School of Mathematics and Systems Engineering

Reports from MSI - Rapporter från MSI

Trusted Computing & Digital Rights Management – Theory & Effects

Daniel Gustafsson
Tomas Stewén

Sep
2004

MSI
Växjö University
SE-351 95 VÄXJÖ

Report 04086
ISSN 1650-2647
ISRN VXU/MSI/DA/E/--04086/--SE

Preface

First we want to thank Pernilla Rönn, Lena Johansson and Jonas Stewén that made it possible for us to carry out this master thesis at AerotechTelub.

A special thank to Lena Johansson, our supervisor at AerotechTelub, for her help with ideas and opinions.

A special thank to Jonas Stewén, employee at AerotechTelub, for his help with material and his opinions on our work.

We also want to thank Ola Flygt our supervisor and Mathias Hedenborg our examiner at Växjö University.

Växjö, June 2004

Abstract

Trusted Computing Platform Alliance (TCPA), now known as Trusted Computing Group (TCG), is a trusted computing initiative created by IBM, Microsoft, HP, Compaq, Intel and several other smaller companies. Their goal is to create a secure trusted platform with help of new hardware and software. TCG have developed a new chip, the Trusted Platform Module (TPM) that is the core of this initiative, which is attached to the motherboard. An analysis is made on the TCG's specifications and a summary is written of the different parts and functionalities implemented by this group.

Microsoft is in the development stage for an operating system that can make use of TCG's TPM and other new hardware. This initiative of the operating system is called NGSCB (Next Generation Secure Computing Base) former known as Palladium. This implementation makes use of TCG's main functionalities with a few additions. An analysis is made on Microsoft's NGSCB specifications and a summary is written of how this operating system will work. NGSCB is expected in Microsoft's next operating system Longhorn version 2 that will have its release no sooner than 2006.

Intel has developed hardware needed for a trusted platform and has come up with a template on how operating system developers should implement their OS to make use of this hardware. TCG's TPM are also a part of the system. This system is called LaGrande. An analysis is also made on this trusted computing initiative and a sum up of it is written. This initiative is very similar to NGSCB, but Microsoft and Intel are not willing to comment on that.

DRM (Digital Rights Management) is a technology that protects digital content (audio, video, documents, e-books etc) with rights. A DRM system is a system that manages the rights connected to the content and provides security for those by encryption. First, Microsoft's RMS (Rights Management System) that controls the rights of documents within a company is considered. Second, a general digital media DRM system is considered that handles e-commerce for digital content.

Analysis and discussion are made on what effects TC (Trusted Computing) and DRM could result in for home users, companies and suppliers of TC hardware and software. The different questions stated in the problemformulation is also discussed and considered.

There are good and bad effects for every group but if TC will work as stated today, then the pros will outweigh the cons. The same goes for DRM on a TC platform. Since the benefits outweights the drawbacks, we think that TC should be fullfilled and taken into production. TC and DRM provides a good base of security and it is then up to the developers to use this in a good and responsible way.

1 Introduction

This master thesis will be carried out on AerotechTelub in Växjö, during the spring 2004. Our supervisor from AerotechTelub is Lena Johansson and our supervisor from Växjö University is Ola Flygt.

1.1 Problem background

Today it is hard to have any control over how PCs are used by home users and companies. Trying to change this, an alliance of several big companies (Microsoft, Intel, IBM, HP and Compaq) created TCPA (Trusted Computing Platform Alliance), which now has become TCG (Trusted Computing Group). This group's goal it to create a TC (Trusted Computing) standard that put more trust into the PC. This is necessary since more and more sensitive information has becomes digital and the range of software attacks on computers is steadily increasing.

TC enables DRM (Digital Rights Management) to run in a secure way. DRM specifies rights for digital content that is enforced by an underlying PKI (Public Key Infrastructure). For instance, rights that a file cannot be printed or copied can be specified for a word document. Rights can also be connected to digital content so it for instance will "self-destruct" after certain amount of time and/or only be used a certain number of times. DRM documents and software can be tied to a specific computer/user. How will this affect the different users?

The different members of the TCG are using different names on this new standard but the most known term is Microsoft's NGSCB (Next Generation Secure Computing Base, formerly known as Palladium). The hardware needed for this standard is built into the computers motherboard and some modification to the CPU also needed. This new standard shows much promise but many users are worried that it will violate their integrity and that it will give the suppliers too much power. This new technology will probably not be an option but unavoidable if you want to buy new hardware in the future.

We choose this problem since there are many rumours and speculations in these areas and we want to make an effort to describe and analyse TC and DRM from an objective point of view. This subject is very important since TC and DRM will affect a wide range of people and not much is known about their effects.

1.2 Problem formulation

There are many questions that need to be answered with these new technologies.
Our main problem is: *How will TC and DRM affect companies, home users and suppliers?*

We divide the main problem into sub problems. The sub problems are chosen in collaboration with AerotechTelub and are questions in need of examination.

- *What new costs, dependencies and consequences will TC and DRM give for its different users?*
- *Will TC stop software piracy, viruses and spam?*
- *What effect will it have on companies specializing on security?*
- *Will companies and users be "forced" into using TC and DRM?*
- *How can TC and DRM violate the integrity and privacy of its users?*
- *How safe is TC and DRM?*
- *What kind of operating system support TC?*
- *Which software and hardware will be "trusted", how and by whom will it be approved?*

- *Will governments get special access to criminal's data secured by a trusted platform and/or DRM?*
- *What happens with open source programs?*

1.3 Realization

The first big part of this thesis is to go through the technical part of TCG (TC and TPM), analysing how it is meant to work and how it is achieved. The next step is to examine some of the current implementations of TC that are underway (NGSCB, LaGrande etc) and examine DRM. The next major part of this thesis is to examine the effects of TC and DRM. After this the main problem and sub problems will be discussed, finally the last part of the thesis is to reach a conclusion based on our results.

1.4 Limitations

We will only go into the PC platform aspect and disregard from cellular phones, PDAs and other platforms.

1.5 Report structure

This thesis is divided in one technical part and one part with more of our own conclusions and standpoints. The first technical part includes chapter 3, 4 and 5 which gives the reader an understanding how TC and DRM works. Chapter 3 describes TCG's TC, chapter 4 describes two TC OSs initiatives and chapter 5 describes DRM (Digital Rights Management). The second part of the thesis includes chapters 6, 7 and 8 that consider and discusses the effects of TC and DRM in different areas. In chapter 6 the effects for different groups of TC and DRM users are considered. In chapter 7 the questions in the problem formulation is discussed and finally in chapter 8 a conclusion on TC and DRM is reached.

2 Method

2.1 Possible methods

[43] In this section we describe different possible method approaches to solve our problem.

2.1.1 Conceptual-analytic approach

In conceptual-analytic research the basic assumptions behind constructs are first analysed. After that, theories, models and framework used previous studies are identified and thereafter, logical reasoning is applied.

This method could be suitable in our thesis. We first want to analyse the theory behind TC and DRM and thereafter use logical reasoning to identify their effects in different areas.

2.1.2 Theory-testing approach

This model tries to the answer the question: Do observations confirm or falsify a particular theory, model or framework?

Since we don't intend to confirm or falsify the theory behind TC and DRM, but identify the effects of these, this model will not be suitable in our thesis.

2.1.3 Theory-creating approach

This model tries to answer the question: Which kind of theory, model or framework best describes or explains a part of reality?

This approach is used to create new theories, model or frameworks and therefore does not suite our problem there we will reason around an undergoing developments of a new technologies.

2.1.4 Constructive approach

This model tries to answer the question: Can we build a certain innovation and how useful is a certain innovation?

This model should be used to evaluate before a new innovation is fulfilled and since we do not intend to construct anything in this thesis this approach is not suitable for us.

2.1.5 Mathematical approach

The mathematical approach is based on the rules of math and isn't suitable at all for our thesis there analysing and reasoning will be used.

2.2 Selected method

We will use conceptual-analytical research approach in this thesis since this is the most suitable for our problem.

We will first analyse the theory behind TC and DRM and present a summary of these theories. Thereafter we will discuss their effects on different areas considering the underlying theory. Finally we will reach a conclusion based on our results in from the effects and discussion.

2.3 Collection of data

Our sources of information will be books, articles, Internet, employees at AerotechTelub and maybe other contacts from different companies. For technical information we will mostly consider the information found from the companies involved in the different projects, because that information will be most trustworthy. We would have liked to use interview and surveys

to gather information but due the general limited knowledge in these areas and due the limited time we decided to rely on written material.

3 TCG (Trusted Computing Group)

In this section TCG will be investigated and an overview will be presented.

3.1 History

[1] One of the biggest issues facing computer technology today is data security. Users are working more and more with sensitive information, while the number of threats is growing and hackers are developing new types of attacks. That is why many technology experts want to develop trusted computing (TC) into the core operations rather than in add-on applications. TC systems would cryptographically seal off the parts of the computer that deal with data and applications and give decryption keys only to programs and information that the system knows is trustworthy. [2] Many big software and hardware companies have worked individually for many years to get more trust available into the PC platform. They all realized that this was a very big and difficult task to accomplish by them selves. In 1999 Compaq, HP, Microsoft, IBM and Intel founded TCPA (Trusted Computing Platform Alliance). They founded this alliance to try to realize a common goal, to put more trust into today's PC platform. TCPA formulate their mission like this;

> "Through the collaboration of hardware, software, communications, and technology vendors, drive and implement TCPA specifications for an enhanced hardware and OS based trusted computing platform that implements trust into client, server, networking, and communication platforms".

Many other companies joined this alliance and there were approximately 200 members in 2003. [3] In 2003 TCG –Trusted Computing Group was founded and they adopted TCPA's specifications and will both enhance these specifications and extend the specifications across multiple platforms such as servers, PDA's, and digital phones. Their mission is formulated as follows;

> "TCG specifications will enable more secure computing environments without compromising functional integrity, privacy, or individual rights. The primary goal is to help users protect their information assets (data, passwords, keys, etc.) from compromise due to external software attack and physical theft."

TCG have currently approximately 40 members but they are expanding all the time. All members from TCPA are encouraged to be a member of TCG as well.

TCG has established operational technical Work Groups for future enhancements on the TC specification. Work Groups for server, PDA, and mobile phone platform specifications will be available soon.

3.2 Overview

In this section a short overview of TC fundamentals will be presented.

Secure boot

If a computer is to be trusted it must be booted in a secure manner. The way TC achieves this is to start from an implicit trusted component (CRTM). This trusted component validates the next code to be executed and passes control over to it if it is approved. This continues until the OS is loaded and holds the trust. This process of passing control is called the chain of trust.

Key management

- Endorsement key (EK). In every trusted platform there is an EK pair that is created during manufacturing of the security chip. EK is the permanent identity of the platform and is used to acquire attestation identity keys (AIK).
- Attestation identity key (AIK). AIKs are used to attest the platform state and configuration to another party without revealing the identity of the platform. A trusted platform can have an unlimited number of AIKs. Many AIKs are necessary to give the user anonymity.
- Storage root key (SRK). In every trusted platform there is one SRK that protects all other keys. The SRK can be seen as the root of a tree where all keys are protected by the key one level over it in the hierarchy.

Trust

There are several roots of trust that are used for starting larger operations. To be able to trust an operation it must be started in a trusted state. In the TCG specification there are three different roots of trust. One for booting, one for protected storage and one for reporting, these are called Core Root of Trust for Measurement, Root of Trust for Storage and Root of Trust for Reporting.

Attestation

A platform can attest itself for another party to prove that the platform is in a trusted state and is running the requested trusted software. This is done to make sure that interaction only occurs between trusted platforms

Attestation is performed by sending the recorded platform state, encrypted with a private AIK, to another party. The other party evaluates if the platform is in a trusted state and if the running software is ok. If the information is approved interaction can begin, otherwise the requested service is rejected. Attestation can be done with the help of TTP (Trusted Third Party) or by sending mathematical proof between two parties.

Protected storage

Protected storage allows the TC user to store arbitrary data on the computers hard disk in a secure manner. Data are encrypted to protect it from unauthorized persons/programs. Encrypted data can also be given certain platform state requirements that must be fulfilled to be able to decrypt.

TPM (Trusted Platform Module)

The TPM is the heart of the trusted platform. The TPM provides key generation; encryption, hashing, key protecting and many more functions that make all TC functions work in a secure way. The TPM is integrated on the motherboard of a computer and can never be moved.

3.3 Examples

In this section a few common scenarios is described to give the reader an understanding of how the different functionality works.

Successful booting and attestation

A trusted platform user starts his computer and looks up his bank web page in Internet explorer. Figure 3.1 is an illustration of the example below.

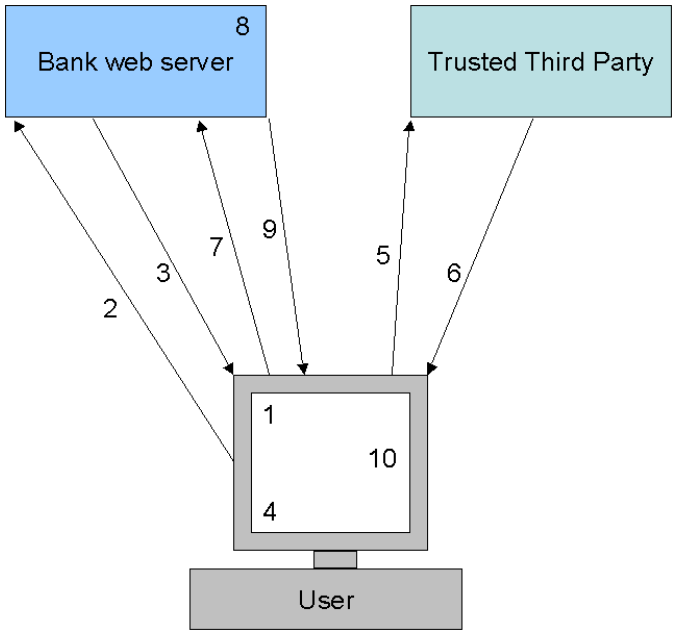

Figure 3.1 Overview of boot and attestation example

1. The user starts the computer with authenticated boot, when completed the computer is in a trusted state. In memory there is a log of all occurred events and in the state registers there are hashed values of the events.
2. The user starts Internet explorer and looks up his banks web page. The new event is added to the log and the state registers are updated.
3. The bank web server who is a TC platform requests attestation from the user.
4. The user's security chip (TPM) generates a new AIK pair.
5. The AIK public key together with endorsement credential and other credentials that proves the platforms origin is sent to a trusted third party (TTP).
6. The TTP evaluates the information and if approved an AIK certificate is returned to the user platform.
7. Now the user's computer sends attestation information (event log and state values signed by private AIK) and AIK certificate to the web server.
8. The web server first decides if the TTP is trusted, then the information is evaluated (log events are redone to test the state values correctness).
9. The server accepts the attestation. Now the server knows that the user platform is in a trusted state and has not been infected by malicious code. The server gives the user the requested service.
10. The user gets the requested web page with a slight delay due to the background security.
11. The user logs into his bank page with his security box (the same procedure as without trusted platform)

Using protected storage

The user wants to have a summary of his accounting on his home computer; he downloads account information from the bank and stores it in a secure manner on the hard drive.

