A New Cloud Architecture of Virtual Trusted Platform Modules

Dongxi LIU†‡, Member, Jack LEE††*, Julian JANG†, Surya NEPAL†, and John ZIC†, Nonmembers

SUMMARY We propose and implement a cloud architecture of virtual Trusted Platform Modules (TPMs) to improve the usability of TPMs. In this architecture, virtual TPMs can be obtained from the TPM cloud on demand. Hence, the TPM functionality is available for applications that do not have physical TPMs in their local platforms. Moreover, the TPM cloud allows users to access their keys and data in the same virtual TPM even if they move to untrusted platforms. The TPM cloud is easy to access for applications in different languages since cloud computing delivers services in standard protocols. The functionality of the TPM cloud is demonstrated by applying it to implement the Needham-Schroeder public-key protocol for web authentications, such that the strong security provided by TPMs is integrated into high level applications. The chain of trust based on the TPM cloud is discussed and the security properties of the virtual TPMs in the cloud are analyzed.

key words: TPM, cloud, virtualization, trust service

1. Introduction

Trusted computing is a category of technology developed by the Trusted Computing Group (TCG) [1] to facilitate the development of trusted systems. The standards of trusted computing specify the hardware and software components needed to build trusted systems. In particular, the hardware component is a chip called Trusted Platform Module (TPM), which is used as the hardware root in system trust.

The trust of TPM lies in its capabilities of secure key management (i.e., key generation, storage and use), and secure storage and reporting of platform configuration measurements. A TPM can generate RSA key pairs. The private RSA keys are always used within the TPM, never leaving it without encryption. A TPM stores measurements in a set of Platform Configuration Registers (PCRs), which are physically protected. The platform measurements stored in PCRs are signed with a key in the TPM when reporting the integrity of a system for remote attestation [2].

TPMs are currently provided by embedding them into computer motherboards. For an application to benefit from the TPM functionality, there are several requirements. First, the computer running the application must have a TPM. Second, the supporting softwares (e.g., TPM drivers and TCG Software Stack (TSS) [3]) must be installed for the application to access the TPM. Third, the application users must be willing to use the TPM since the computer being used might not be owned by them.

However, the above requirements are sometimes not easy to satisfy, thus making TPMs not so usable and hampering the wide acceptance of TPMs by secure applications [4]. A lot of computers (new or legacy) are not manufactured with TPMs. Most IBM blade servers [5] do not contain TPMs, nor do many resource-constrained embedded systems due to the size and cost overheads of a separate TPM [6]. There is the software-based TPM emulator [7]. However, it is only developed for Unix and not as secure as physical TPMs. For the requirement of supporting software, the current TSS is mainly implemented in C language such as TrouSerS [8] or in Java such as jTSS [9]. Hence, it is hard for the applications developed in other languages to access the TPM functionality. For example, Javascript code in web applications cannot access the TPM functionality easily. At last, application users may be reluctant to use TPMs in the computers they use but do not own (e.g., a public computer in an Internet bar) since the TPMs there may not have proper keys and PCRs.

In this paper, we propose a cloud architecture of virtual TPMs (the TPM cloud), embodying the concept of infrastructure as a service in cloud computing [10]. Our motivation is to improve the usability of TPMs. A usable security mechanism is more likely to be widely and effectively used [4], [11], [12]. The TPM cloud will help applications to benefit from the strong security provided by TPMs.

From the TPM cloud, users can apply for their own TPM instances (or virtual TPMs) on demand, as exemplified in Fig. 1. The TPM cloud contains a cluster of physical TPMs and virtualizes them to provide TPM instances for a large number of users. Consequently, applications can access the TPM functionality irrespective of the availability of TPMs in the underlying computers. Moreover, even if users run their applications in different computers, they still can access the same TPM instance since it is provided as a service in the cloud. Hence, there is no need to migrate private keys among computers, avoiding inconvenience and potential security problems. Cloud computing advocates the de-
livery of services in standard protocols, such as Simple Object Access Protocol (SOAP), so there is no interoperability problem between the TPM cloud and its applications. Our contributions of this work are summarized as follows:

- We propose an architecture of the TPM cloud (Sect. 2). The TPM cloud facilitates the usability of TPMs since it addresses the requirement problems as discussed above. The TPM functionalities in the cloud are provided as Web service operations, so they can be easily integrated into application design and implementation. Easy access to TPM functionalities is also realized as a key step to promote the wide acceptance of TPMs and regarded as a future working direction in the report [4].

- We formalize the functionality of the TPM cloud (Sect. 3). The formal functionality specification lays a foundation for analyzing and implementing the TPM cloud. To support a large number of users, the TPM cloud virtualizes a cluster of physical TPMs. Based on the formalization, we suggest some improvement to the TPM specification, which can make TPMs more flexible to be used in various contexts.

- We implement a prototype of the TPM cloud and apply it to implement the Needham-Schroeder public-key protocol [14] for Web authentication (Sect. 4). In the implementation of this protocol, both the user and the Web server depend on the TPM cloud to decrypt encrypted messages. That is, their private keys are manipulated only by the TPMs in the cloud. Hence, even if the user logs onto the server through a public computer, the private key is still secure and not released to the public computer. The TPM cloud brings strong security to Web applications without sacrificing users flexibility of using different computer platforms.

- We analyze the security properties of the virtual TPMs in the TPM cloud by comparing them with physical TPMs and software-based virtual TPMs. We conclude that the virtual TPMs in the TPM cloud are as secure as physical TPMs in practical applications (Sect. 5).

A preliminary version of this paper was presented in [13]. In this new version, we revised the key hierarchy in the TPM cloud, and formalized all components of the cloud architecture and more cloud commands to cover the typical use of the TPM cloud. The new key hierarchy makes it easier to migrate keys among physical TPMs, hence making the cloud architecture simpler as discussed in Sect. 3. We also discussed the establishment of trust chains based on the TPM cloud and analyzed the security properties in this version.

2. Architecture

The architecture of TPM cloud is shown in Fig. 2. This architecture includes a virtual TPM service, a user management component, a cryptographic service, and a number of physical TPM services and their management.

