求，里面包含新的策略版本号。客户端比较服务器版本号和本地策略文件版本号，如果版本号更大则启动更新流程，否则拒绝服务器请求。

为了防止重复更新（一个客户端可能注册在多个服务器列表上面），确认启动更新之后，首先立即把本地文件版本号升级，这样就避免了重复数据。一旦更新失败，回退策略版本到本地策略文件，并发一个报文通知审计服务器请求重传。新的策略更新完毕之后，刷新到本地策略文件。接着通知负载均衡模块，重新计算审计服务器地址，并且更新通知给客户端框架。

2.2.3 负载均衡

负载均衡模块结构如图4所示。

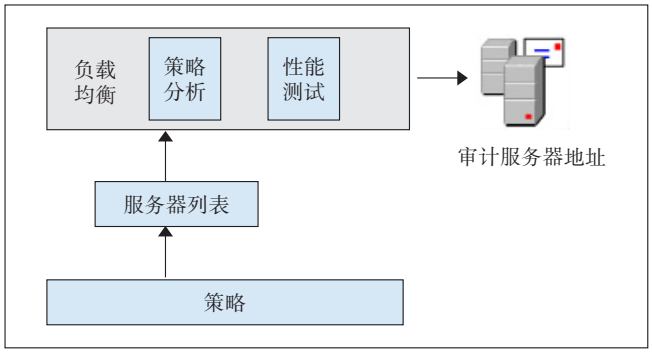

图4　负载均衡模块结构

由于客户端可以注册在多个审计服务器上的工作线程上，负载均衡就成为客户端里面非常重要的一个功能组建，它要选择出最合适的服务器作为接收方。负载均衡模块输入是策略模块，输出是合适的审计服务器地址。负载均衡模块主要完成策略信息的分析以及服务器的传输性能测试。之后，该模块也会以固定的时间间隔来计算传输时间，做到动态的负载均衡。

客户端所使用的UDP/TCP socket也是由负载均衡模块创建的，在传输的策略发生变化时（例如丢包情况严重），负载均衡程序透明地完成socket替换，客户端框架只感知数据发送的成功与否，然后采取相应措施，连接情况都是透明的。

2.2.4 工作线程

客户端工作线程根据策略收集节点上的审计信息，包括数据完整性、隐私数据保护、各种日志等，除可靠性要得到保证外，还需要一定的灵活性。

在客户端实例启动成功，并正确接收到服务器发

送的策略后，客户端会创建出工作线程。此后，工作线程在策略指定的时间窗口内不断采集审计信息，并将它们作为报文放在发送队列上等待发送。如果在运行过程中客户端和审计服务器间的连接出现问题（出现网络拥塞，或者服务器端出现故障），此时将审计报文放入自己的私有队列中，等到连接正常后再将它们放入发送队列，如果长时间出现连接故障，则私有队列内的报文会超过某一阈值，此时客户端就会认为审计工作无法完成，自行销毁并释放相应空间，不会给系统其他方面带来风险。

2.2.5 发送队列

如上所述，在客户端也存在公共、私有两级队列，但它和服务器端的两级队列功能略有不同。

考虑到一般来说服务器端负载较大，服务器端的工作线程使用自身私有队列接收审计报文并进行各种检查，审计数据操作集也是从私有队列里读取检查后的审计信息。如果队列长度超过阈值后，工作线程会直接将新到达的报文暂存公共队列中，以后择机再对它们进行处理。

而对于客户端来说，单个所在节点发送的报文相对较少，有一个公共的发送队列已经够用，各工作线程之所以还设置私有队列是由于：当审计服务器和客户端间出现连接故障时（一般由负载均衡模块告知），工作线程会将新的报文放入自己的私有队列中，并不会发送出去，之后再根据情况来决定是将私有队列中的报文移入公共队列发送，还是工作线程自毁。

2.2.6 报文格式

审计报文的格式定义如下：

<timestamp><ipaddress><severity><customs> : <message body>。

<timestamp>是客户端自行添加的，时间是从本机读取的，全局上来看可能不同步，但是对于本机特定应用而言，顺序是保证的。

<ipaddress>包括服务器的地址以及客户端的地址。

<severity>客户端可以制定默认的严重级别，当超过该级别后则认为本节点出现异常情况，服务端接收到该报文后立即发出报警。

<customs>自定义字段是可选的，如果用户没有指定，则该区域为空。

<message body>，消息体，即需要发送的审计信息，用户自定义，支持变参传递。

3 结束语

通过上述措施，在一定程度上保障了基于数据中心的云计算审计系统高可用性。云计算审计系统收集云服务方和用户的审计数据，保证审计数据的真实性和完整性。云计算审计系统需要根据不同环境、运营经验不断完善。安全可靠的云计算审计系统有利于提高云计算服务的安全保障能力，促进云计算产业的快速健康发展。

参考文献：

[1] 张剑，陈剑锋，王强. 云计算安全审计服务研究[J]. 信息安全与通信保密，2013(6)：62-64.
[2] 王剑锋. 云环境下外包数据的高效检索及安全审计技术研究[D]. 西安：西安电子科技大学，2016.
[3] 徐洋，朱丹，张焕国，等. 云环境下基于代数签名持有性审计的大数据安全存储方案[J]. 计算机科学，2016，43(10)：172-176.
[4] 杨阳. 云计算时代安全体系防御措施探讨[J]. 网络安全技术与应用，2016(2)：24-24.
[5] 郑鑫. 云计算安全体系架构与关键技术研究[J]. 通讯世界，2015(20)：227-227.
[6] 王永建，朱运起，徐杨，等. 基于云计算的智慧政务安全体系设计研究[J]. 通信技术，2016，49(4)：462-468.
[7] 余伟明，王伟. 基于融合式安全设备的云计算安全域体系设计[J]. 邮电设计技术，2016(1)：39-44.
[8] 陈鄂湘，裴俊豪，项晖. 云计算环境下信息安全体系架构研究[J]. 电信工程技术与标准化，2016，29(12)：72-77.
[9] 程宏兵，赵紫星，叶长河. 基于体系架构的云计算安全研究进展[J]. 计算机科学，2016，43(7)：19-27.
[10] 翟友钧，赵旦谱，台宪青. 一种面向主动安全的动态采集策略设计与实现[J]. 信息技术与网络安全，2018，37(4)：36-40.
[11] 王磊，齐明，李江泓，等. 基于公有云存储技术的医疗大数据建设实践与思考[J]. 中国数字医学，2017，12(1)：12-14.
[12] 陈飓. 云计算安全保密体系架构及其关键技术[J]. 山东工业技，2016(16)：289-290.
[13] 邓平. 基于公有云的企业信息化安全风险研究[J]. 中国管理信息化，2016，19(13)：74-76.
[14] 胡昌平，黄书书. 公有云存储服务中的用户权益保障[J]. 情报理论与实践，2016，39(11)：17-21.

作者简介：
张伟，毕业于北京邮电大学，高级工程师，硕士，国资委网络安全专家、中国大唐集团科学技术研究院专家，主要研究方向为计算机应用与信息安全。