1. The user downloads the accounting information and saves it on the hard drive as an Excel file.
2. To protect the data the user encrypts the information using a symmetric key generated by the security chip.
3. The security chip then protects the symmetric key. Since the user knows that he only wants to open this file on his home computer and only with excel, these requirements are added to the encryption.
4. Now the account information is only available on the users home computer, if excel tries to open it and if the platform is in a trusted state.

3.4 Booting and Measurements

[4][5] TCG security services build on integrity protected boot sequence. TCG supports two different boot sequences, authenticated boot and secure boot. This boot service is essential for start the computer in trusted mode. If any of these boot sequences fails your computer cannot be claimed as trusted and hence will not be able to use the trusted platform implementation. The main idea with the protected booting is that all the executable code and configurations will be measured (hashed and added to state registers) before it is executed.

CRTM (Core Root of Trust Measurement)

CRTM is the place where the booting and measurement process starts running; hence its integrity must be absolutely assured and implicitly trusted. When the computer is turned on a boot process starts in a pre-defined state with the execution of the BIOS boot block code. This block is called the CRTM. It must not be modified in any way for the system still to be considered secure, and a condition is given that every reset must make the processor start executing inside the CRTM. CRTM is updateable by the vendor of it and it is the vendor who is responsible for this part. It is not specified in the TCG specification how this update is performed.

The Chain of Trust

The main security aspect of the TCG specification is to create a totally secure trusted platform. This is accomplished by having the CRTM implicitly trusted. It is in the CRTM the chain of trust is started. CRTM can then claim the BIOS as trusted which in turn can claim the boot loader as trusted which can claim the OS as trusted and the OS can then give trust to applications. Every stage in this chain have to proof that they can be trusted, this is done by measurements and matching. An integrity measurement is taken on each stage; result in a value that is matched against an expected value stored in the platform. If one stage metric fails the matching, then that stage cannot be a part of the chain and will hence not be able to make another stage trusted or even run in the trusted mode of the platform. This process ensures that each part of the platform and software running on it can be trusted and this is established during the boot process.

Regular boot

In an ordinary OS boot a small piece of code in the Boot ROM executes when the computer is turned on. This chunk of code reads in the BIOS code from the boot block and executes it. The BIOS reads in the code for the operating boot process and executes that as well and so

forth until the entire operating system is up running. Then control is passed to the operating system.

Authenticated boot

As we see in figure 3.1, the control is first passed to the TPM that checks the CRTM integrity. CRTM then takes a hash of the BIOS code and stores that value in the Trusted Platform Module (TPM) in a Platform Configuration Register (PCR) and in a full measurement history log. These PCRs cannot be deleted or overwritten within a boot cycle only updated with concatenation of the old and the new values. After the hash is calculated a comparison is made with this hash value to a stored hash value of the BIOS. If the values match, the CRTM passes control to the BIOS code, which will execute. The BIOS then measures the system components, the Option ROM of the peripherals via the TPM and stores these values in the PCRs, and reads in OS loader, calculate the hash and match the values, pass on the control. The OS loader does the same thing with OS and OS with applications. If the code has been altered in any stage, the change will be detected in the hash value, otherwise the user knows that code has not been tampered with and control can be passed on. Those measurement steps, CRTM - BIOS - OS loader - OS, are called: “the chain of trust” (see section above, chain of trust). The operating system can at anytime use the TPM to measure other applications. The figure 3.2 is an illustration of the boot sequence and the associated steps are listed below.

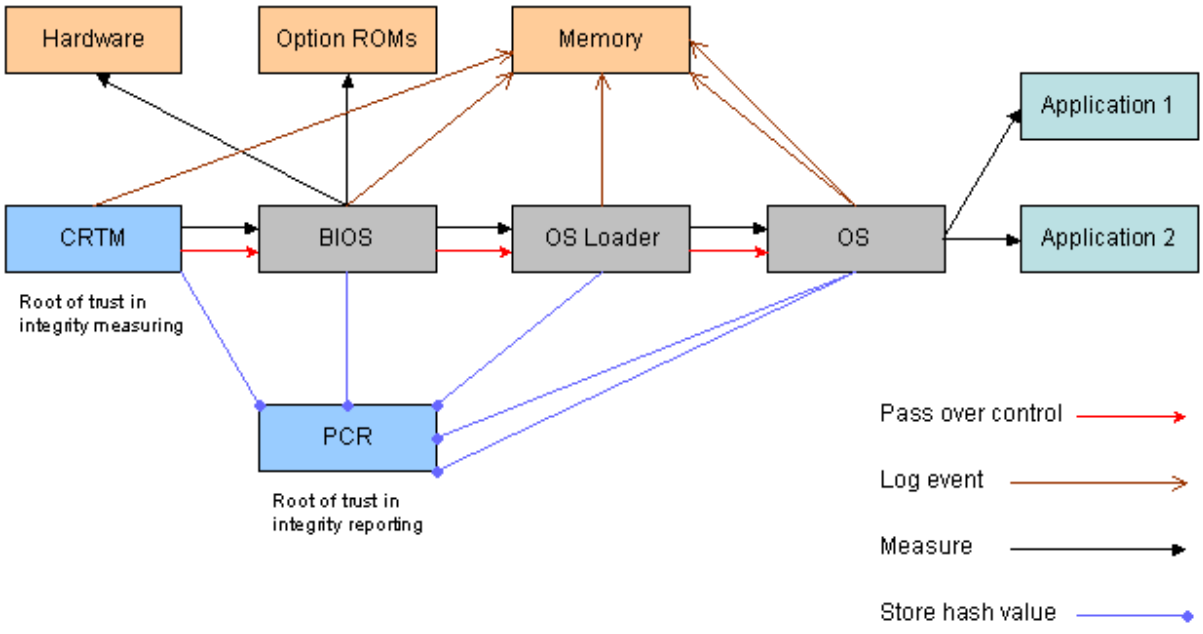

Figure 3.2 Authenticated boot

1. CRTM measures its integrity and stores the value in the PCR and in the log.
2. CRTM measures the BIOS integrity and stores the value in the PCR and in the log, passing control to the BIOS.
3. The BIOS measures the hardware and option ROMs and the boot loader and stores the values in the PCR and in the log, passing control to the boot loader.
4. Boot loader measures the OS and stores the value in the PCR and in the log, passing control to the OS.
5. OS can use the TPM to measure other applications

Secure boot

The secure boot works almost as the authenticated boot except that the platform owner can define expected PCR values that are stored in the TPM. If the hash value in the PCR does not match with the value expected for that stage of the boot process, the boot can be terminated.

All layers along this measurement chain have their own public-private key-pair, which is certified by the layer immediately preceding it in the chain. This in turn is used to certify the layer immediately above it. Each layer signs two things of the layer above it: a hash of its executable image, and its public key. This binds the public key to the software. The succeeding layers are not authorized to read the private key of preceding layers.

That key must be kept secret. Thus, the BIOS must be unable to read the TPM's private key, and so on. After this process is finished with all values matching, a non-tampered trusted operating system is running on trusted hardware.

Storage of integrity metrics

When integrity metric is measured it is stored in a sequence in the PCR. The states of all sequences inside a TPM are set to a known value at power-up. When a new metric is measured it must be appended to a sequence and it must modify the sequence value. The TCG architecture uses a method that concatenates the value of the new integrity metric with the existing value of the sequence. Then the TPM compute a digest of this concatenation, and uses this digest as the new sequence. In this case a sequence could represent many integrity metrics and their updates. The storage process is illustrated in figure 3.3.

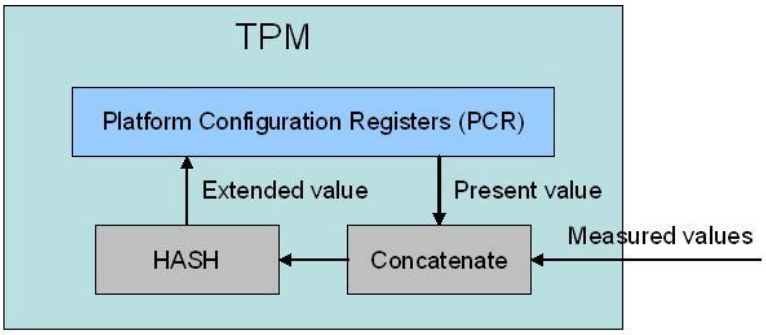

Figure 3.3 The concatenate-hash-storage process for the PCRs

When the measured values are sent to the TPM, they will not be stored one by one because that would demand a lot of memory in the TPM and when an update is made on a value it is not sufficient to just overwrite that value, instead that value will be appended to a sequence.

A log is also used as storage in this process. The values of the integrity metrics are stored in the log while a digest of the values is stored in the PCRs. Each entry in the log inside the Trusted Platform Measurement Store contains a description of a measured entity plus an appropriate integrity metric that has been recorded inside a TPM. The log can then be used to reproduce the value of each sequence of integrity metrics inside the TPM. If the log and the TPM are consistent and the TPM is trustworthy, the log can be trusted as well. If the values derived from the log and the values reported from the TPM are the same, the log can be assumed to be an accurate record of the steps involved in building the software environment of the target platform. The descriptions in the log, the measured entities, represent the actual entities that contributed to the software environment inside the platform. If the values in the log and the values reported from the TPM do not match, then there is an undesirable

inconsistency in the state of the target platform hence it cannot be trusted. A reboot could solve this problem because the measurement process restarts on reboot.

3.5 Root of trust

[6] There are several roots of trust that are used for starting larger operations. To be able to trust an operation it must be started in a trusted state. The three general roots of trust processes in TCG are described in this section.

RTM (Root of Trust for Measurement)

All trust in the measurement process has its origin from this point. It is the component that can be trusted to reliably measure and report to the RTR (Root of Trust for Reporting, see section below) what software executes at the start of the platform boot. The RTM is implemented as the CRTM

RTR (Root of Trust for Reporting)

The RTR is responsible for establishing platform identities, reporting platform configurations, protecting reported values and establishing a context for attesting the reported values. The RTR shares responsibility of protecting measurement digests with the RTS (Root of Trust for Storage, see section below).

The RTM makes reliable measurements about the platform and send measurement results into the RTR.

The RTR prevents unauthorized changes to the measurement results, and reliably reports those measurement results. The RTR must have a cryptographic identity in order to prove to a remote party that RTR messages are coming from genuine trusted platform.

The RTR is a cryptographic identity used to distinguish and authenticate an individual TPM. In the TPM, the EK (Endorsement Key, see section 3.8.4) is the RTR and hence bound to the TPM and the platform.

RTS (Root of Trust for Storage)

The RTS provides protection on data used by the TPM but held in external storage devices. The RTS provides confidentially and integrity for the external blobs (encrypted data or keys) protected by the TPM. In the TPM the RTS is the Storage Root Key (SRK, see section 3.7).

3.6 Attestation

Attestation is the process of vouching for the accuracy of information. By attestation a platform can prove to another party that its TPM is authentic, that the platform is in a secure state and which configuration that is used on the platform.

3.6.1 AIK creation

[5][7] The creation of AIKs is an important part of attestation. It is these keys that give the platform its anonymity as the same time as it can identify itself as a trusted platform. In other words, AIK allows a platform to prove its trusted state to another party without revealing the identity of the platform.

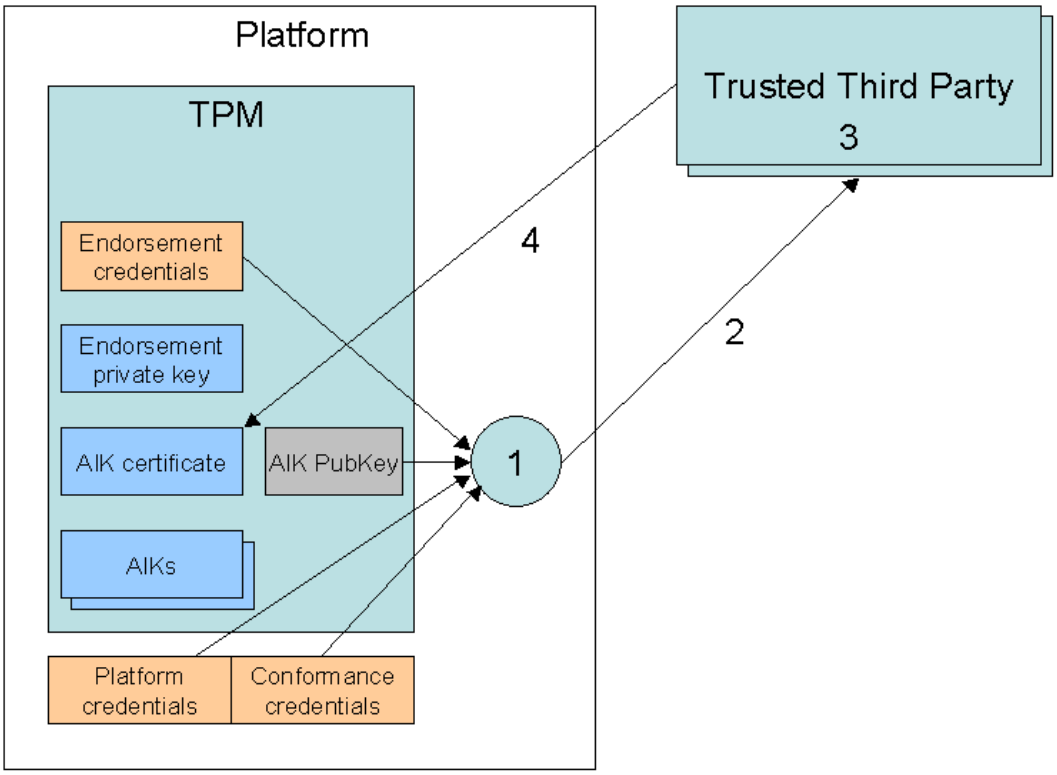

Figure 3.4 Creation of AIK keys.

To describe the creation of AIK we divide the process into four steps (see figure 3.4).

1. First an AIK key pair is generated in the TPM. To get the public AIK verified and signed by a TTP the platform bundles the public AIK (AIK PubKey in figure 1.1) and the endorsement, platform and conformance credentials into a request.
2. To bind the request to the AIK pair the TPM creates a hash of the TTP's public key. This hash is encrypted with the private AIK to create a signature, which is sent with the request to a TTP.
3. When the request is received the TTP verifies the signature and credentials. If everything is approved the TTP creates the AIK certificate by signing public AIK with the TTP's private key. With this AIK certificate, any party that trust the TTP can trust the AIK signed by the TTP.
4. Then the AIK certificate is sent back to the platform, encrypted by the public endorsement key. By using the public endorsement key to encrypt the response from the TTP it is guaranteed that only the TPM that send the request can understand the response. The AIK certificate, which contains public AIK signed by a TTP, can now be used for attestation.