2.1 Virtual TPM Service

The virtual TPM service has the cloud port as the interface to users or applications, such as the Javascript Web pages and the Web server in the implementation of the Needham-Schroeder public-key protocol. Through this port, users or applications send cloud commands to consume TPM services, such as registering new TPM instances, creating keys and signing data.

The secure I/O module in this component protects the communication between the TPM cloud and its users. The TPM cloud is designed to have a Cloud Key (CK), which is a RSA key. We denote the public part of CK as PKc. Every command cloud received should be encrypted with PKc. In the TPM cloud, all operations involving private keys must be performed within physical TPMs (a security feature of TPM), so the secure I/O module relies on physical TPMs to decrypt the encrypted cloud commands. In other words, the secure I/O module needs not to know the private CK.

The result of virtual TPM service should also be protected, since the result (such as a decryption result) might contain valuable data. If needed, the result of cloud commands is encrypted with symmetric encryption algorithms before sending back. By using symmetric encryption, users do not keep and manage private keys on their platforms. The symmetric key for encrypting the result is generated by users and provided as an argument in the cloud command.

The execution planner determines the execution of cloud commands after being decrypted by the secure I/O module. It needs to select the component (the TPM management, the cryptographic service or the user management) and may need to divide a cloud command into a sequence of steps, suitable for other components to execute. As an example, to execute a cloud command for data decryption, the execution planner needs to load the key into a physical TPM and then call the physical TPM_UnBind command.

2.2 User Management

The user management component manages the state of users and maintains a TPM instance for each user. A TPM in-

†This report appears after our paper [13] is presented.
stance includes the virtual TPM state, a key hierarchy and a set of virtual PCRs.

The virtual TPM state indicates the state of a TPM instance by simulating state flags in the physical TPM, such as the flags indicating whether a TPM instance is enabled or disabled, activated or deactivated. The virtual state does not need to support all state flags in the physical TPM since some flags might not make sense for TPM instances in cloud, such as the state flag indicating the physical presence of a human since virtual TPMs in the cloud are accessed as a network service. On the other hand, TPM instances may have flags not included in physical TPMs. As a future work, a delegated flag could be used to indicate whether a TPM instance is delegated by its owner to other users.

Before introducing the key hierarchy for a virtual TPM, we describe briefly the key management in a physical TPM. Each TPM has a unique endorsement key (EK), which is generated by the chip manufacturer. Before using a TPM, users need to take the ownership of the TPM and create a storage root key (SRK). Both EK and SRK are RSA key pairs. Other RSA keys in the TPM are created under a parent key. Their private keys are always used within TPM and when released outside TPM they are encrypted with their parent keys. The SRK can be used as a parent key.

The key hierarchy for a virtual TPM is shown in Fig. 3. A virtual TPM has a virtual EK and a virtual SRK, which are both generated by a physical TPM with a Virtual TPM Root Key as their parent key. As a migratable storage key created under SRK, the Virtual TPM Root Key is created on one physical TPM, and then migrated and loaded into every physical TPM. Therefore, the virtual EK and SRK can be loaded into every physical TPM since the Virtual TPM Root Key is already there. Using Virtual TPM Root Key is an important improvement over the old cloud architecture in [13]. The Virtual TPM Root Key facilitates the management of physical TPMs and the migration of keys among them, with more details discussed in Sect. 3.1. In addition, the Cloud Key is also created on one physical TPM, and then migrated and loaded into every physical TPM.

The virtual EK is created as a binding key since its public key is used to encrypt data, for instance, by Privacy CA when activating an identity key. The virtual SRK is a storage key and used as the parent key to create other keys. All keys in the virtual TPM are migratable, such that they are not bound to a physical TPM. This feature is useful to maintain the availability of virtual TPMs in case of physical TPM failure and balance the load among physical TPMs. A key can move from one physical TPM to another for unbinding data if the target TPM is not busy.

A virtual PCR is a sequence of bytes with the same length as a physical PCR. Virtual PCRs stay outside physical TPMs. To protect their contents, we encrypt virtual PCRs with the virtual public EK. Before executing a PCR dependent TPM command, the virtual PCRs need to be decrypted with the corresponding virtual private EK. Since virtual PCRs are not in physical TPMs, the TPM commands dependent on PCRs need to be specially treated, as discussed below.

2.3 Cryptographic Service

The cryptographic service implements some cryptographic operations (e.g., symmetric encryption algorithms) that are not provided by physical TPMs. The cryptographic operations are similar to vendor-specific commands in some physical TPMs. For example, an Atmel TPM has the specific commands TPM_BindV20 and TPM_VerifySignature for public-key encryption and verification, respectively.

The cryptographic service also deals with the TPM commands relying on PCRs. A critical feature for physical PCRs is that they can only be extended and not all of them are resettable. Hence, it is hard to swap a set of virtual PCRs into the physical PCRs as in classic virtual memory management systems. However, we realize that some PCR dependent TPM commands can still be implemented in physical TPMs after rewriting, while others have to be simulated in software. Our solution is that if a command only refers to the state of PCRs, then it is rewritten for execution on a physical TPM; if a command needs the actual values of PCRs, then it is simulated. A command is said to refer to the state of PCRs if it does not change PCRs and only use them to affect the execution of subsequent TPM commands; otherwise we say it needs the actual values of PCRs. For example, the command TPM_Extend needs the actual values of the argument PCRs since it changes the values of PCRs; the command TPM_CreateWrapKey refers to the state of the argument PCRs since the resulting key can only be loaded into a TPM by a subsequent command TPM_LoadKey if the PCRs are not changed. That is, the commands TPM_CreateWrapKey and TPM_LoadKey do not care about the actual values of PCRs, just requiring them to have the same values. Moreover, the command TPM_CreateWrapKey has to be executed by a physical TPM since it generates RSA keys. The method of rewriting PCR dependent TPM commands is described in Sect. 3.3.

2.4 Physical TPMs and Their Management

The TPM cloud depends on physical TPMs to manipulate private keys and execute TPM commands (except some commands relying on PCRs). The TPM cloud might have a large number of users. For the scalability of the TPM cloud, the cloud architectures supports a number of physical TPMs.
The physical TPM management component coordinates all physical TPMs to execute multiple TPM commands in parallel. The scheduler in this component determines which physical TPM is selected to execute a new command, such that the workload of physical TPMs is balanced. The executor accepts the requests to execute TPM commands and issues the commands to physical TPMs. The key loader loads a key into a physical TPM to be used by other TPM commands. The key is not necessarily created on the same TPM.

3. Formal Functionality of the TPM Cloud

We describe formally the functionality of each TPM cloud component and their interactions to explain the processing of cloud commands. The formal functionality specification lays a foundation for the analysis and implementation of TPM cloud. We start with the description of several simplified physical TPM commands to be used. Details of them can be found in the TCG TPM specification [1].

TPM_LoadKey(keHandle, pUsageAuth, key): Loads key into a physical TPM and returns a key handle. The parent of key is specified by the handle keHandle and has the usage authorization (or password) pUsageAuth.

TPM_BindV20(key, data): Encrypts data with key.

TPM_UnBind(handle, usageAuth, encdata): Decrypts the encrypted data encdata with the key specified by handle. The usage authorization of the key is usageAuth.

TPM_Extend(index, data): Extends the PCR index with data.

TPM_PCR_Reset(index): Resets the PCR index to its default initial value.

TPM_CreateWrapKey(keHandle, pUsageAuth, usageAuth, indices): Creates a new key having the usage authorization usageAuth under the parent key keHandle with the usage authorization pUsageAuth. The new key is locked by PCRs indices.

TPM_Sign(handle, usageAuth, data): Signs data with the key specified by handle. The key has the usage authorization usageAuth.

The Cloud Key and Virtual TPM Root Key are only used inside the cloud, so in the following we give them a default usage authorization, represented by an underscore _.

3.1 Physical TPM Representation and Access Operations

In the TPM management component, a physical TPM is represented as a tuple (id, ckhandle, rkhandle), where id is an identifier, ckhandle the handle for the Cloud Key, and rkhandle the handle for the Virtual TPM Root Key. Recall that the Cloud Key and Virtual TPM Root Key are loaded into every physical TPM. Let pTPM be a set of physical TPMs managed by the TPM management component. Suppose pTPM includes a TPM (id, ckhandle, rkhandle). Then, we refer to this TPM by pTPM[id], its components ckhandle and rkhandle by pTPM[id].ckhandle and pTPM[id].rkhandle, respectively.

Due to the use of the Virtual TPM Root Key, the management of physical TPMs in this new cloud architecture is simpler. Unlike the representation of physical TPMs in [13], users are not linked to physical TPMs any more in the above TPM representation, since the Virtual TPM Root Key is used as the parent of all virtual Eks and virtual SRks. That is, by using the Virtual TPM Root Key, a user is no longer bound to a particular physical TPM in this new cloud architecture because the SRK of a physical TPM is not used as the parent key for creating his virtual EK and SRK.

The TPM management component provides the following operations for other components to access physical TPMs. The operation schedule() returns the identifier of a physical TPM, which is scheduled to execute a command. The operation execute(id, tpmcmd) executes the TPM command tpmcmd on the physical TPM id. For example, execute(id, TPM_Extend(i, data)) sends the command TPM_Extend(i, data) to physical TPM id for executing.

The operation load(name, id, vsrkpwd, key) defined in Fig. 4 loads key into the TPM id, and returns a handle. key is supposed to be created under the virtual SRK of user name, which has the usage authorization vsrkpwd. Briefly, this operation first loads the virtual SRK of user name into the TPM id, and then loads key into the same TPM. When loading the virtual SRK, the TPM id uses the Virtual TPM Root Key on it as the parent key. The virtual SRK is then used as the parent key to load key. The virtual SRK vsrk is retrieved through the operation getVTPM(name).keyhrk.srk, which is described later. Compared with the load operation in [13], the operation in Fig. 4 is simpler since we do not need to rewrap key on the source physical TPM. The Virtual TPM Root Key is loaded into every physical TPMs, so we can directly load key and its parent virtual SRK into the target physical TPM.

3.2 Representation of Users and Virtual TPMs

A user is described by a pair (name, vtpm), meaning that the user name has the TPM instance vtpm. A TPM instance vtpm is represented by a tuple (ownerauth, state, keyhrk, {pcrs}pubek), where ownerauth is the owner authorization of the TPM instance, state its state, keyhrk the key hierarchy created for this TPM instance, and pcrs a set of PCRs encrypted with the public virtual EK pubek.

For a TPM instance, we consider only the ownership state for simplicity. That is, the state field in vtpm is a boolean to indicate whether the command TPM_TakeOwnership is executed or not. Other state flags

```java
load(name, id, vsrkpwd, key)
1. rkhdl = pTPM[id].rkhandle;
2. vsrk = getVTPM(name).keyhrk.srk;
3. vsrk = execute(id, TPM_LoadKey(rkhdl, vsrk));
4. keyhrk = execute(id, TPM_LoadKey(vsrkhdl, vsrkpwd, key));
```
can be supported similarly. For example, we can add other boolean-valued fields to indicate whether a TPM instance is enabled or activated.

The key hierarchy in a TPM instance is described by the tuple \((ek, srk, (\{key_1, handle_1\}, \ldots, \{key_n, handle_n\}))\), including the virtual EK \(ek\), the virtual SRK \(srk\), and a set of pairs of \(key_i\) and \(handle_i\). The key \(key_i\) is created under \(srk\), as shown in Fig. 3. A key is a pair \((usageauth, prikey)\) \(_{vek}\).pubkey, pubkey\), consisting of the public key \(pubkey\) and a blob of private key \(prikey\) and its usage authorization \(usageauth\) encrypted with the public key \(srk\).pubkey. Hence, the private key is protected when stored in a virtual TPM.

Let \(U\) be a list of users. The operation \(getVTPM(name)\) returns the TPM instance owned by the user \(name\). That is, if \((name, vtpm) \in U\), then \(getVTPM(name) = vtpm\). Given a TPM instance \(vtpm\), its four fields are referred to by the notations \(vtpm\.owner\), \(vtpm\.state\), \(vtpm\.keyhrk\) and \(vtpm\.encpcrs\). For the key hierarchy in \(vtpm\), the virtual EK and the virtual SRK are accessed by using the notations \(vtpm\.keyhrk\.ek\) and \(vtpm\.keyhrk\.srk\), respectively. Other keys in the hierarchy are referred to by \(vtpm\.keyhrk\.keys\). We use \(vtpm\.keyhrk\.keys[\{handle\}]\) for the key indexed by the handle \(handle\), and \(vtpm\.keyhrk\.keys[\{key\}]\) for the handle of \(key\). That is, if \((key, handle) \in vtpm\.keyhrk\.keys\), then \(vtpm\.keyhrk\.keys[\{handle\}] = key\) and \(vtpm\.keyhrk\.keys[\{key\}] = handle\). For a key \(key\), its public key and private key are accessed by \(key\.pubkey\) and \(key\.privkey\), respectively.