A requirement for the AIK certificates to work is that the PCR values are the same on usage as on creation. In other case a new AIK must be created.

There are three different credentials involved in the process above. The endorsement credential is issued and signed by the manufacturer and contains public EK and TPM model and TPM manufacturing information. The platform credential is issued and signed by the platform manufacturer and contains information of the platform and a reference to the Endorsement credential. And at last there is the conformance credential that is issued and

signed by a credible party (could be the manufacturer) and indicates that the TCG guidelines are followed.

3.6.2 Remote Attestation

[7][8] Remote attestation is used to prove to a remote party that a trusted platform is used and to show which configuration that is run. This is proved with attestation that sends PCR values and corresponding logs to the party requesting the attestation (this process is illustrated in figure 3.5). This aims to allow unauthorized changes to software to be detected, “good” or “bad”. If the user or an attacker has altered one of the applications, or a part of the operating system with new code, not necessary malicious, the user and third party should be able to tell. The interaction between the two parties is as follows:

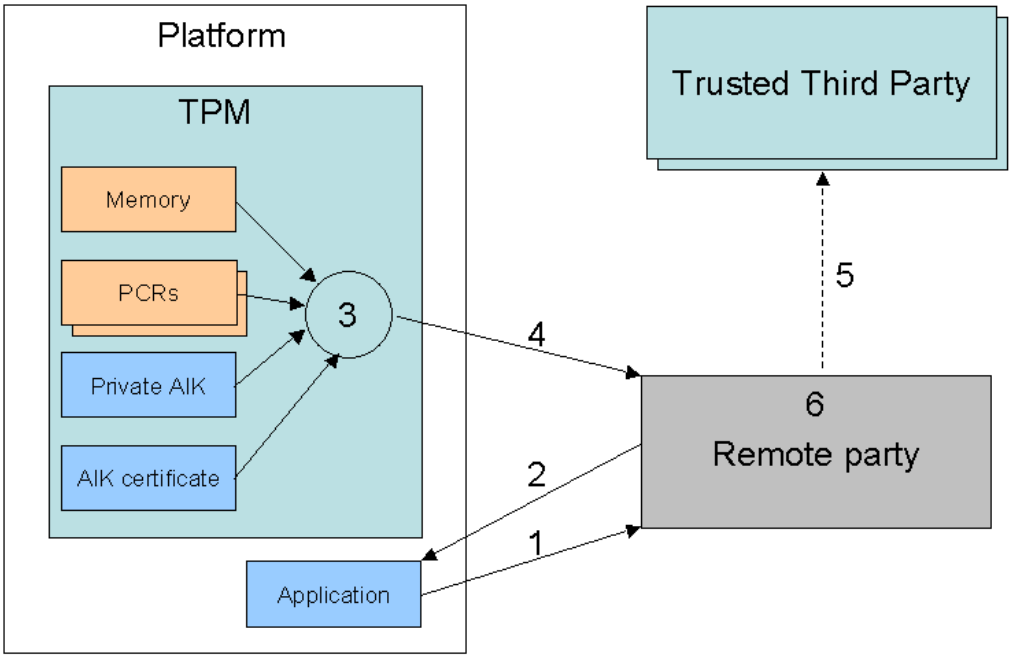

Figure 3.5 Interaction between a remote party, TC platform and a TTP during an attestation.

1. Service requested by the platform owner via an application to a remote party.
2. The remote party asks the requester for attestation.
3. PCRs with corresponding logs are signed by a private AIK.
4. Attestation is sent to the remote party.
5. The remote party examines the AIK certificate signature and evaluates if the signature is from a TTP. Then the signature of the PCR with corresponding logs is inspected.
6. The remote party checks if the requester’s attestation is correct (calculates the PCR values from the log and matches them against the received PCR values) and that the request comes from an approved application. If all is good then the remote party can claim the requester as trusted and continues the interaction.

This process lets the party that requests attestation avoids sending sensitive data to a compromised system or to a non-approved application.

3.6.3 DAA (Direct Anonymous Attestation)

[9] DAA is an option to the regular attestation that could in some manner reveal the identity of a TPM owner. The identity could be revealed by the endorsement credential sent to TTP. To prevent this, a new function is made in the TPM v1.2, the DAA. DAA reliably communicates information about the static or dynamic capabilities of a computer with TPM to a remote party. These capabilities do not require the disclosure of personally identifiable information and is under the control of the platform owner, who can be an individual user or an organization. Users can generate multiple AIKs for interaction with different parties to maintain anonymity; these keys can be used without requiring the involvement of a Trusted Third Party (TTP). DAA can be implemented with Direct Proof Protocol (DPP) alias “zero knowledge protocol” and is based on two TPM commands, Join and Sign.

Direct proof protocol

This protocol is a two-part protocol with a prover and verifier. The prover has some knowledge and is supposed to prove this to the verifier. It provides anonymity without a TTP. Instead of getting an AIK certificate from a TTP, the TPM generates an AIK certificate that can be used. In this case, the public EK does not have to leave the platform.

Mathematical proofs are used to show that the protocol does have the claimed properties.

3.7 Protected storage

[6][10][11][12] The TPM is used to protect arbitrary data and keys with encryption. Some of the secrets can be stored inside the TPM but since the storage there is limited the TPM have the capability to protect data outside of the TPM. The outside storage has the advantages that protected data can be migrated between computers and that data can be backed up in case of a hard drive failure. To achieve the outside secure storage the TCG specify “blobs” of secure data. There are data blobs that can contain arbitrary data and there are key blobs that contain a key that can be imported back into the TPM. Encryption of data bigger than 2048 bits results in a key blob and a data blob that can be stored securely on any media. Aside from the protected data a data blob contains a 20-byte field that can be used for authorization data. For convenience this field has the same size as the output of SHA-1 hash algorithm. This authorization data is used to check for errors after the decryption.

For small pieces of data (less than 2048 bits) the encryption can be done inside the TPM using its RSA engine. With larger pieces of data there are two possibilities:

1. The platform can use a one time symmetric key (not bigger than 2048 bits) to encrypt the larger piece of data and then use the TPM to protect the symmetric key.
2. The second alternative is to divide the big data chunk into several small ones (max 2048 bits) that the TPM can encrypt.

Of these alternatives the first one is generally the most efficient and best.

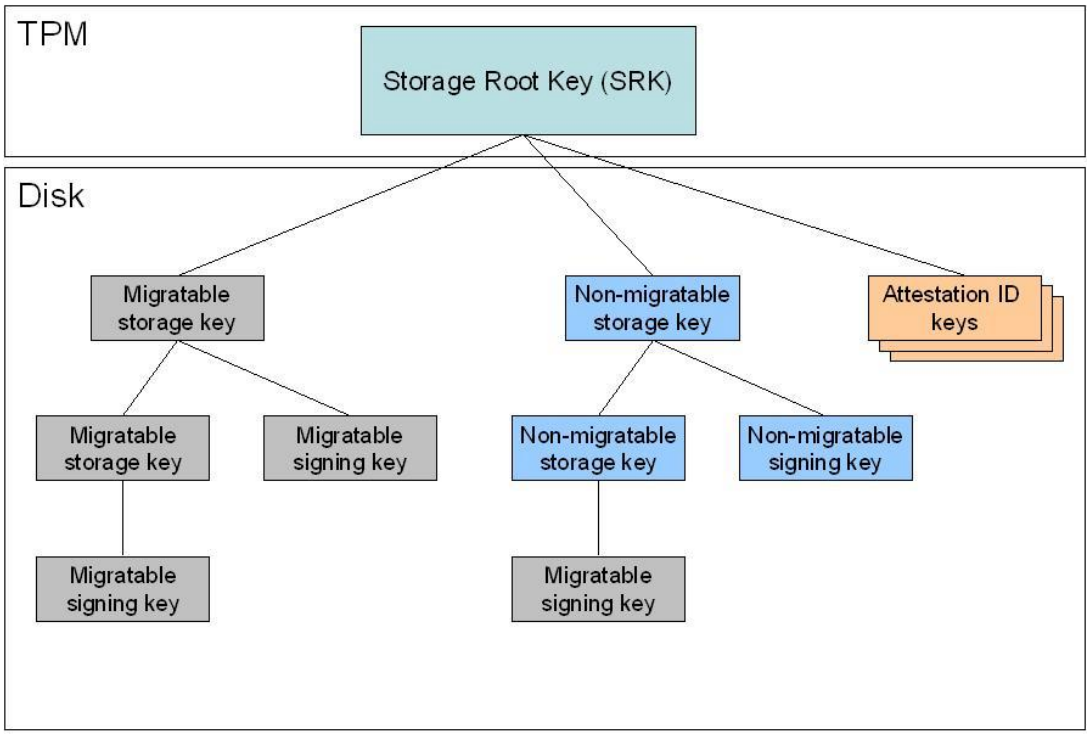

Figure 3.6 The key hierarchy for protected storage.

The TPM implements a key hierarchy for all keys used for protected storage. An example key hierarchy is shown in figure 3.6. The SRK (Storage Root Key) is the root of this tree structure and is the only key that the TPM protects directly. Each key in the hierarchy is encrypted by the key on the succeeding hierarchy level. Only leaf nodes can contain signing keys since the TPM will not use a signing key to protect another node. Nodes that are migratable cannot protect a non-migratable since this makes the non-migratable node migratable. A non-migratable node on the other hand can protect a migratable. This gives rise to two different branches as can be seen in figure 3.6

The TPM supports different types of encryption functions. The most important are Bind/Unbind, Seal/Unseal and WrapKey/UnwrapKey.
Bind uses a parent key to encrypt external data and Unbind uses the other key in the key pair to decrypt the data. Bind/Unbind is simply encrypt/decrypt.

Seal gives some more options. Together with the data, selected PCR and/or a unique TPM identifier (tpmProof) are encrypted. This allows the software to decide that this information must only be opened on this computer since the tpmProof is unique. It also allows the software to specify, with the PCRs, the future state the computer must be in when the data is decrypted. This is a powerful tool and gives the sealer the opportunity to specify in detail what is expected, with PCR values, to be able to unseal. The sealer can choose not to make any demands on the PCR state.

WrapKey is used to protect externally generated keys. The key must not be any bigger than 2048 bits and will be protected by the TPM until it is needed again. It is also possible to include an expected PCR state that must be fulfilled on decryption. The function is called WrapKeyToPcr.

3.8 TPM (Trusted Platform Module)

[13] The TPM has four major functions:

1. Asymmetric key functions for key pair generation using random number generator. Digital signatures, public key encryption and private key decryption of keys gives more secure storage of files and digital secrets. This is done with hardware based protection of symmetric keys, associated with software-encrypted files (data, passwords, credit card numbers, etc.), and keys used for digital signatures. Private keys created in the TPM are always protected, even when they are in use.
2. The protected storage of HASH values corresponding to platform configuration information. These values are stored in PCR (Platform Control Registers) and can be reported in a secure manner for verifiable attestation of the platform configuration.
3. An endorsement key pair (a unique permanent identity key pair that every TPM has) that can be used to establish to another party that the AIKs were generated in a TPM without revealing the identity of the owner. In this way the quality of the AIKs can be confirmed without knowing which TPM created them.
4. Initialisation and management functions, which give the owner the ability to turn functionality on and off, reset the chip and take ownership. These functions must have strong control to protect privacy.

3.8.1 Interoperability

[14] TPM must support at least RSA, SHA-1 and HMAC to follow the TCG specification. More algorithm or protocols may be available to the TPM. All algorithms and protocols accessible in the TPM must be included in the TPM- and platform credentials. There are two reasons for specifying the algorithm. The first is to understand the security properties of the selected algorithm such as appropriate key sizes and use of protocol. The other reason is to specify a base level of interoperability.

3.8.2 Components

[14] In this section the different TPM components will be considered.

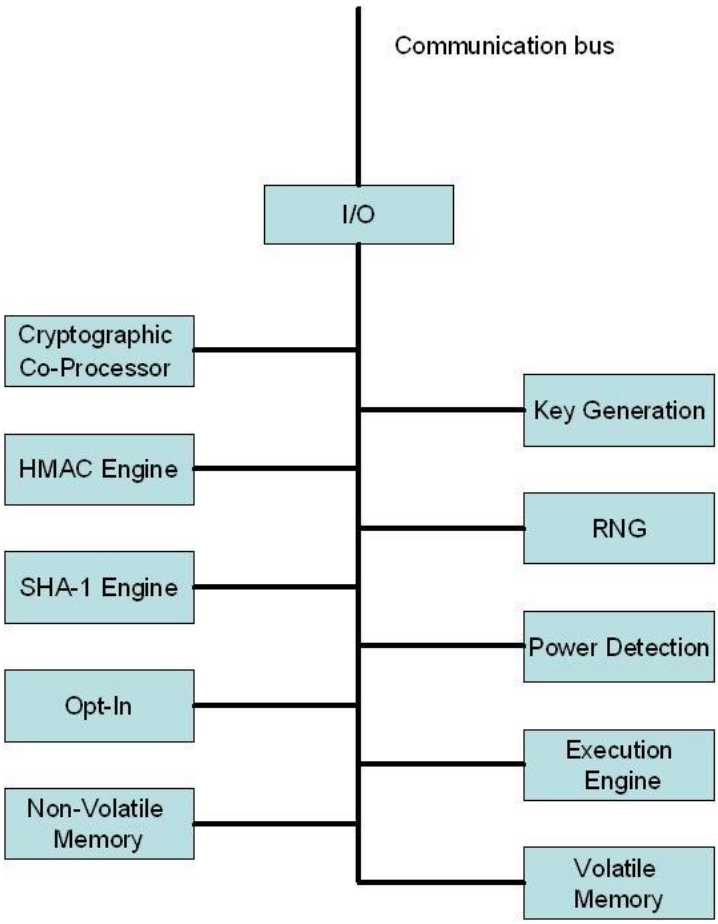

Figure 3.8 Components in the TPM.

I/O

The I/O component handles the flow of information over the communication bus. It is performing protocol encoding/decoding, routing of messages to the correct component and enforcing access policies associated with the Opt-In component (see section 3.8.2.8).

Cryptographic Co-Processor

The Cryptographic Co-Processor implements cryptographic operations within the TPM. The supported operations are:

- Asymmetric key generation (RSA)
- Asymmetric encryption/decryption (RSA)
- Hashing (SHA-1)
- Random number generation (RNG)

These capabilities are used to generate random data, create asymmetric keys, signing and keeping stored data confidential.

Other algorithms such as DES can be added to a TPM but there is no guarantee that other TPMs will understand and can decode the data.

Key generation

In the key generation component RSA key pairs and symmetric key are created. Both RSA signing and encryption key pairs are created here. TCG puts no maximum limit on creation time for the keys.