A new user or a new TPM instance is created by using the registration operation \(register(name, passwd)\) in Fig. 5. The operation starts by creating a virtual EK \(ek\) under the Virtual TPM Root Key on TPM \(id\). We do not require PCRs to lock \(ek\), so the last argument of \( TPM\.CreateWrapKey\) is an empty PCR index list \([\]\). Next, the initial PCRs \(initpcrs\) (an array with each entry initialized as 20 bytes of 0s) is encrypted by the virtual EK \(ek\.pubkey\). At last, a tuple for the new user is added into \(U\), where the password is stored after hashing. Since the ownership of the new TPM instance is not taken, the ownership state is false and accordingly the virtual SRK is not existing (represented by an underscore \(_{\_}\)).

### 3.3 Simulation and Rewriting of TPM Commands

The implementation of cryptographic operations (e.g., symmetric encryption) is straightforward. In this section, we describe how the TPM commands that are dependent on PCRs are specially treated. The commands \( TPM\.Extend\) and \( TPM\.CreateWrapKey\) are taken as examples.

The command \( TPM\.Extend\) is simulated since it needs the actual PCRs values. The simulation in Fig. 6 follows firmly the semantics of the physical \( TPM\.Extend\) command. In the simulation, the virtual PCRs of user \(name\) is first decrypted with the virtual EK by calling the command \( TPM\.UnBind\). Next, the concatenation of virtual PCR \(index\) and data (i.e., \(pcrs[\{index\}]\)) is hashed, resulting in the new (or extended) virtual PCRs. At last, the extended virtual PCRs are encrypted with the public virtual EK of user \(name\) and put back to the TPM instance.

The operation \(createkey\) in Fig. 7 creates a key under the virtual SRK of user \(name\) and returns a handle \(khdl\). The usage of virtual SRK is protected by \(passwd\), and the new key has the usage authorization \(usageauth\). Different from \(pcextend\), the operation \(createkey\) relies on the corresponding physical TPM command \( TPM\.CreateWrapKey\) to create keys within a physical TPM, such that the private keys are not released outside physical TPMs during generation. There are two cases of invoking \( TPM\.CreateWrapKey\). If the argument \(indices\) is an empty list \([\]\), meaning that the new key is not locked by any PCR, then the arguments of \(createkey\) are passed to \( TPM\.CreateWrapKey\) directly. Otherwise, the argument \(indices\) is rewritten (as described below) before passed.

```java
pcextend(name, passwd, index, data){
    1. id = schedule(); rhdl =TPM[id].rkhandle;
    2. vek =getVTPM(name).keyhrk.ek;
    3. vekhdl =execute(id, TPM.LoadKey(rhdl, _ vek));
    4. epcrs =getVTPM(name).encpcrs;
    5. pcrs.execute(id, TPM.UnBind(vekhdl, passwd, epcrs));
    6. pcrs[\{index\}] = hash(pcrs[\{index\}]\[data\]);
    7. TPM.CreateWrapKey(name, passwd, usageauth, \[\]);
}

createkey(name, passwd, usageauth, indices){
    1. id = schedule(); rhdl =TPM[id].rkhandle;
    2. vsrk =getVTPM(name).keyhrk.srk;
    3. vsrkhdl =execute(id, TPM.LoadKey(rhdl, _ vsrk));
    4. if indices[\] then
    5.   khdl = execute(id, TPM.CreateWrapKey(vsrkhdl, passwd, usageauth, \[\]);
    6.   else
    7. vek =getVTPM(name).keyhrk.ek;
    8. vekhdl =execute(id, TPM.LoadKey(rhdl, _ vek));
    9. epcrs =getVTPM(name).encpcrs;
    10. pcrs.execute(id, TPM.UnBind(vekhdl, passwd, epcrs));
    11. execute(id, TPM.PCR_Reset(rindex));
    12. for each index in indices do
    13.   execute(id, TPM.Extend(rindex, pcrs[\{index\}]));
    14.   khdl = execute(id, TPM.CreateWrapKey(vsrkhdl, passwd, usageauth, rindex));
}
```

Fig. 6 Simulation of \( TPM\.Extend\).

Fig. 7 Rewriting of \( TPM\.CreateWrapKey\).
We cannot simply put the virtual PCRs specified by  
indicies into the corresponding physical PCRs since PCRs  
can only be extended. Our solution is to use a resettable  
physical PCR rindex to record the current state of virtual  
PCRs indicies. To this end, we reset the PCR rindex and  
extend this PCR with each specified virtual PCR. Then,  
we invoke the physical command TPM_CreateWrapKey  
with rindex as its argument for specifying a physical  
PCR. For the command TPM_Load, it can also refer to  
PCRs when loading a key (the argument of TPM_Load  
for specifying PCRs is not given in this paper). The virtual  
PCRs for TPM_LoadKey are also treated in the same  
way. Hence, when the same virtual PCRs are specified  
for TPM_CreateWrapKey and TPM_LoadKey, the resettable  
physical PCR rindex would have the same value, allowing  
a key locked by the specified PCRs to be successfully  
loaded. Our solution shows an indirect way of using virtual  
PCRs when considering the restriction of physical PCRs.

3.3.1 Limitations of TPM Specification

From the formalization of createkey, we find that the TPM  
specification should allow an extra argument for the com-  
mand TPM_TakeOwnership, which can configure whether  
PCRs are or are not resettable when taking the ownership  
of a TPM. This will bring much flexibility for the usage  
of PCRs in various scenarios. For example, a user of a PC  
might want all PCRs in its TPM not to be resettable, while  
the TPM cloud wants all PCRs to be resettable. If all PCRs  
are resettable, the virtual TPM commands using PCRs can  
be implemented more easily.

In addition, the TPM specification requires the com-  
mands TPM_Seal and TPM_UnSeal use nonmigratable keys.  
Since the TPM cloud is only supposed to support migrat-  
able keys for better load balancing, such commands pro-  
vided by physical TPMs become useless. On the other hand,  
the TPM_Seal command binds the sealed data with the se-  
cret value tpmProof, which is only known within a TPM.  
That is, a sealed data cannot be unsealed by another phys-  
tical TPM even if the two TPMs have the same PCRs. If  
the TPM specification could allow TPM_Seal to use migrat-  
able keys to seal data only with the values of PCRs, then  
the physical commands TPM_Seal and TPM_UnSeal can be  
used in the TPM cloud. Moreover, it also makes it possible  
to unseal data when it is sealed to a TPM that is broken.

3.4 Cloud Commands Processing

Cloud commands are encrypted with the public Cloud Key  
PKc. To decrypt a cloud command ecmd, the secure I/O  
module executes the following two steps: selects a physi-  
cical TPM id and then executes the command TPM_UnBind.  

```c
id = schedule();
ecmd = execute(id, TPM_UnBind(pTPM[id].ckhandle, _, ecmd));
```

After a cloud command is decrypted, it is passed to the  
execution planner, where the steps of processing cloud com-  
mands are defined. In the following, we take several cloud  
commands as examples to describe their processing. These  
commands cover the typical uses of TPM cloud, ranging  
from registration of new users, generation and use of keys  
to extension of PCRs.

3.4.1 Preparation of New Virtual TPMs

A new virtual TPM is created when users send a registra-  
tion command cTPM_Register(name, passwd) (i.e., the  
cloud command after decryption). The prefix “cTPM_” is  
used to indicate cloud commands. The implementation of  
cTPM_Register is straightforward. The execution planner  
just needs to call the operation register in the user man-  
gement component with the same arguments.

After a new virtual TPM is created, users need to take  
it’s ownership, like using a physical TPM. The cloud com-  
mand cTPM_TakeOwnership(name, passwd) is for this pur-  
pose. Figure 8 gives the steps to take ownership. The execu-  
tion planner first checks the user password and makes sure  
the ownership of TPM instance for user name has not been  
taken. After these checks, a new virtual SRK is created and  
recorded into the TPM instance. Like the virtual EK, the  
virtual SRK is also protected with the usage authorization  
passwd. Hence, when a user is authenticated to own a TPM  
instance, he can access the virtual EK and SRK without pro-  
viding extra passwords. This design makes the cloud com-  
mmands more convenient to use. At last, the state of the TPM  
instance is changed accordingly to indicate that the owner-  
ship has been taken.

3.4.2 Key Creation and Loading in TPM Cloud

The cloud command cTPM_CreateWrapKey(name, passwd,  
usageAuth, indicies) is used to create a key for user name  
under his virtual SRK. The new key is protected by the  
usage authorization usageAuth and locked by virtual PCRs  
indicies. The steps implementing this command is shown in  
Fig. 9. First, the user name is authenticated by checking the  
password passwd, and then the createkey operation in the  
cryptographic service is invoked with the same arguments.  
At last, the new key is added into the key hierarchy. There  
is no key handle yet (indicated by a underscore _), since the  
key has not been loaded into a TPM.

A key is loaded into a virtual TPM by the cloud command  
cTPM_LoadKey(name, passwd, key), defined in

```c

cTPM_LoadKey(name, passwd, key)[
  1. vTPM = getVTPM(name);
  2. assert(hash(passwd) = vTPM.ownerauth);
  3. assert(vTPM.state = false);
  4. id = schedule();
     rkhdl = pTPM[id].rkhandle;
  5. src = execute(id, TPM_CreateWrapKey(rkhdl, _, passwd[],));
  6. vTPM.keyhrk.srk = src;
  7. vTPM.state = true;
]  
```

Fig. 8 Steps for taking TPM ownership.
cTPM_CreateWrapKey(name, passwd, usageAuth, indices) {
  1. vTPM = getVTPM(name);
  2. assert(hash(passwd)=vTPM.ownerauth);
  3. key = createkey(name, passwd, usageAuth, indices);
  4. vTPM.keyhrk.keys = vTPM.keyhrk.keys U f(key, ...)
}

Fig. 9 Steps for creating keys.

cTPM_LoadKey(name, passwd, key) {
  1. vTPM = getVTPM(name);
  2. assert(hash(passwd)=vTPM.ownerauth);
  3. id = schedule();
  4. hdl = load(name, id, passwd, key);
  5. vTPM.keyhrk.keys[sk] = hdl;
}

Fig. 10 Steps for loading keys.

cTPM_UnBind(name, passwd, hdl, usgAuth, encdata, sk) {
  1. vTPM = getVTPM(name);
  2. assert(hash(passwd)=vTPM.ownerauth);
  3. assert(vTPM.state = true);
  4. id = schedule();
  5. newhdl = load(name, id, passwd,vTPM.keyhrk.keys[hdl]);
  6. data = execute(id, TPM_UnBind(newhdl, usgAuth, encdata));
  7. result = encrypt(data, sk);
}

Fig. 11 Steps for unbinding data.

cTPM_Sign(name, passwd, hdl, usgAuth, data) {
  1. vTPM = getVTPM(name);
  2. assert(hash(passwd)=vTPM.ownerauth);
  3. assert(vTPM.state = true);
  4. id = schedule();
  5. newhdl = load(name, id, passwd,vTPM.keyhrk.keys[hdl]);
  6. result = execute(id, TPM_Sign(newhdl, usgAuth, data));
}

Fig. 12 Steps for signing data.