HMAC Engine

The HMAC engine function is to provide information in two cases. The first is to provide proof of the authorization data and the second is to provide proof that an arriving request is authorized and has not been modified.

RNG (Random Number Generator)

The RNG gives the TPM its randomness. The random numbers are used for creating nonces, key generation and randomness in signatures.
The RNG shall be able to produce random data without a genuine source of entropy since that often is expensive. To achieve this, the RNG consist of a state-machine that accepts and mixes unpredictable data. This is done in the following way: When the TPM is produced a non-volatile storage in the state-machine is initialised with random data. The state-machine can mix this data with unpredictable data from sources such as mouse movements, thermal noise; keyboard keystrokes etc to further improve the unpredictability. Neither the owner nor the manufacture of the TPM can derive the state of the state-machine ones the external unpredictable is mixed with the initial data. The data from the state-machine is run through a one-way function (SHA-1) on the pre-processor in the RNG. In this way good randomised data is produced. What is described above is a PRNG (Pseudo Random Number Generator) algorithm but if a hardware random generator is available the PRNG does not need to be implemented.

The TPM should be able to produce 32 bytes of randomised data at every request while larger request may fail due to insufficient randomised data available.

SHA-1 Engine

Since SHA-1 is a trusted algorithm it is used as the primary hash function in the TPM. The SHA-1 engine allows users outside the TPM to support measurement during boot and to give environment with limited resources the hash capabilities. No minimal requirements on throughput are given by the TCG.

Power Detection

The Power detection component supervises the TPM and platform power state. The TPM must be notified if any change in the power state occurs, according to TCG's specification. In some vulnerable states the TPM will restrict its functionality. An example of this is during the computer POST (Power On Self Test).

Opt-In

The Opt-In component gives the TPM the ability to be turned on/off, enabled/disabled, activated/deactivated etc. This is done by using persistent and a volatile flag contained within the Opt-In. Changing any of these flags requires either the authorization from the TPM owner or the assertion of physical presence at the platform. Which techniques that should be used to represent the physical-presence is up to the manufacturer.

Execution Engine

The execution engine runs program code to execute the TPM commands received from the I/O port. The execution engine is a vital component since it ensures that the operations are properly segregated and that protected locations are shielded.

Non-volatile memory

In the non-volatile memory the state and identity of the TPM is persistently stored.

3.8.3 PCR (Platform Configuration Register)

[14][15][16][17] The PCR is a 160 bits storage location used for integrity measurements. There are at least 16 PCRs and they are in a shielded location inside the TPM. The PCRs can be placed either in volatile storage or non-volatile storage since they are reset on reboot. The first 8 is reserved for TPM use (measuring during boot, etc) and the last 8 is reserved for operating system and applications.

There are many integrity metrics that should be measured in a platform. Some of them may change often and therefore a special method is used for updates. This is done by running the old value together with the new through a HASH and thereby gets a new value that is derived from both values. By using this method the PCR can store an unlimited number of measurements. Since a HASH function is a one-way method it is computational infeasible for an attacker to get the input message given a PCR value.

The function used for updating the PCR looks like this: [PCR] = SHA-1 {[PCR] + Extend value}

3.8.4 EK (Endorsement Key)

[5][11][14] In every TPM there is a unique RSA key pair that is called EK. This 2048 bit key pair is set when the chip is manufactured and is never changed. By the EK a specific TPM is identified, there is one private and one public EK. The private EK is at all times kept secure within the TPM. This is very important since the trust in the platform depends on this key's security and uniqueness. If this private EK is compromised all trust in the computer is lost.

EK are never used in ordinary transactions, instead it is used to obtain AIK certificate from a TTP. This is described more in section 3.6.1.

3.8.5 AIK (Attestation Identity Keys)

[11][14][15] Since the public EK cannot be used in general operations, due to privacy, some other keys are needed. The AIK are of the same size (2048 bits) as the EK but there can be an unlimited number of AIK. AIK are RSA keys created in the TPM and signed by a TTP. The process for this is described in section 3.6.1. The AIK has only one intended use: digitally sign the platform configuration during attestation. The reason to have an unlimited number of AIK is to achieve anonymity. If the same AIK is used on every attestation the anonymity is lost.

3.8.6 Tamper protection

[14] The TCG specifies that the TPM must be physically protected from tampering. This includes binding the TPM chip to other physical parts of the platform. The TPM must be glued to the motherboard in such a way that a removal is evident by visual inspection. This is to make hard to disassembly TPM chip to use it on other platforms.

3.8.7 Taking ownership

[14] In order to be able to use a TPM in an effective way the owner must take ownership. The user takes ownership by showing his/her physical presents by e.g. pressing a key and then type in a password. When ownership has been taken the SRK (Storage Root Key) is created in the TPM.

3.9 TSS (Trusted Software Stack)

[13][18] The TSS provides a standard interface for accessing the functions of a TPM. This facilitates application development and interoperability across different platforms. To make full use of the TPM capabilities applications need to write directly to the TSS. There are some functions that allow creation of interfaces that can interact with existing crypto API.

4 Trusted computing initiatives

In the following section a brief overview of Microsoft's NGSCB and Intel's LaGrande is presented. This to give an understanding of how a future trusted system could look like. Neither one of these initiatives are completed hence there are still some questions unanswered about their functionality. There also exist device drivers and API for TCG's TPM v1.1 for Linux [19][20]. We did not find any more information about an implementation or integration of a TPM and supporting software for Linux; therefore we leave this for future papers.

First we are going to look at the software threats these implementations aim to prevent.

4.1 Security goals

[21] In addition to the goals of TCG, TC-OSs has aims to prevent the following threats:

User output

Attacking software could get access to the graphics frame buffer. This software could then see and/or change what the user sees. It could also be able to take screen dumps.

User input

Attacking software could get access to the keyboard and/or mouse. This software could then see and/or change the inputs made by the user.

Memory

Attacking software could get access to the memory and hence be able to compromise secrets like data, keys, passwords etc. The attacker could alter these secrets and also change the settings on the computer.

DMA (Direct Memory Access)

Attacking software could get access to the DMA controller that in turn has access to protected memory. The attacker could then access the protected memory directly via this DMA controller and compromise the secrets.

4.2 NGSCB (Next Generation Secure Computing Base)

NGSCB formerly known as Palladium is Microsoft's Trusted Computing initiative.
It is a new security technology for the Microsoft windows platform. Microsoft's goal for the NGSCB is to raise the bar on PC system security. NGSCB will be integrated into future Microsoft OS as a subset of the general functionality.

4.2.1 Time aspect

According to Microsoft, NGSCB was first expected in Microsoft's next operating system Longhorn version 1, which will be released in 2006. After some problems Microsoft expects to integrate NGSCB in Longhorn version 2 instead. Microsoft has not mentioned any release date for that version yet.

4.2.2 Overview

[22][23][25][26] [27] An overview of NGSCB architecture can be seen in figure 4.1

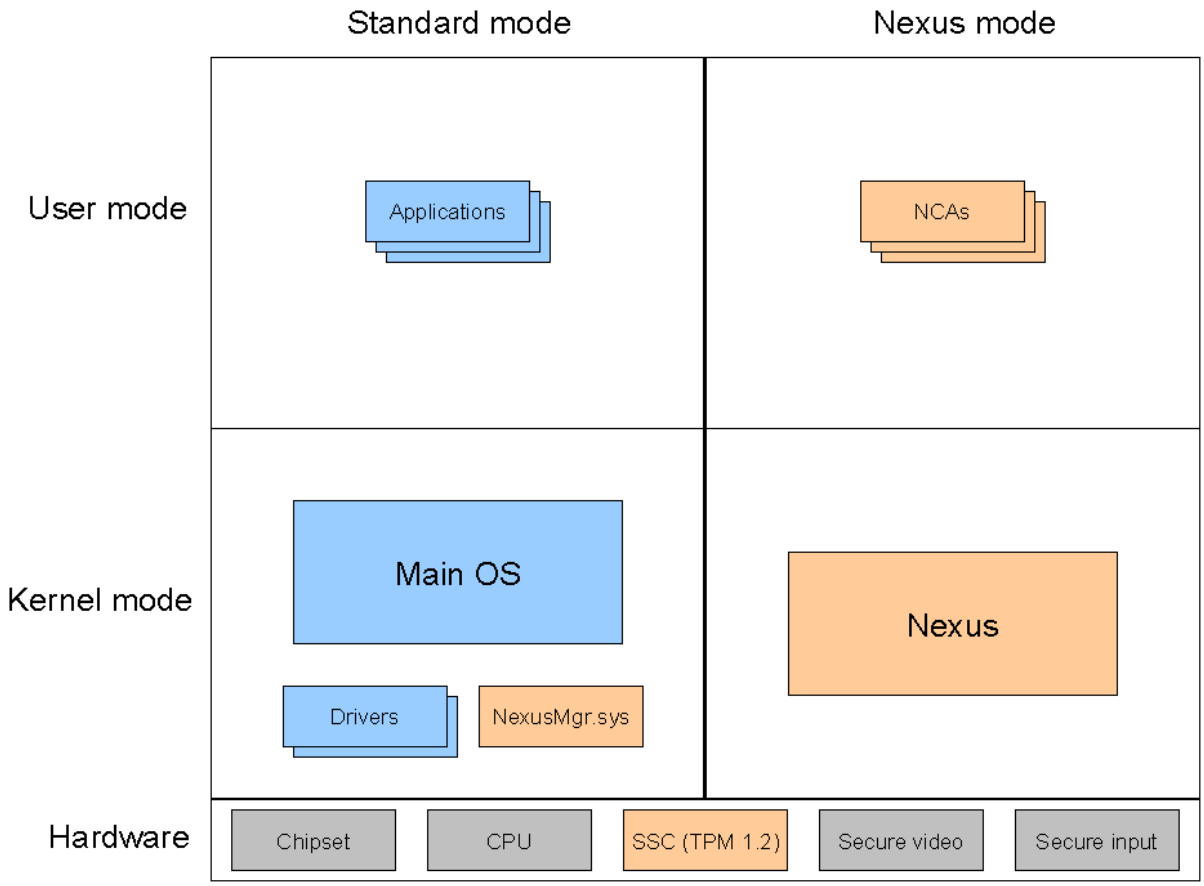

Figure 4.1 The NGSCB architecture.

NGSCB divides the system into four quadrants. In the left side of figure 4.1 main OS and ordinary applications are run. This side is called the standard mode and works as a normal OS does today except for the NexusMgr. NexusMgr is responsible for loading the secure OS kernel and interacting with it for services.

The right side of figure 4.1 is the secure mode. Here all the applications that require a secure environment are executed. The nexus is a small security OS and the NCAs (Nexus Computing Agents) are applications running in secure mode. In the bottom of figure 4.1 the special hardware required for NGSCB are listed.

The nexus is authenticated during start-up. When authenticated, the Nexus creates a protected operating environment within Windows. The Nexus manages security hardware and protected operating environment. When an application wants to run in the protected environment the Nexus provides system services such as starting a NCA. Other functions offered by the Nexus:

- Service to store cryptographic key and encrypt/decrypt information.
- Identification and authentication of NCAs.
- Access control to trusted applications and resources by using a security reference monitor.

- Managing all essential NGSCB services such as: memory management, exclusive access to device memory and secure input/output, and access to non-NGSCB system service.

The Nexus itself executes inside the curtained memory (see section 4.2.4, Strong process isolation). When the Nexus is authenticated successfully it is allowed to access "secrets" by the NGSCB hardware (SSC). These secrets can be used for decryption and authentication for trusted applications. The Nexus also has privilege access to keyboard; monitor etc which gives secure communication between user and Nexus.

Each nexus generates a random key set on first load. The key set is protected by the SSC. NGSCB allows the user to choose which nexus to run and the nexus source code will be open for inspection by anyone. The nexus has a limitation that it should not be any bigger than 100 000 to 300 000 lines of code to keep review and provability.

The Nexus is not a complete operating system kernel; it only implements operating system services that are vital for preserving its integrity. This is to minimize the Nexus size and by relying on the main operating system for the rest of functionality it can be achieved.

The Nexus system components were chosen to give it the ability to guarantee the confidentiality and integrity of Nexus data, even when the main operating system is malfunctioning or has been compromised by malicious code.

In the nexus there is a security reference model that keeps track of the user-configured policies that specifies which permissions have been granted to trusted applications. These policies are enforced using the NCAs identities described in following part.

The Nexus allows NCAs to run in its secure memory. A NCA can be an application, a part of an application or a service. The NCAs runs in the protected operating environment and is isolated from all other NCA and other software except if they have a trusted relationship with a NCA. Within the protected operating environment the NCAs can conceal information and provide user access policies to that information.

Each NCA has a signed XML "manifest" that describes the application. The manifest identifies the program HASH, defines the code modules that can be loaded for each NCA, version numbers and debugging policy. A NCA is authenticated and identified by its code identity, which is computed by the nexus. The identity code is a digest of the application manifest. The user can define policies for different NCA using the identity code.

When an NCA needs a security related service or an essential NGSCB service (e.g. memory management) a request is sent to the Nexus. The digest of the application manifest is used when a NCA decides which other NCA to trust. The NCAs controls its own trust relationship and does not have to trust or rely on each other.

One important thing to note is that all current programs will be able to run in standard mode and NGSCB will not have any affect on this. If developers want software to take advantage of NGSCB the software must be adapted to utilize the protected computing environment.

4.2.3 Hardware requirements

[22] The most important hardware requirement for NGSCB is the SSC (Security Service Component) that is a TPM of version 1.2. Mouse and keyboard that gives a secure path from user to the secure OS and secure graphic adapter are also required. To work with NGSCB the CPU needs an extra mode flag that forces the CPU to either run in standard or in nexus mode. The chipset requires a function to prevent devices from accessing curtained memory used for strong process isolation (see section 4.2.4, Strong process isolation).

4.2.4 Fundamentals

This section will describe the technology of NGSCB divided in four fundamental parts. The four key parts in NGSCB is showed in figure 4.2

Figure 4.2 [24] The four fundamentals of NGSCB

Strong process isolation

[22] In NGSCB a protected operating environment restricts and protects memory spaces for applications and services with high security demands. This secure memory area in an otherwise open operating system is called curtained memory.

On today's computers the RAM (Random Access Memory) is often divided in two sections, the operating system and the user space. User programs running on a computer is located in the user space but can access important functions in the operating system space when they need to. With this current scheme it is relatively easy for an attacker to use both the user and operating system memory space to add malicious programs. To address this problem NGSCB isolates a specific portion of the RAM for curtained memory. An NGSCB addressing-mode bit is set to address this portion of memory, and the bit is added to the NGSCB CPU.

To prevent any DMA (Direct Memory Access) device from accessing the curtained memory directly the NGSCB block all DMA read and write requests to the curtained memory. The blocking is done by a structure called the DMA exclusion vector, which specifies on page level what is curtained memory and therefore cannot be accessed. If information from the hard drive need to be read into curtained memory it must first be read into the operating system or the user space and then be moved by a NCA into the curtained memory. By hiding pages of memory, NGSCB allows trusted applications to run without risk of being tampered with or spied on by unauthorized application or even the operating system. No application in user or operating system memory space can access or even find out that an application exists in the curtained memory due to the RAM blocking. By using curtained memory, isolated from the open hardware and software environment, trusted applications can be protected from software attacks.

Sealed storage

[22] NGSCB provides sealed data storage by using a special SSC (Security Service Component, knows as TPM in the TCG specification). The SSC provides encryption for key and data protection for the nexus. There are public/private key pairs and an AES (Advanced Encryption Standard, known as SRK in TCG specification) key that protects all the keys

lower in the tree hierarchy. When a NCA needs to encrypt data, a key derived from the AES (Advanced encryption Standard) key are used and the storage of the encrypted data is managed by the main operating system. When data is protected the NCA can be sure that only the NCA itself and those who the NCA trust can access that information. To identify applications, the SSC takes a HASH of the applications manifest that works as an ID. Hence the sealer must specify the ID or a list of IDs that are allowed to access the data.

Since protected data only can be accessed if the SSC that protected it is present, the NGSCB provides functions to backup and migrate protected information to other computers.

Data that needs to be stored and protected longer than process durations can also be handled by the SSC in combination with the nexus. Long-lived data can be confidential banking records or some credentials that a browser needs to store.

Attestation

[22][23][27] Attestation in NGSCB works pretty much as described in the TCG section. In NGSCB the attestation can be done without a trusted third party but it is recommended by Microsoft that third party identity service providers be used as an in-between. The use of third party gives users the ability to prove that the platform is trusted without revealing its identity. If third party is not used the RSA platform identity key pair must be used and since this poses a risk, which parties that should be interacted with must be considered carefully. The nexus lets the user restrict the parties to which attestation information should be revealed.
A third option is Direct Anonymous Attestation (described in section 3.6.3) that allows a platform to authenticate itself without TTP and without using the RSA platform identity key pair. Microsoft is currently investigating this approach that would allow two NCAs almost perfect anonymity in AIK creation.

When requested, a nexus can prepare a chain that authenticates:

- Agent (NCA) by digest, signed by nexus
- Nexus by digest, signed by TPM (SSC)
- TPM public key, signed by OEM (Original Equipment Manufacturer) or IT-department

Secure paths to the user

[22] NGSCB provide secure paths from keyboard and mouse to applications and from applications to monitor. To create these secure paths a special I/O device is used to ensure that user data that is transferred between authorized locations without being intercepted. These paths helps protect a computer from keystroke recording programs as well as programs that enable a remote user to act as the legitimate local user.

Since graphic adapter often is optimised for performance instead of security, it is hard to solve this vulnerability. New secure NGSCB output devices will take advantage of advances in graphic adapter technology to protect the video memory. Without this protection software can easily read and write data in video memory.

4.3 LaGrande

LT (LaGrande Technology) is Intel's trusted platform initiative. It is a specification on how Intel's enhancements to their hardware should be used combined with software to run a trusted platform. Any operating system developer can develop their OS to use LT hardware as long as they follow the guidelines on the software that LT gives.

4.3.1 Time aspect

[29] Because LaGrande is a template on how Intel want OS developers to implement OSs' making use of Intel's hardware and no direct implementations have been started, to our knowledge, we cannot give any hint of a release date. Some of Intel's TC hardware is already available and the rest will arrive in one or two years.

4.3.2 Overview

Intel has made security enhancements to their hardware for a more trusted execution. Enhancements have been made on the CPU, memory, I/O devices and a TPM is attached to the motherboard. Except these hardware enhancements LT specifies some special software that is needed to use the new capabilities of the hardware.

As figure 4.3 shows, the LT is divided into a standard partition and a protected partition. The standard partition works as today's OSs without some of the LT enhancements, on the other hand, the protected partition makes use of all the enhancements made.

A choice can be made for an application between running in the standard or protected partition or both. If the application wants to run in the protected partition that partition is booted, a domain manager is loaded for that application which helps it to make use of the LT hardware for a more secure execution.

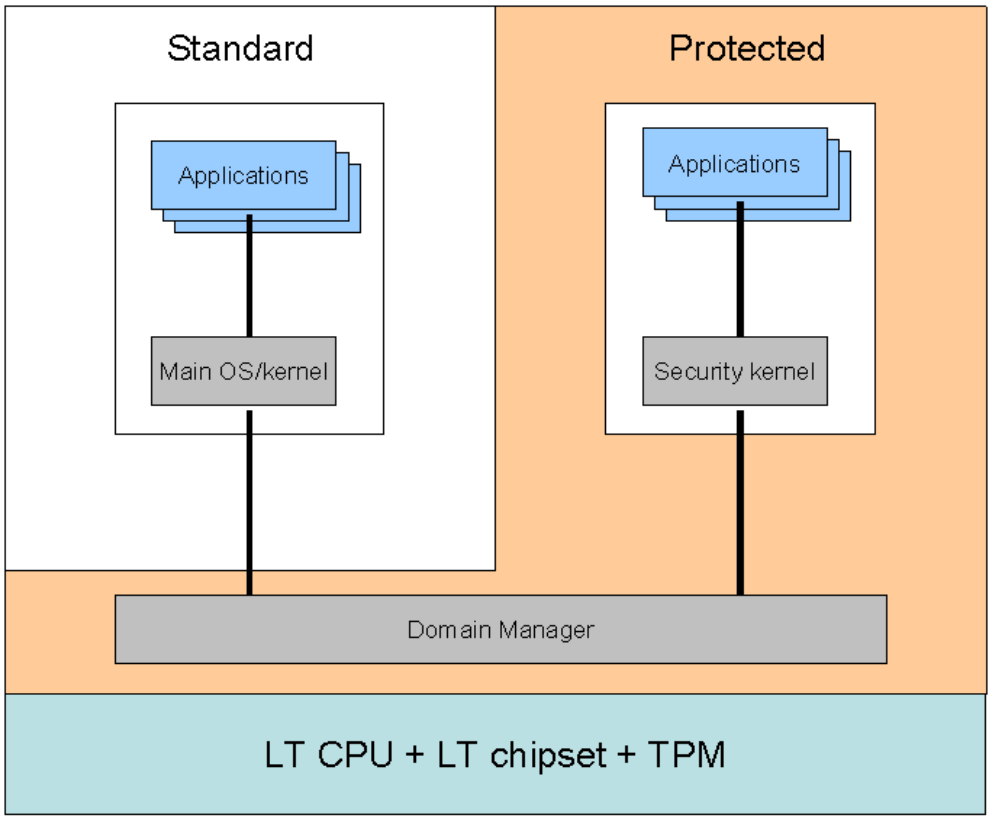

Figure 4.3 LaGrande architecture of a TC platform.

4.3.3 LT objectives

[30] LT objectives are to protect confidential data, communications and E-commerce transactions from attacking software on the system and network. These objectives will be made without compromising ease of use, manageability, privacy, performance, versatility and backwards compatibility.

4.3.4 Key capabilities of LT

[21][30] A platform based on LT technology has a number of key capabilities. When these are combined the platform will deliver a protected trusted environment. The capabilities include:

Protected execution

Applications are able to run in a protected execution environment, this means that no other unauthorized software on the platform can observe or compromise the information used by the running application. The applications runs totally isolated from each other.

Protected storage

The TPM is used to encrypt and store keys, data or other secrets. The secret data are encrypted with a symmetric key along with environmental integrity metrics. If these secrets are going to be released, a decryption key is needed and the environmental integrity metrics to be the same as when the encryption took place. For a more detailed description on sealing and unsealing data see section 3.7.

Protected input

A mechanism is used to protect the communication between the keyboard/mouse and applications running in the protected execution environment from being observed or compromised by any other unauthorized software running on the platform. If the input connection is via USB, all the keystrokes and mouse clicks are encrypted using an encryption key shared between a protected domain's input manager and an input device. Only those applications with the associated decryption key can decrypt and use the data.

Protected graphics

A mechanism is used to let applications running in the protected execution environment send protected display information to the display frame buffer so that no other software can observe or compromise that information. A protected pathway is created between an application or software agent and the output display.

Attestation

Enables a system to provide assurance that the LT protected environment was correctly invoked. It also provides the ability to provide measurement of the software running in the protected area. These measurements with the corresponding log signed by an AIK are used to establish trust between two platforms. This functionality works as in the TCG specification. For a more detailed description see section 3.6.

Protected launch

Protected launch provides for the controlled launch and registration of the critical OS and system software components in a protected execution environment.

AC (Authenticated Code module)

When a user loads the protected partition, the processor loads the authenticated code module into protected memory. This AC detects improper hardware configurations, validates code (e.g. the domain manager) and stores the validation values in the PCRs. This AC can be compared with the measurement component CRTM (see section 3.4) used by TCG.

The chipset manufacturer signs this AC for tamper resistance and integrity reasons.

Domain manager

The domain manager handles the interaction with LT hardware in the protected partition for each application running in that environment. It handles the domain separation and no other domain manager can read data belonging to another domain manager.

4.3.5 LaGrande Technology Hardware Overview

[21][30] An LT platform requires a set of new hardware to be able to operate as expected. The key hardware elements processor, chipset, mouse/keyboard and TPM are illustrated in figure 4.4 and are described below.

A processor is needed that allow the creation of multiple execution environments. This is to give the user a choice between running the software in the standard partition (see section 4.3.6), or in the protected partition (see section 4.3.6), where software can run isolated, free from being observed or compromised by other software running on the platform, or both of these environments.

Access to hardware resources such as memory is strengthening by enhancements in the processor and chipset hardware. The processor have other enhancements such as: domain separation, event handling to reduce the vulnerability of data exposed through system events, instructions to manage the protected execution environment, and instructions to establish a more secure software stack.

The chipsets has support for key elements such as: capability to enforce memory protection policy, enhancements to protect data access from memory, protected channels to graphics and input/output devices, and interfaces to the TPM.

Enhancements on mouse and keyboard will make the communication between these and applications running in the protected environment secure. No unauthorized software will be able to observe or compromise the keystrokes or mouse clicks.

The TPM is bound to the platform and connected to the motherboard. The TPM is used to handle the keys, encryption/decryption, signatures, storage of sensitive data and report platform attestations. For a more detailed description see section 3.8.

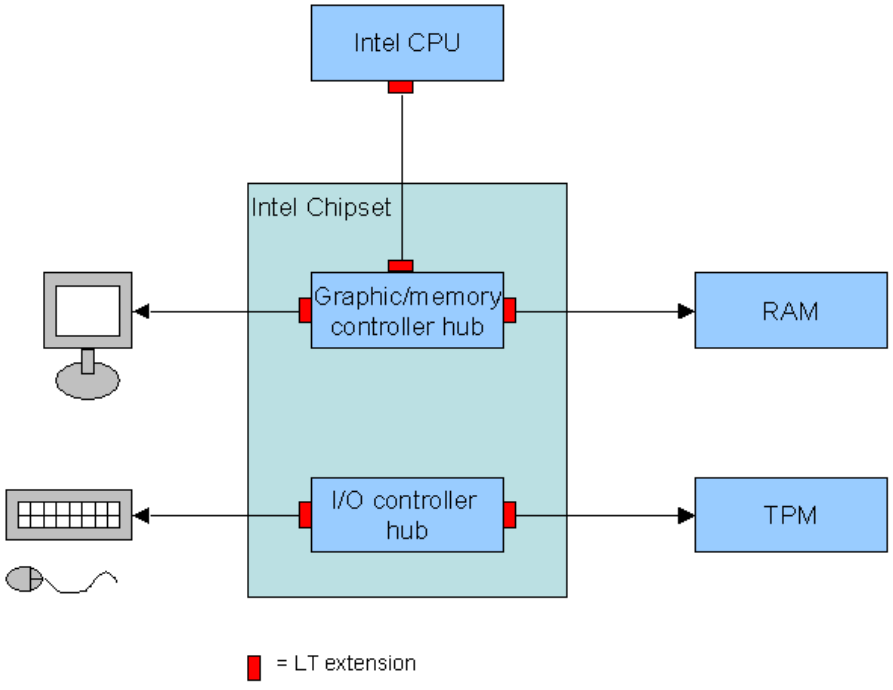

Figure 4.4 LaGrande hardware overview.

4.3.6 Execution environments

[21][30] LT supports applications to run in different environment or partitions depending on which security demands they have.

The standard partition provides an execution environment that is identical to today's OS environments (see figure 4.3). Applications and other software can run as they do today on this partition. This has the advantage that you can run today's software without modification and if creation of new software does not have high security demands, then they can run in this partition (they do not need to integrate support for LT in the software).

The protected partition (see figure 4.3) provides a parallel and co-existing environment that will run LT supported software that makes use of the new LT hardware (see figure 4.4). In the protected partition, applications can run isolated from each other; hence no unauthorized applications can observe or compromise each others events and data. A protected partition requires an LT-capable processor, an LT-capable chipset, and a domain manager to provide domain separation. The TPM protects secrets stored in an LT-enabled platform, even when the protected partition is not running. The LT protection model can support any domain manager and enhanced OS kernel.

Applications can be developed to execute within the standard partition, the protected partition or as in most cases, make use of both the partitions. For instance, if both partitions are used, the code that manages human interfaces and handles I/O could run in the standard partition while security-demanding services could run in the protected partition.

4.3.7 Loading the protected environment

[21][30] Loading the protected environment is needed to use the LT.

Initial trust

Before loading the protected environment initial trust is needed. It is quite hard to attest for initial trust on the local computer but Intel has three different ways to do this, described below.

The user may choose to trust the initial configuration. The user would be advised to create a secret (a phrase or quote) to be sealed to this environment. The user should have confidence that the same environment is running as long as this secret can be displayed to the user on subsequent boots.

A portable device (PD) with support for cryptographic operations can be used as a remote party. It can be loaded with measurements of valid environment configurations at the local retailer and then be connected to the PC. When connected, it can perform attestation of the user system and report a pass or fail.

The user may request that a remote party performs attestation on the user's platform. This on the other hand leaves the problem of how the remote party safely reports this information back to the user. The user cannot trust the environment yet so report back to that environment would not be safe. Instead the remote party should report back by either mail or by using an automated phone menu. A PD can also be used to cryptographically secured protocols to the PD and then let the PD display the result.

Booting of an protected partition

LT have the ability to start the protected partition without the need for platform reboot, and software in the standard partition will still be able to run at the same time as the protected partition boot but they will not have access to the CPU at start up time. The protected partition is invoked by a request to an OS component, which allocates memory spaces for the protected

partition and marks them as “protected”. Then the domain manager is loaded into this memory and registered by an authenticated code module (AC).