The cloud command cTPM_LoadKey(name, passwd, key) extends the virtual PCR index with data for user name. The implementation of this command is straightforward, as shown below. First, the user name is authenticated. Then, the pcrextend operation provided by the cryptographic service is invoked with the same arguments.

```
assert(hash(passwd)=getVTPM(name).ownerauth);
pcrextend(name, passwd, index, data);
```

4. Applications of The TPM Cloud

We have implemented a prototype of the TPM cloud to demonstrate its functionality. Currently, physical TPMs are usually embedded into computer motherboards, with one computer having at most one physical TPM. Our prototype uses a cluster of Dell Optiplex GX620 to provide physical TPMs. The TPM functionalities on these computers are exposed as web services, and they are accessed by the TPM management component with SOAP. The cloud commands are also exposed as web services, such that the TPM functionality provided by the TPM cloud can be easily accessed. The web service engine we used is Apache Axis2 [15]. TCG advocates that trusted computing technology can be used to improve the security of computer systems (e.g., improved authentication and network access control) [1]. In this paper, we apply the TPM cloud to the implementation of the Needham-Schroeder public-key protocol (NS protocol) for web authentication, as it is a simple and well-analyzed authentication protocol. As a result, the TPM functionality is incorporated into high-level applications, useful to improve the acceptance of TPM functionalities by users.

To use the TPM service, users need to provide their user names and passwords, as shown by the cloud commands. When using the TPM service on a public computer, the malware there might steal the passwords and thus can access the secrets of users in the cloud. This problem can be solved for instance by authenticating cloud commands with some second factor (e.g., smartcard issued by the TPM cloud provider). Hence, when just keeping one second factor for the cloud rather than many second factors for different web servers, a user can securely exploit the cloud to run strong protocols to authenticate him to many different web servers.

3.4.3 Data Decryption and Signing

The cloud commands using private keys for data decryption and signing are defined in Fig. 11 and Fig. 12, respectively. The command cTPM_UnBind(name, passwd, hdl, usgAuth, encdata, sk) decrypts encdata for user name by using the key hdl, which is protected by usgAuth. The decrypted result will be encrypted with the symmetric key sk. The symmetric encryption is done by using the encrypt operation provided by the cryptographic service. Similarly, the command cTPM_Sign(name, passwd, hdl, usgAuth, data) signs data for user name.

These two commands have the same first five steps, in which the key with handle hdl is reloaded into a physical TPM, resulting in a new handle. The reloading is needed because a physical TPM has limited slots to accommodate loaded keys and thus a previously loaded key might have been unloaded by the physical TPM management component. Hence, we need to reload it without users noticing their key have been unloaded. After a new handle is generated, the physical TPM command TPM_UnBind and TPM_Sign are used to do the actual decryption and signing since private keys must be used within physical TPMs.

3.4.4 Extension of Virtual PCRs

The cloud command cTPM_Extend(name, passwd, index, data) extends the virtual PCR index with data for user name. The implementation of this command is straightforward, as shown below. First, the user name is authenticated. Then, the pcrextend operation provided by the cryptographic service is invoked with the same arguments.

```
assert(hash(passwd)=getVTPM(name).ownerauth);
pcrextend(name, passwd, index, data);
```
In the following, we introduce the implementation of NS protocol for web authentication based on the TPM cloud. Suppose that user A and the web server B have registered TPM instances in the cloud and created their RSA keys there. Let PK_A and PK_B be the public keys of A and B, respectively. Then, the NS protocol is run by exchanging the following three messages m1, m2 and m3. The two nonces Na and Nb are random numbers generated by A and B, respectively.

\[
\begin{align*}
  m_1 & \quad A \rightarrow B : \{Na, A\}_{PK_B} \\
  m_2 & \quad B \rightarrow A : \{Na, Nb, B\}_{PK_A} \\
  m_3 & \quad A \rightarrow B : \{Nb\}_{PK_B}
\end{align*}
\]

We show the web pages used by the user to run the NS protocol. The pages are embedded with the Javascript code that implements the RSA and DES encryption algorithms. Recall that the cloud commands are encrypted with the public Cloud Key, and the result of the cloud command might be encrypted with some symmetric encryption algorithm (e.g., DES algorithm). In this application, the user and the server use the TPM cloud to decrypt the messages in the NS protocol, so that their private keys are always kept inside the TPM cloud.

The first page for logging on the web server is shown in Fig. 13, where the user can click the “Generate m1” button to generate the message m1 and click the Submit button to send this message to the server. The TPM cloud is not used for generating the first message since the public key PK_B is known and the encryption algorithm is provided in the page. However, the server needs the cloud to unbind the received
message since its private key is stored there.

After the server receives the decrypted message, it sends back the second page, shown in Fig. 14. This page shows the second message sent by the server. The user now checks whether the second message is valid. If valid, the user then sends the third message. The “Check m2” button is to check the second message. Clicking on this button will send a cloud command to decrypt the second message after prompting the user to input some necessary information such as the user name and password. If the decrypted message contains correct nonce Na, a popup window will say the server is authenticated. And then, the user generates the third message by clicking the “Generate m3” button and sends it back by clicking the submit button. Upon receiving the third message, the web server uses the cloud to decrypt it and if the decrypted message contains correct Nb, then it replies a page to tell the successful authentication of the user.

This application demonstrated that the TPM functionality can be extended to high-level applications by using the TPM cloud, which provides a way of improving the usability of TPMs. To the best of our knowledge, our work shows the first application of TPM functionality to the Javascript programs in Web pages.

5. The Trust Chain Based on The TPM Cloud

In this section, we discuss the chain of trust that can be established based on the TPM cloud. The TPM cloud has the same functionality as a physical TPM. To implement remote attestation, a platform can store its platform measurements remotely into the cloud and then depend on the cloud to sign and report the measurements. In the following, we discuss how to establish the trust of chain based on the TPM cloud.

Suppose there is a platform that executes a sequence of programs $P_0, P_1, P_2, \ldots, P_n$ after its power is switched on. The program $P_0$ is firstly executed, and then it starts the second program $P_1$, which launches the next program $P_2$, until the $P_n$ is started by $P_{n-1}$. Usually, the program $P_0$ is a boot loader. We assume there is a program $P_{net}(0 \leq net \leq n)$, which enables the platform to access network (hence the TPM cloud) after being executed. The establishment of the trust chain is discussed in two cases.

In the first case, we suppose a platform that does not have a physical TPM. For this platform, the trust establishment technique in [2] is not applicable since there is no physical TPM to store and sign the program measurements. By using the TPM cloud, the platform gets the opportunity to build the trust chain. The method is that before $P_{net}$ launches $P_{net-1}$, it measures the program $P_{net-1}$ and the measurement (i.e., the hash of $P_{net-1}$) is extended into the PCRs in the TPM cloud by using the cloud command cTPM_Extend; the chain of trust is established until $P_n$ is measured and extended into the PCRs in the cloud by $P_{n-1}$.

In this case, the programs from $P_0$ to $P_{net}$ comprise the trusted computing base. By comparison, in [2], only $P_0$ (the boot loader) is included in the trusted computing base if there is a physical TPM on the platform. To make the trusted computing base smaller, the program $P_{net}$ should be executed after $P_0$ as closely as possible. For example, the program $P_1$ can be designed to enable network access and then launch the operating system or virtual machine monitor, such that the TPM cloud can be used to build the trust chain by trusting only $P_0$ and $P_1$.

In the second case, the platform is assumed to have a physical TPM. For this platform, the trust chain from $P_0$ to $P_{net}$ is built into PCRs in the physical TPM, as described in [2]. After the program $P_{net}$ is executed, it sends to the TPM cloud the measurements of programs from $P_i$ to $P_{net}$, together with the PCR values quoted from the physical TPM. The TPM cloud then checks the integrity of the measurements based on the signed PCR values. If valid, then they are extended into the PCRs in the TPM cloud. The programs from $P_{net-1}$ to $P_n$ are dealt with similarly as in the first case. In this case, the trusted computing based is only $P_0$ (the boot loader).

In this case, if $P_1$ is a virtual machine monitor, each virtual machine on the monitor is allowed to have their own trust chain even if the platform is shared by all virtual machines. That is, the trust chain for each virtual machine includes the measurements of the virtual machine monitor, the virtual machine, and other programs managed by the virtual machine. This is important for the cloud computing platforms such as Amazon Elastic Compute Cloud (Amazon EC2), where virtual machines usually belong to different users and their trust chains should be separated.

6. Security Analysis of The TPM Cloud

The TPM cloud improves the usability of physical TPMs by virtualizing the physical TPMs in a cloud architecture. In this section, we analyze the security properties of virtual TPMs in the TPM cloud, which is based on physical TPMs to execute all commands involving private keys and other commands such as $\text{TPM}_\text{Extend}$. The analysis is done by comparing them with physical TPMs and the software-based virtual TPMs (vTPMs) [16], which implement TPM functionality as user-space software.

To compare the security property of the various TPM implementations, we first need to determine an attack model. That is, what information an attacker is allowed to know and what operations an attacker is allowed to do. In different attack models, a system may have different security properties. For example, if an attack model allows attackers to tamper hardware rather than software, then the software-based TPM can provide more secure TPM functionality than physical TPMs and virtual TPMs based on physical TPMs in the TPM cloud.

In this section, we use an attack model that is more possible. In this model, we suppose the attacker has installed some malicious software on a platform and thus the attacker can know all contents semantically in the platform memory and change those contents if he wants. To do our analysis, we assume a platform is a machine that has a phys-
Table 1  Security analysis of three TPM implementations.

<table>
<thead>
<tr>
<th></th>
<th>Physical TPMs (pTPMs)</th>
<th>Software-based virtual TPMs (vTPMs)</th>
<th>Virtual TPMs in TPM Cloud (vpTPMs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Key Generation</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Signing with Private Keys</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>Decrypting with Private Keys</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>PCR Reset</td>
<td>✓</td>
<td>✗</td>
<td>✓</td>
</tr>
<tr>
<td>PCR Extend</td>
<td>✗</td>
<td>✓</td>
<td>✗</td>
</tr>
</tbody>
</table>

dical TPM or a software-based virtual TPM, or the platform that implements the TPM cloud. For example, if a key is put into the memory, then the attacker knows it is a key and its value. This situation can happen when the platform is compromised during execution even after trusted booting, known as the vulnerability of time-of-check time-of-use (TOCTOU) [17]. In addition, we assume that the physical TPM can only be handled by using TPM commands and the attacker cannot decrypt a message if the needed key is not known by him.

After determining the attack model, we choose the security properties of TPM functionality to compare. As introduced in Sect. 1, a TPM lies in its capabilities of secure key management, and secure storage and reporting of platform configuration measurements in PCRs. Hence, the secrecy of private keys and the integrity of PCR values are the most important security properties of TPMs.

We compare the secrecy of private keys and the integrity of PCR values with respect to several kinds of TPM commands implemented in the physical TPMs, the software-based virtual TPMs, and the virtual TPMs based on physical TPMs in the cloud. These commands deal with private keys or PCRs. The comparison result is given in Table 1, where the symbol ✓ means the secrecy or integrity property is respected, while ✗ means not. The comparison is explained below and for brevity we use pTPMs for physical TPMs, and vTPMs for the virtual TPMs based on physical TPMs.

For key generation, private keys are secret in both pTPMs and vTPMs since private keys are generated within physical TPMs and when they are moved to the platform memory they are encrypted with the parent keys, which have their private parts only used within physical TPMs. Hence, the attacker has no way to know the private keys in pTPMs and vTPMs. For vTPMs, the private keys are generated in the platform memory, so the attacker can read the generated private keys in our attack model.

When signing or decrypting with private keys, pTPMs and vTPMs do not expose private keys to the attacker, since they are used within physical TPMs. vTPMs expose private keys when they are used by signing or decrypting software since they have to be first loaded into the platform memory.

For vTPMs and vTPMs, the contents of PCRs need to be put into platform memory before processing. Hence, in both vTPMs and vTPMs, the integrity of PCRs values cannot be guaranteed since they can be tampered by the attacker in the platform memory. For pTPMs, the PCR Reset command, which initializes PCR values, can guarantee the integrity of newly initialized PCR values (i.e., they must be zero after initialization). However, the PCR Extend command in pTPMs cannot guarantee the integrity of extended PCR values since it may use an argument that has already been tampered before it is sent from the platform memory into pTPMs.

Although vTPMs are not as secure as pTPMs with respect to the PCR Reset command, we can argue that this is not a problem for security of applications. This is because PCR values always need to be extended after initialization when they are used in applications and the PCR Extend command in vTPMs and pTPMs are at the same security level in our attack model, as discussed above. Apparently, vTPMs are not as secure as pTPMs and vTPMs since the secrecy of private keys are not guaranteed in our attack model.