The launch of a protected execution environment occurs in stages. These are designed to ensure that the processor and the chipset recognize and participate in the launch, that all participants launch the same environment, and that there is no interference with the launch process. The stages in the launch process can be compared to TCG authenticated boot with the succeeding part measures the preceding one and store the result in a PCR in the TPM. The stages are as follows:

1. Ordinary software running on an LT processor executes a command to initiate the launch process. This command ensures that all CPU activity is stopped; then triggers a sequence of handshakes, and the processor and chipset are brought into a protected environment.
2. The processor loads the AC into internal protected memory, authenticates it, registers its identity in a PCR, and then invokes it. The AC checks that there is no improperly configured hardware, enables memory protection for the proposed domain manager, records the identity of the domain manager in a TPM PCR, then passes control to the domain manager.

Exiting a protected partition

When a user does not need to run the protected environment anymore, LT supports the takedown of it without a platform reboot. This process has the following stages:

1. The domain manager is responsible for cleaning up the protected partitions, ensuring that no secrets are left behind in either memory or registers. Secrets have to be re-sealed and put into persistent storage.
2. The domain manager invokes an exit-command, which triggers a sequence of handshakes that finalize the exit process.

4.3.8 Unexpected events

[21][30] Unexpected events could occur during runtime and even in those cases must the secret data be protected.

If an unexpected system reset occurs, then the CPU and chipset protection would be lost but the memory array may retain content, in this case secrets could be exposed. Re-establish protection may be impossible; hence the page table may be corrupt. LT’s solution to this is to block access to memory on reset or zero out all memory before loading the OS.

When the system is going to sleep, memory protection may be lost because some states remove power to CPU but leave power to chipset. When the system is awake again, the system may be unable to reset the memory protections. LT’s solution is to let the OS requests for a DM (Domain manager) takedown; the DM will encrypt all data in memory and then remove power. On power-up again OS launches the DM, which decrypts the memory pages and continues processing.

4.3.9 LT policy on owner/user choice and control

[31] LT/Intel has a policy on which choices and control options the owner/user must have. This policy includes the following items:

Choice and Control

"LT based platforms must have a straightforward mechanism for ensuring choice in controlling the LT operation".
This means that the PC owner must have a choice whether they want to use the LT protections, and to the degree feasible, maintain control over the various functions.

If a purchaser does not request for the LT capability to be "on", the capability must be "off".

Feature control must only be turned on or off with owner knowledge and consent. The state of the LT must not be changed without owner knowledge.

Visibility

"Users of the system must have clear visibility into the operational state of the LT hardware and LT enabled software"
This means that the system user must have an easy access to information about the state of their LT hardware and the software that is running in the protected partition.

Users must be provided with an interface that reliably identifies the current functional state of the LT hardware.

Users must be provided with an interface that reliably identifies the current state of the protected OS kernel and any protected applications running on it.

Users must have full visibility into attestation information transferred over the network.

Privacy protection

"On the LT platform, privacy protection mechanisms must be made available and must not be associated with any PII (Personally Identifiable Information)".
This means that the high level of security must not be implemented with cost of user privacy.

Unique keys and credentials needed by the LT platform must be protected and used in a privacy-preserving manner. Owners/users must have complete control over the unique keys that the LT creates.

When PII is needed the user must be informed on what information, to what the information is needed and to whom.

PII must only be included in an AIK certificate (see section 3.6, AIK Creation) when the owner of the information provides specific informed consent. PII owners must have the ability to delete any AIK certificate that contains their PII.

5 DRM (Digital Rights Management)

DRM is a technology that protects the digital content with rights and encryption. Digital content can be digital video, audio, e-books, documents, applications etc. In a DRM system digital content can be distributed securely between parties with the rights connected to the content. The rights are described in a special kind of mark up language (there are different options, XrML, ODCL etc).

There currently are not any standard specified for DRM by any standardization agency, hence there are big interoperability problems between the different implementations. There are some initiatives for standardization underway and this is necessary to facilitate for all parties.

5.1 History

[32] Many companies have had problems protecting their digital belongings from being compromised and misused in an illegal way. Secret documents have been copied, printed or emailed. Digital audio, video and e-books have been copied, distributed illegally over the net and software has been hacked so it can run unlicensed. Companies have lost money because of this. To prevent this several DRM methods have been applied:

- Digital watermarking: copyright and other source information is hidden inside the document or audio/video file without the user's knowledge. If a copy is made, the information will still remain in that copy. This technique will not prevent piracy or restrict use, but it can prove authorship and track copies to the original owner.
- Physical copy protection: A physical property of the medium is needed or a requirement of a specific piece of hardware attached to a user's computer, like a USB or smart card.
- Encryption: The content is encrypted and the distribution of keys or certificates is made separately. In some cases a third-party validation is required.
- Product activation: The product comes with an activation/identification code that has to be registered with the publisher before it can run properly.

These DRM techniques have not been that efficient and new techniques to manage the rights on digital content are needed. Most new initiatives are based on certificate-based protection and watermarking plus distribution of rights for each file and user.

5.2 DRM and TC

A common misconception is that TC and DRM goes hand in hand, this is not true. DRM is an independent technology but is strengthening by TC. The TC encryption, secure booting and the ability to detect unauthorized changes put DRM on a higher level of security. Microsoft and other TC companies state that TC was not intended for DRM, but few people have any doubts that they will be combined. Without TC behind DRM, DRM is not secure. The weaknesses in DRM without TC are that there are no secure ways to handle the keys and store the keys on the hard drive, and the applications do not run isolated from each other. "Screen sniffers" and "keystroke sniffers" could compromise secret information as well as other kinds of attacks. TC prevents this by handling and stores the keys in a secure way, applications are running isolated from each other and secure I/O channels are used.

5.3 Overview

The main idea with DRM applications is to use a central server that handles the rights. DRM in a simple form is showed in figure 5.1.

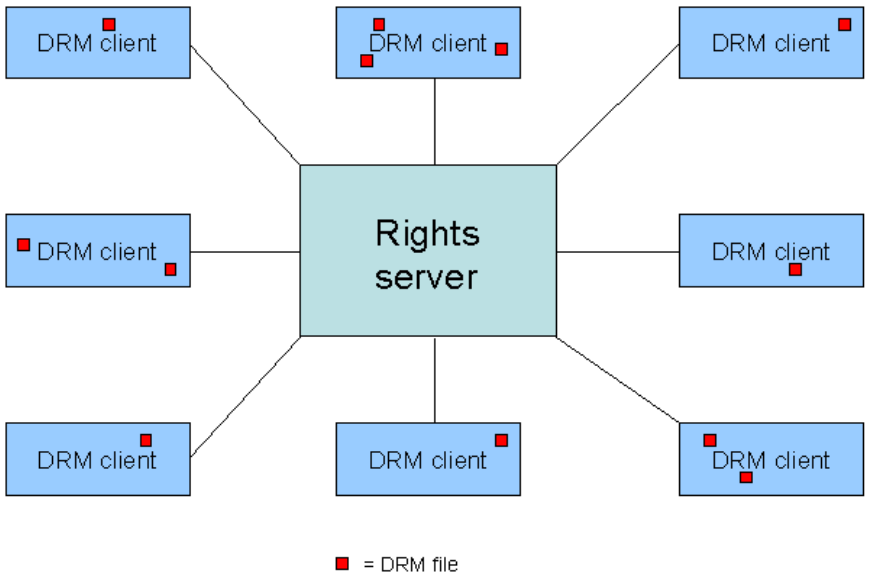

Figure 5.1 DRM system with a central server.

Whenever a DRM client wants to open a file protected by DRM, the server handling the rights must be contacted. The server allows or denies the opening of the file considering the rights attached to the file. The server can be an internal company server or a server on the net. This model is an abstraction and more detailed examples will follow.

5.4 Microsoft's Rights Management Service

As often is the case, Microsoft has many companies ready to back them up on their new technologies. Microsoft DRM has come far and is available both for media over Internet and manage rights for documents inside a company. The DRM system for protecting files and emails inside a company is called RMS (Rights Management System). RMS is a system intended for companies and not home users. In the following sections RMS will be examined.

Needed components

[33] If you want to start using Microsoft RMS in your company, the following products must be bought:

- For server
  - Windows server 2003
  - Database (Microsoft SQL server 2000 or Microsoft SQL server 2000 desktop engine)
  - Windows Rights Management service 1.0
- For clients
  - Windows server 2003 client access license (one per user or computer)
  - Office 2003 standard edition

Office 2003 standard edition can only use RMS documents not create them. To create RMS documents Office 2003 professional edition is required. There are many required components and it will not be cheap to introduce RMS in a company [34].

Microsoft has released software development kits that enable developers to start building applications supporting rights management capabilities.

Rights

Many different rights can be specified for a DRM file. Which groups and users can open, print, edit, forward and copy documents or files. Time based rules can also be added, e.g.: a word document can only be opened between 2004-05-18 and 2004-05-29. Rights are added to a file by the XrML language described in section 5.7.

Creating a DRM file

[33] When a DRM protected file is created in e.g. word, the user specifies the rights inside word with a graphic user interface. The document can be given a certain predefined security level where only some groups can open it and even fewer can edit and print it.

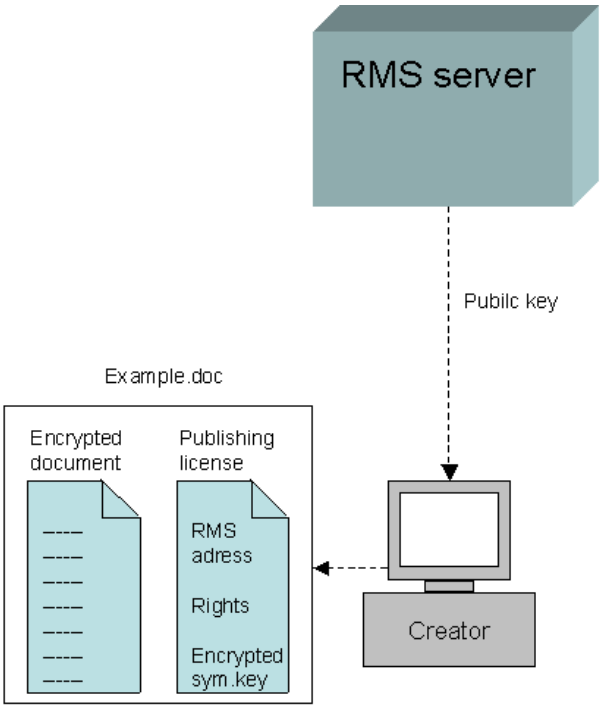

Figure 5.2 Document creations with rights for a RMS system

The steps for creating a DRM document are illustrated in figure 5.2 and are described below.

1. The user specifies the rights
2. A publishing license containing the rights is created by used application together with the RMS client.
3. A symmetric key is created that is used to encrypt the document.
4. The RMS server's public key is fetched and used to encrypt the symmetric key. The encrypted symmetric key is placed in the publishing licence that in turn is placed with the encrypted document.

When the document is DRM protected it still have the same extension.

Opening a DRM file

[33] When opening a RMS protected document the application opening the document must support RMS. Currently this means you must have a windows machine and office 2003. There is a plug-in available for Internet explorer that allows users to open files protected by RMS but for convenient usage office 2003 is needed.

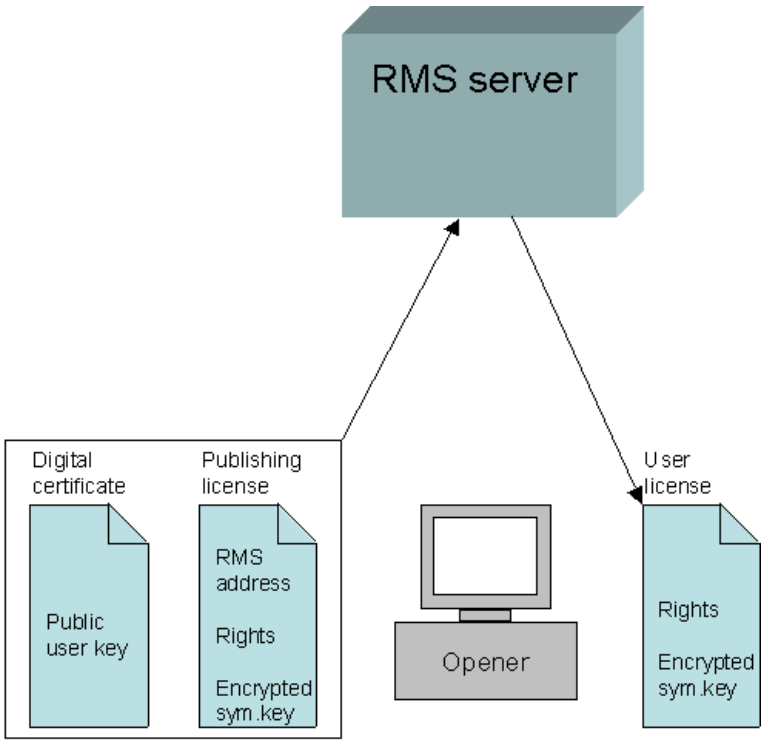

Figure 5.3 Opening of an RMS protected document.

The steps for opening a RMS protected document are illustrated in figure 5.3 and are described below.

1. When trying to open a file the used program discovers that the file is protected
2. The program gets the RMS address from the publishing license. The publishing licence and digital certificate that identifies the user are sent to RMS server
3. The RMS server controls the user and the program trying to open the file.
4. If approved the RMS server decrypts the symmetric key in the publishing license using the RMS server private key. After that, the RMS server encrypts the symmetric key with the user's public key, found in the digital certificate.
5. The RMS server issues a user license, containing the encrypted symmetric key and rights, and sends it to the user's computer.
6. The program running on the user computer gets the user license and decrypts the symmetric key using the private user key. When the symmetric key is in plain text it is used to decrypt the document.
7. The user can use the document as the creator of the document specified. The next time the user opens the same document in the same program, the program sees that a user license is already available and opens the document without any interaction with the RMS server.

5.5 A basic DRM architecture for distribution of digital content

[35] Figure 5.4 presents a basic architecture of a DRM system for distribution of digital content. Most DRM systems today (e.g. ActiveMARK, Electronic Media Management System etc) are using this architecture or parts of it with some modifications.