At last, we mention that pTPMs and vTPMs do not have good usability as virtual TPM in our TPM cloud. The usability problem was discussed in Sect. 1. For example, even if a host has a physical TPM and the correct driver, a Javascript program running in a Web browser cannot access the TPM functionality on the same host due to the interoperability problem between TSS and Javascript, and the security limitation to Javascript enforced by the Web browser. Hence, the NS protocol cannot be implemented with pTPMs and software-based vTPMs.

7. Related Work

The system of vTPMs [16] provides virtual TPMs for virtual machines on a single hardware platform. The vTPMs are implemented in software that is running in the host memory as user-space processes or in a secure co-processor. A vTPM has its own virtual EK and virtual SRK. However, all keys of vTPMs are created not by the physical TPM, instead by software. On the contrary, our TPM cloud depends on physical TPMs to generate and manipulate keys. The benefit is that private keys are always protected by TPMs even when the TPM cloud is intruded by malware. More importantly, the system of vTPM and other work of TPM virtualization [18], [19] cannot address all the usability problems of TPMs discussed in the first section.

A software-based TPM (SW-TPM) [6] is developed for resource-constrained embedded systems because they cannot afford the size and cost overheads of a separate TPM. Similarly, the work [20] proposes the customizable trusted modules for mobile phones to address the resource limitation of mobile phones when using trusted modules. The
TPM cloud can be applied to the mobile embedded systems if they can access network. Using the TPM cloud is also a way to reduce the energy overheads of embedded systems caused primarily by 2,048-bit RSA operations [6].

The work [21] deals with the complexity and inflexibility of TPM functionality by proposing the μTPM architecture. In this architecture, some TPM functionality is implemented outside the physical TPMs. On the one hand, the physical μTPMs do not have to be very complex to implement too much functionality. On the other hand, the functionality outside physical μTPMs can be changed more easily to get flexibility. Similarly, in our architecture of TPM cloud, we move some TPM functionality from physical TPMs to the cryptographic service module.

The Bind attestation service [22] has been used to perform fine attestation of applications such as several distributed computation applications and the implementation of Border Gateway routing Protocol (BGP). The Bind service needs the TPM on the platform running those applications. By redirecting the TPM commands (such as TPM_Extend, TPM_PCR_Reset, TPM_Sea1 and TPM_UnSea1 given in [22]), the Bind service can be used on the platform without TPMs. This also applies to other TPM-based applications like [23], [24]. The remote attestation is a key component to build trusted distributed systems (such as computation grids) [25]. If a host in a distributed system does not have a physical TPM, the trust among hosts can still be built by relying on the TPM cloud, which can store and sign the measurements of the host applications.

The work [26], [27] analyzes how to build trust into cloud computing (in particular, how to build trusted virtual machine monitors). A typical architecture of cloud computing is a virtual machine monitor supporting many virtual machines for many users. So users can apply for computing resources in an on-demand, pay-as-you-go manner. The TPM cloud can be used together with the mechanisms in [26], [27] to attest the integrity of virtual machines in cloud computing. To attest the integrity of a virtual machine, the underlying virtual machine monitor, whose integrity is guaranteed by the mechanisms in [26], [27], can hash the image of the virtual machine before launching it and then extend the hash value into a virtual TPM in the TPM cloud.

The virtual smart cards [28], [29] are proposed as a secure and convenient way of managing private keys. By using virtual smart cards, users store their private keys in the Trust Provider [28] or in the Virtual Smart Card Server [29], which are the only places where the private keys can be used for making signature and decryption. To use their own private keys, users need to authenticate to the the Trust Provider or the Virtual Smart Card Server by using some second authentication factor. The TPM cloud is similar to virtual smart cards in the style of managing private keys and authenticating users with second factors. However, the private keys in TPM cloud are used within the physical TPMs, rather than by software running on a server as in the virtual smart card systems. Moreover, the TPM cloud has other functionality, such as the management of platform measurement with virtual PCRs.

8. Conclusion

TPM has been applied to build trusted systems. However, it has been limited in its use and adoption as we analyzed before. We proposed the TPM cloud to promote acceptance of TPM functionality. From the TPM cloud, users can apply for their own TPM instances on demand. It allows users to benefit from the security properties of TPM without owning a physical TPM. Moreover, the TPM functionality in the cloud is easy for users to access when they move to platforms not owned by them. The TPM functionality also can be easily integrated into applications developed in various languages without the interoperability problem. The embedded devices can use the TPM functionality in the cloud to reduce the energy overheads caused by physical TPMs.

We formalized the functionality of the TPM cloud, and made it clear how cloud commands are executed. Some possible improvements to the TPM specification are proposed to make TPMs more useful in various usage contexts. We demonstrated the functionality of the TPM cloud by using it to implement Needham-Schroeder public-key protocol for web authentication. This example showed that using the TPM cloud is a viable way to incorporate strong security into high-level applications. At last, we discussed the establishment of trusted chains by using the TPM cloud and analyzed the security properties of virtual TPMs in the TPM cloud, concluding that the virtual TPMs in the TPM cloud are as secure as physical TPMs in practical applications in our attack model.

References


John Zic is currently a research team leader and science specialist for Trustworthy Systems within the ICT Centre, CSIRO and holds visiting positions the University of New South Wales and at Macquarie University. He is a co-inventor of the world’s first portable USB-based trusted computing platform to incorporate a TPM cryptographic microcontroller: the patented Trust Extension Device. He has recently given invited presentations and keynotes at the EU FP7 INCO-TRUST workshop in New York, MIT Kerberos Consortium Conference, Vanguard/TTI CyberINsecurity Conference and participated in the 11th Joint EU and Australian Science and Technology Cooperation Committee. John was a member of the Australian Academy of Technological Sciences and Engineering Working Group on Cloud Computing from 2009–2010, advising on privacy, security and trust. Prior to the current position, John was the Research Director for the Networking Technologies Laboratory for a period of two years, during which time he oversaw the second Centre for Networking Technologies for the Information Economy project (CeNTIE 2). This project was funded by the Australian Government through the then Department of Communications, Information Technologies and the Arts (DCITA) that brought in some $3M p.a. into the ICT Centre. John has previously held research positions at Motorola’s Australian Research Centre from 1999 to 2003, and has been a faculty member from 1982 to 1999 at both the University of New South Wales and Sydney University.