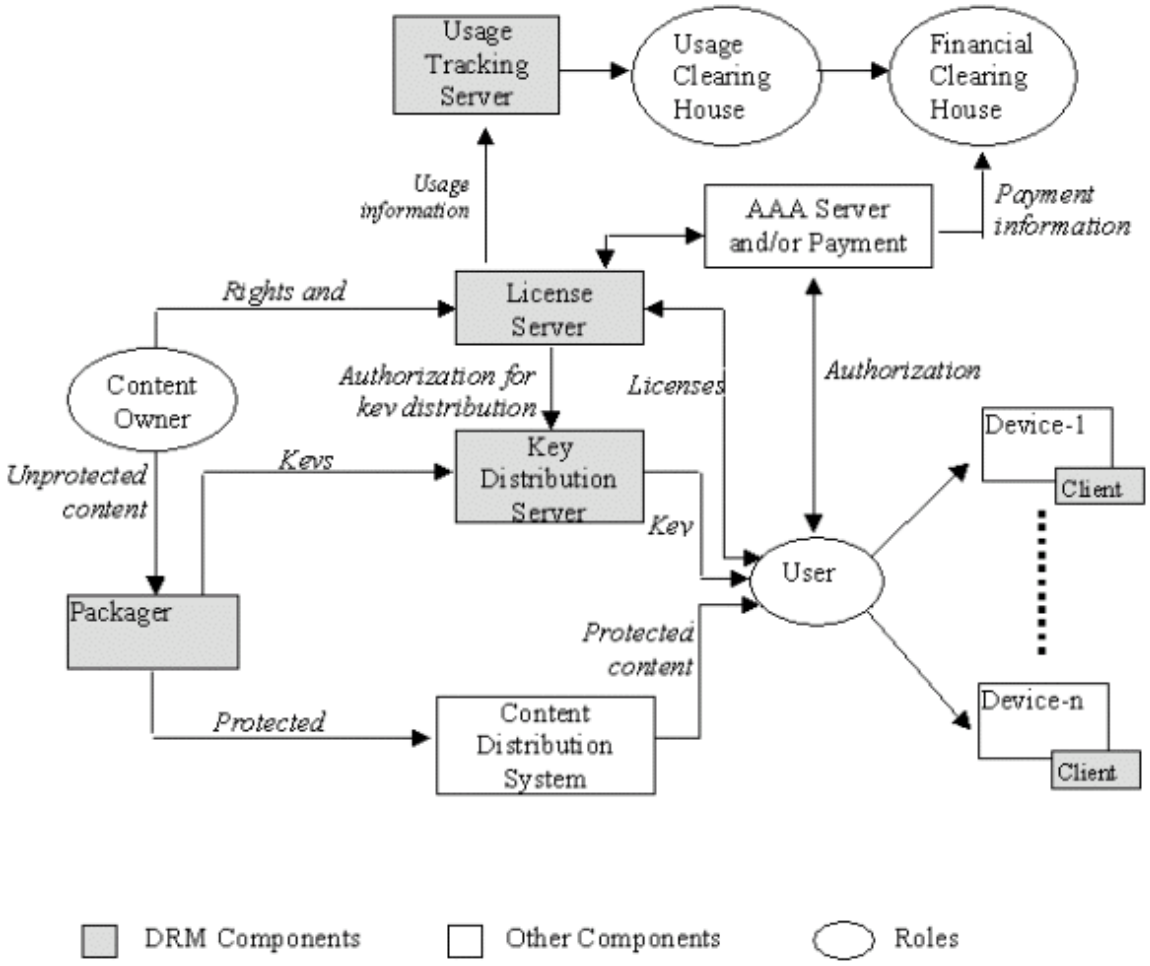

Figure 5.4 [35] General DRM system for distribution of digital content

A user may receive the protected content using any of the methods of delivery supported by the content distribution system. The user then accesses the AAA server for authorization, authentication and access control and/or may have to pay a fee. The AAA server then signals the license server to deliver the license to the user and also authorize the key distribution server to send the keys to the user. The key distribution server delivers the keys when the user begins to use the content. A DRM client is needed in the display device for the decryption and usage control of the content.

5.5.1 Component description

[35] The following components are a part of the basic DRM system. All these components are not necessary in some cases and the architecture may be changed slightly without any efficiency loss.

Content owner

The content owner provides content to the packager. The content owner also provides contracts and rights to the license server. These determine how the content is to be used by a given user.

Packager

The packager can, if it is needed, compress the data, encrypt and watermarking it. The packager gets content as input and outputs keys to the key distribution server and protected content to the content distribution system.

Key distribution server

This server gets keys as input from the packager and outputs these to the users. The license server obtains authorization for key distribution.

License server

The license server gets contracts and rights from the content owner, and outputs licenses to users based on contracts, rights and the input from the AAA (authorization, authentication and access control) server and/or payment. The license server then provides authorization information to the key distribution server.

Content distribution system

This system may distribute the content over the Internet as streaming or download, distribute over physical media or broadcast.

Users, devices and clients

Rights can be associated with users or a device, depending on the implementation, but users association is preferred. Super-distribution (distribution of protected content between users) may be supported.

AAA

The AAA server carries out authorization, authentication and access control. On top or instead, the user may have to pay a license fee for the service. After completion, this server signals to the license server to send licenses to the user.

Usage tracking server

This server gets usage information from the license server, and outputs to the usage clearinghouse.

Usage clearinghouse

The usage clearinghouse keeps information of the usage of the content and is responsible for the distribution of that information to all involved parties (content owner, financial clearinghouse etc).

Financial clearinghouse

It enables financial transactions to be carried out. This covers the collection and contracted distribution to all involved parties (content owner etc).

5.6 Other implementations

For protecting document etc with DRM on a company there currently is not much to choose from. Microsoft solution is the only one available here. On the other hand, there are plenty of different solutions for distributing digital media over the net. Many of these have the weakness that they cannot interoperate.

There are also many different languages for marking up the rights. We chose to present XrML since Microsoft has adopted it.

5.7 XrML (eXtensible rights Markup Language)

[35][36][37] XrML is an XML-based language designed for specifying rights and condition to control digital content and services. XrML has its root in Xerox PARC's DPRL (Digital Property Rights Language), first introduced 1996. DPRL became XrML when the language used to construct the language changed from lisp to XML in 1999. Microsoft has adopted XrML and uses it to specify rights in their DRM solutions.

The XrML data model consists of four entities and relations between those.
The basic relationship is defined by the XrML assertion "grant". An XrML grant is structured as illustrated in figure 5.4.

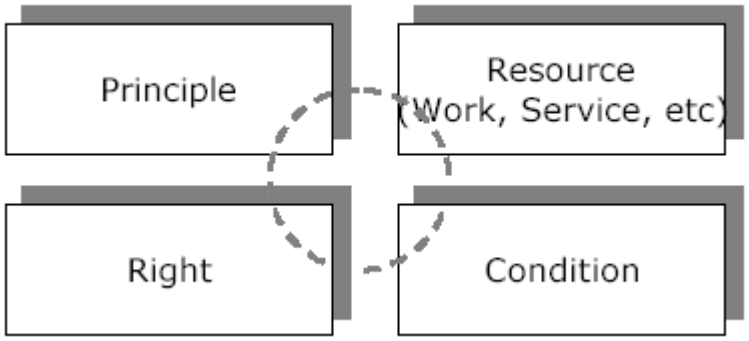

Figure 5.5 [38] The XrML data model.

The entities in figure 5.5 consist of the following:

- Principle specifies by whom the grant is issued.
- The right specifies which rights the grant involves.
- The resource specifies which resources the rights referees to.
- The condition specifies what requirements that must be fulfilled for the rights to be exercised.

To prevent unauthorized changes to the rights associated to a file, a digital signature is used. The signature is checked every time the file is opened.

In figure 5.6 a simple example of XrML is shown where rights for copy, printing, view and extract an e-book is specified. Price per use is also specified in the code sample.

```
<grant>
    <!--The principal who will be allowed to obtain
    the license to copy, print, view, and extract. -->
    <keyHolder licensePartIdRef="issuedToParty"/>
    <obtain/>
    <grantGroup>
        <!--The principal who will be allowed to play,
        print, copy, and extract. -->
        <keyHolder licensePartIdRef="issuedToParty"/>
        <!--The right to play (view) is granted-->
        <grant>
            <cx:play/>
            <cx:digitalWork licensePartIdRef="eBook"/>
        </grant>
        <!--The right to print is granted-->
        <grant>
            <cx:print/>
            <cx:digitalWork licensePartIdRef="eBook"/>
        </grant>
        <!--The right to copy is granted-->
        <grant>
            <cx:copy/>
            <cx:digitalWork licensePartIdRef="eBook"/>
        </grant>
        <!--The right to extract is granted-->
        <grant>
            <cx:extract/>
            <cx:digitalWork licensePartIdRef="eBook"/>
        </grant>
    </grantGroup>
    <sx:fee>
        <sx:paymentPerUse>
            <sx:rate currency="USD">25.99</sx:rate>
        </sx:paymentPerUse>
    </sx:fee>
</grant>
```

Figure 5.6 Example of rights specified with XrML, follow the comments in the figure.

6 Effects

In this section we look at the effects of TC and DRM combined with TC for home users, companies and suppliers.

In this section we refer to TCG, NGSCB and LaGrande as TC. All these three technologies have the same fundamentals.

6.1 Home users

Home users accounts for people using their computer for personal business and pleasure in their homes.

6.1.1 TC

In this section we consider the effects of TC.

Costs

It probably will not be economic defendable for a user to try to update his computer to a trusted computer by adding and replacing hardware, since much of the hardware must be replaced. New motherboards with a TPM, a CPU supporting secure and non-secure mode, secure paths from mouse and keyboard, and a secure graphic device. Converting a non-trusted platform to a trusted will be expensive and cumbersome. On the other hand when a user needs to purchase a new computer the additional cost for a trusted computer will be relatively small. We think that most home users will buy a trusted computer when they need a new one, without knowledge of the additional trusted parts.

To benefits from TC an operating system supporting it must be used. This will probably not be more expensive than those on the market today. Special applications are also needed to make use of the TC functionality and we cannot se any reason why these should cost more than applications today. Note that old applications still can be used on the trusted platform but they will not take advantage of the TC functionality.

Dependencies

There is a possibility that more and more applications will be developed with TC in mind, hence users without TC will get less functionality from the applications than users who has a TC computer. There is also a risk that companies will only develop software for TC platforms that does not work on a regular computer and thereby forces users into using TC. The companies behind TC have interest in getting users to adopt the new technology and specifying TC hardware and OS as a requirement for an application is a good way to achieve this.

If a user wants to run a TC OS he needs the TC hardware and if he wants to run an application supporting TC he needs a TC OS. In other words if you are planning on running trusted applications you have to buy the other parts first.

Consequences pros

- Data/key protection
- Ability to turn TC on and off
- No malicious code in secure mode
- Applications in secure mode is isolated from each other
- User has control over applications access policies
- Lower prices on software

The most obvious gain for a user using TC is the protection of files with encryption and the secure storage of the keys (see section 3.7). Being able to tie a file to a certain program by sealing the data, on his home computer also gives the user more control. He can for example decide that only a certain economy program has access to his protected accounting information. The encrypted file can also be tied to the user's computer by encrypting the file together with a unique identifier (tpmProof), to further control its usage. Normal home users usually will not need to protect that many files but when the need occurs there is secure protection with many options available.

One important aspect about user control is that a computer purchaser may get the trusted hardware without wanting it. It is optional to turn it on; actually the TC functionality is turned off as default and is not turned on until the user demands it. A user can run an old OS and old software on the new computer with trusted hardware. Since the undergoing implementations supports both secure and standard mode a user can continue to run any program he or she wants to, in standard mode. This is important for users. When they convert their current system to a TC system they should still be able to run and use their old applications and data.

Another good thing is that after booting the secure part of the operating system and applications in secure mode the user can be sure of that no hidden malicious code are running on that part of the computer. There can still be viruses in standard mode but these cannot do any damage to applications running in secure mode since they executes in curtained memory. Malicious programs in standard mode can also try to record keystrokes and print screens from applications running in secure mode; the secure I/O channels and secure graphic device prevent this.

Even if a program is implemented for TC it can still mean harm to the system and try to access secret information from other applications in the system. Since an application running in secure mode specifies which other applications it allows interaction with, these kinds of malicious programs cannot do any harm since they can only access its own resources.

The user himself specifies the policies that applications running in secure mode must follow when using sealed secrets. Exactly how this is done is unclear but it gives the user better control over his system.

TC can improve method trying to prevent piracy (see section 6.1.1, Consequences cons); hence the prices on software should go down as more people start buying software. Therefore, decreasing piracy has some good effects for the home users.

Consequences cons

- Some servers may demand a TC platform
- User can be forced into using a specific software when interacting with a TC server
- Lock in for users by tying files to specific application or supplier
- Helping to prevent software piracy
- Users cannot tamper with applications undetected
- Cumbersome to switch computer

A problem for users is if web servers and other servers start to request attestation (see section 3.6, Remote attestation). Firstly this means that a user must have TC platform and TC OS to be served by the server. Secondly the user could loose his freedom in choosing which applications to run, since attestation gives the server information about what application made the request. This application may be unauthorized to interact with the server; hence the server would deny it service. For example can a user surf the web with Netscape when a server asks for attestation, the server sees that the user's computer is in trusted state but does only allow Microsoft Internet Explorer to interact with it and therefore denies the user the requested

service. This is a very serious drawback since it allows companies and governments to decide which software a user must run. This leads to less freedom for the user.

Attempts to control which applications a server interacts with and vice versa, has occurred before, but without TC there has been ways to get around it. E.g. Microsoft tried to force all Windows clients to only interact with a Microsoft servers but the Samba server was made to act like a Microsoft server to get around this. This would not be possible to do if TC were used since attestation would have proved the server was not a Microsoft server.

Another serious drawback for TC users is that the software developers can start tying files to a certain application by encrypting the file with a mutual key. Running the application in secure mode and at the first start up connects to a server and get a key can do this. This key can be downloaded in a secure manner using attestation and PKI (Public Key Infrastructure). When the program gets the key it protects the key by sealing it so only that application can decrypt it. Now the application has a key that encrypts all created files and only applications with the same key can decrypt them. What is described above can for example be used by Microsoft Word to ensure that files written in Word can only be opened by other word applications. This really makes it hard for a user to switch to another word processing program since all old files cannot be opened in the new program. It also means that if user A wants to send a word document to user B, user B must have a word application to be able to read and edit the document. History shows that big companies often want to tie its customer to its products. TC presents an alarmingly good way of doing this. If software companies will implement these kinds of functions is not sure but if it is profitable the risk is imminent.

A similar approach is if the program running in secure mode has to connect to a server at first start up to check that the licensing key is valid and in other case prevent the program from running. Attempts to prevent piracy in this way has occurred before, but the manifest (see section 4.2.2) following a TC application allows identification of any modification of the application and therefore stops any attempts to change the software to circumvent the licence key control. Using piracy software is a doubtful activity but since most home users do so to some extent it must be considered as a disadvantage for the user that it is prevented. Both these suggested scenarios demand an Internet connection when first starting up the application, but this is not an unthinkable requirement in coming applications.

Since a hash of the application or a hash of its manifest identifies the applications running in secure mode, the user cannot tamper with the software if he intends to use them in secure mode. This limits the users control over bought software and restricts him from installing plug-ins and other add-ons not approved by the software developer.

If a user decides to start using another computer or just want to move his files to another computer the process will be more cumbersome and complex than without TC. All files keys protected by the TPM on the first computer must be migrated to the second. There will be support for this event in the platforms (see section 3.7) but how it is done and how complex it is remains to be seen.

The future

When TC is new and it is a big step to convert to it, the option to turn it on and off is available to calm the users and get TC accepted. In the future we think that option will not be available and all software applications must be run in TC mode. This will probably happen when TC has become popular and when many users are depending on it. This because of two aspects; first, many suppliers would gain from it and give them more control. Second, the TCG specification does not say anything else and does not divide up the platform in a standard part and a secure part.

If the standard mode disappears it will diminish the users control over his computer and will make it harder to develop shareware and similar programs since they must be implemented towards TC.

Making the entire OS secure will be a daunting task but if completed it will put a stop for today's viruses and other malicious code.

6.1.2 TC with DRM

By using DRM on top of TC it can be made secure and resistant of tampering. DRM that will affect the home user is mainly the DRM associated with digital media.

Costs

A home user will not have any direct cost from DRM. DRM clients will be integrated into windows media player and similar programs.

Dependencies

The users must have an application with DRM client to be able to use a DRM file. There is also a risk that the user must have a specific application, decided by the owner of the DRM file, to open the file.

Consequences pros

- Larger range of digital media available
- Cheaper media

With DRM protected media possibly many more companies will start to sell digital media over the net, hence the range of media available for the home user will increase.

Buying music and video over the net tends to be cheaper than buying it in the stores. The user can also begin to rent software and other products over the net.

Consequences cons

- Less user control over media/files
- Seller decides supported player/players and devices

When digital media is distributed and sold over the Internet the seller specifies the rights for the media files. This allows the seller to restrict on which computer and on which software the file can be used. This really diminishes the users control over his bought media. If a person for example buys a CD today with music, he or she can copy the music using it on an mp3 player, tape player or other owned devices. The ability to do this with DRM protected media can easily be prevented by the seller. That the seller can decide with which player and on what platform the media can be played will limit the users freedom.

There is a risk that music and video suppliers decides that their music only should be played with their media player. This forces the users into using a media player he or she may not like. The required player can very well be free to download and use but it still limits the users option to select what player to use.

The future

As TC and DRM become standard we think that more and more digital media will be distributed and sold over the net. Bookstore, music and video stores where users can download a movie that work for one day and cheaper books under a limited period or more expensive time unlimited books is possible scenarios. It is likely that the prices for media can

decrease when more users start paying for it. Paying nothing is of course the “best” for the user but we do not think that is possible in the long run.

6.2 Companies

This section will consider users of the TC platform at companies and companies as a whole.

6.2.1 TC

In this section we consider the effects of TC.

Costs

If a company decides to switch to TC in one step it will become very costly. All the computers and OS must be replaced and new software must be bought. This is probably not the most likely scenario. An alternative and more likely scenario is that a company buys TC computers when new computers are needed and TC OS and applications are purchased when need arises. By doing it like this, the company will not be standing with lots of still working non-TC computers and the cost of the switching will be distributed over a longer period of time.

Dependencies

A risk with TC is that there will be few OS supporting the TC platform in the beginning hence the number of choices will be limited. Companies in need of TC would have to choose the existing TC-OS and to switch to another TC-OS later could prove difficult and cumbersome since new applications must be bought and data must be migrated.

Consequences pros

- Control over users with attestation
- File and key protection
- Company can control users using their servers
- Secure application execution
- Better protection for laptops
- API for TPM available for anyone

Attestation is generally most profitable for the one who requests it. With the help of attestation (see section 3.6, Remote attestation) it is relatively easy for a company to force all employees and other users to use the programs decided by the company when connecting to a company server.

The TC platform will make it safer for each user within the company to protect their files, documents and other secrets like keys etc (see section 3.7). Files and documents can be encrypted with keys that in turn will be encrypted and protected by hardware (TPM). This will make it almost impossible for unauthorized staff to get access to them.

Companies can use the TC platform on servers and can demand that users who want to interact with the server have to prove for the server by remote attestation that the user are running a TC platform and running an application that is authorized to interact with the server. This will give the companies control in what applications and platforms may interact with their servers.

When a TC user wants to transfer a secret file to another TC user, the sender can be confident in that the receiver runs a trusted program. The trusted program that is not subject to attack from other software (because they are isolated), can know the private keys used in the process. This gives strong confidence when secret files are transferred between users and

companies. It is also possible to extend this, for instance, a group of trusted programs on a group of machines can share the private key used to decrypt the file so that all of these programs can open the file. Also, once this group of programs trusts the public key, one of them could digitally sign a statement with the public key in it, then all that is needed is for the receiver to hand this certificate to the senders trusted program.

By executing applications in the secure mode the company users can have confidence in that no malicious code has been hidden in the applications because of the measurements done on the applications before executing. They can be sure of that their information will not be compromised by any spyware because of the secure I/O devices and channels. They can also have confidence in that different applications running in the secure mode at the same time do not interfere with each other or compromise each others data due to the isolated execution (see section 4.2.4, Process isolation).

The secure protection of files is especially important when secret files are stored on laptops because laptops are easier and more common to get stolen. The protection of files by the TC platform will be more efficient and secure than today's file protection systems, like hard drive crypto which by the way will take ages to decrypt for the user every time the laptop is booted. In TC, just the secret files are encrypted. If the computer/decryption key is protected by a pass phrase then the key are stored on the hard drive. With TC, the protection of this key will be better because there is hardware that protects it, without TC this key could be compromised. The security above also depends on that the authentication is reliable since if a thief can get into the computer he can get access to the secret files. TC recommends that authentication should be done by some biometric test (e.g. fingerprint), or with smart card that in turn are protected with a secret, or a combination of these.

The API to the TPM will be available to everyone; this means that who ever wants to develop a program that will run in the secure mode and make use of TC functions are free to do so. No authorization is needed for developing TC applications.

Consequences cons

- Company can be controlled when using server beyond their control
- Company locked to a supplier by files tied to software
- Insecure key creation process on manufacturing the TPMs

Attestation can be an advantage for a company when they set up a server and wants to have control over it's' users. On the other hand, if the company uses a server beyond their control they are themselves a target for controlling and possible requirement of special software.

As stated in the user section (see section 6.1.1, Consequences cons), it is possible to tie files created by a program to the creating program. We will continue on the Microsoft Word example here. Companies will be even more locked by the used program than users are, since companies can have vast amount of documentation that they simply cannot afford to loose. It is also likely that the company works with several other companies who may use the same word processor product. If the company should decide to change to another word processing program it will affect the company since much internal and external information would be unreadable. Problems like these would probably make many companies chose not to change the software unless they have a good reason. An alternative would be to keep some licences of the old word processing program for backward compatibility and in the long run complete the change.

We have a question mark for how secure the process is where the manufacturer is responsible for giving each TPM its unique key pair (EK see section 3.8.4). Who controls this and what prevents the manufacturer to make several chips with the same key pair? Can

already used key pairs possibly get into wrong hands? This seams like a weak link in the chain and should be reconsidered thus buyers of TC platforms can put more trust in their platforms from the beginning.

The future

Companies buy computers relatively often and most companies will probably have many TC computers in a few years time. Even if companies will get the TC hardware relatively fast it may take time to switch to an OS supporting TC and new applications taking advantage of it.

6.2.2 TC with DRM

By using DRM on top of TC it can be made secure and resistant of tampering. DRM that will affect the companies are both DRM associated with digital media and DRM for document handling.

Costs

Buying a software solution for document rights handling is not cheap if you look at today's only option of if, Microsoft office 2003 [34]. Microsoft even charges extra for the professional version that is the only version that can create DRM documents.

Cost for setting up a media distribution system supporting DRM will be more expensive than a system not supporting DRM but we think the difference will not be substantial since it does not include many new components. Companies wanting to sell media can use an already working DRM media system such as windows media player and following their syntax in specifying the rights.

Dependencies

There are few DRM solutions for document handling and the interoperability is poor hence companies can have problems when wanting to change from one product to another. This could be especially hard if many documents exist since they could be lost in a product switch. The poor interoperability is could potentially make a company dependent on an application.

Consequences pros

- Controlled media distribution over Internet
- Rights management for documents

DRM for e-commerce companies makes it possible for them to control the usage of the content they sell by specifying rights that are connected to the content. For instance, rights can be that a purchaser does not have the rights to copy, print, forward etc the content or there are a time limit on the content and who will be able to use it etc. This gives rise to a whole new e-market; software companies could for example let software to other companies for a fee, rights could be tagged so a specific fee was withdrawn from the user for every week, month etc that the content was used. A larger fee may be taken for full functionality and smaller fee for restricted functionality. It will be easier to rent or buy audio and video files in a secure way as well as e-books or PC games etc. A fee could be taken every time you watched a movie or read an e-book or to let you use the content for a certain amount of time. When the time limit is due, rights could be tagged to make the file unusable. For instance, if the digital content industry (e.g. music-, movie-, e-book industry etc) would use DRM on its digital content and thereby stop the copying and illegal distribution of their content; it would be an enormous change in their income. The music industry regarded that the loss on music piracy was around $4,5 billion in 2002 [39] and the global movie industry looses around $3 billion

on piracy each year [40].

By using a DRM system document handling system such as Microsoft RMS, a company can gain much control over different documents. By being able to specify rights, groups and users connected to the rights a company can control which employee who has access to different information. By having different levels of secrecy predefined allows the creator to quickly give a document or email the correct rights. This is a function that many companies today have uses for. By letting the creator of a document specify in detail how a document can be used is a valuable function. For instance specifying that an email cannot be forwarded, copied or printed pretty much ensures the email does not leave the recipients computer. DRM for document handling internally in a company shows much promise and will probably be used by many companies in the near future.

Consequences cons

- No standard, poor interoperability
- Small range of DRM products for document handling

Today there are many different implementations and languages to specify rights, but there is no standard hence interoperability between these different approaches is a problem. With limited interoperability companies choosing to use a product will have a hard time switching to another without starting from scratch. Today it is mainly Microsoft who provides DRM for document handling within a company and that does not give companies wanting to use this much choice. Adobe has plug-ins for its product (e.g. Adobe reader) to allow DRM to be utilized to their products. Hopefully more alternative will arrive soon that can interoperate with each other.

The future

We think that all media distributed over Internet in the future will be protected by some kind of DRM. With DRM companies owning media can make sure that people using the media pay them, hence it is very unlikely the companies will sell media without DRM rights.

We think that companies will start using DRM for specifying security levels on documents. Microsoft's RMS solution can be used for this and we think more options for this will follow.

Some standard or interface between different approaches of DRM will hopefully be specified in the future to improve the interoperability.

6.3 Suppliers

Suppliers in this aspect are those who supply hardware and/or software that support trusted computing and/or DRM.

6.3.1 TC

In this section we refer to TCG, NGSCB and LaGrande as TC. All these three technologies have the same fundamentals.

Costs

There will be an initial cost for suppliers wanting to develop software that supports TC. Employees must be trained and new development software is needed.

Dependencies

The suppliers are dependent on that the users and companies accept TC and starts using it.

Consequences pros

- Increasing demand for new hardware
- Less or no piracy
- More control over software products
- Lock in

In the short run the following could be a positive thing for suppliers. Today's I/O devices and other hardware do not support the TC implementation specified by NGSCB or LaGrande. This means that the suppliers have to develop new hardware like a new CPU, graphic cards, a new mouse and keyboard, display devices and "connected" chips for these. In the beginning, these developers will get higher revenue because every one who wants to use the TC platform has to exchange their old hardware with new TC compatible ones. In long run the old not TC compatible hardware would be faced out and the market would go back to normal.

The same aspect applies for software companies who develop TC supporting software. In the short run, the sales curve would rise because everybody has to buy new TC software. But in the long run companies have already exchanged their software to TC compatible ones and the market would go back to normal.

If the problem with software piracy can be "solved" for those programs running in trusted mode (see section 6.1.1, Consequences cons), software suppliers would gain a lot. If everyone has to buy the software instead of just download a cracked version from the net, it will raise the income for most software companies, we are talking enormous amounts (Worldwide software industry estimated a loss of 13 billion American dollar annually due to piracy, [41]) for some companies and less for others. Maybe users will switch to cheaper programs when they cannot run cracked versions of good expensive ones anymore. This can be good for the supplier of smaller and cheaper software since they can get a bigger market share. It may be easier for some software companies to establish themselves on the market when they know that their software products will not be a subject to piracy where a loss of income otherwise could occur.

Software developers will have much more control over their software since they can specify in the manifest e.g. which plugins that are allowed (see section 4.2.2, NCA).

As described in before (see section 6.1.1, Consequences cons) a mutual key can be used to only allow the desired application to open a file. This gives the lock in ability for the software developers. The very big suppliers can decide that their product only shall work with other products from them; this gives the big suppliers more power.

Consequences cons

- Lock in is negative for small suppliers
- More complex applications

Hardware/software suppliers have to develop TC compatible hardware/software so they do not loose any market shares. For a customer that wants to buy an application there may be several to choose from and the choice may rely on the support for the TC platform.

The small suppliers will not gain anything from lock in, on the contrary since big suppliers using lock in will have a firm grip on its customers, it will be harder for small suppliers to gain customers.

TC will add more complexity in the already complex process of software development; hence there is a risk that more faults can occur in applications. The added complexity could also lead to higher need for user support from the suppliers.

The future

Software suppliers may in the future develop their software so they have to run in secure mode. This would make the application secure from piracy. The TC platform could check that the application has valid serial number and that the application has not been tampered with.

6.3.2 TC with DRM

By using DRM on top of TC it can be made secure and resistant of tampering.

Costs

We cannot see any special cost for developing DRM software except for that programmers need to learn the language used to specifying rights.

Dependencies

DRM applications are dependent of TC to really be considered secure. Other than that we cannot se any dependencies for suppliers.

Consequences pros

With a new technology the suppliers can make much money from selling new licences. The first suppliers delivering a good DRM solution can tie lots of customers to them since the interoperability is bad.

DRM opens big possibilities in digital content distribution market and there will be a high demand on DRM solutions.

Consequences cons

We cannot se any major drawbacks here, but one thing that may be difficult for the suppliers is to introduce a product with DRM since there has been much bad publicity around DRM.

The future

We think that DRM solutions for different document handling will be common in the future.

7 Discussion

In this section the effects and problem formulation questions from section 1.2 will be discussed.

7.1 Problem formulation

What new costs, dependencies and consequences will TC and DRM give for its different users? TC and DRM will have many different effects on its users (effects considered in section 6). As a general abstraction it can be stated that the higher up in the chain and the bigger party the more advantages TC and DRM will give. Suppliers gain more control over their customers, companies get controlled by suppliers but can gain more control over its employees/users and home gain little control but can be controlled by companies and suppliers. This is from a control aspect and there are still functions in TC and DRM that is good for home users. As we can see it TC and DRM are developed to give suppliers more control over their software and digital media, and at the same time add needed security enhancements for users and companies.

A very smart move by the suppliers is that they allow the TC hardware to be turned on and off. This is very reassuring for users and companies but there is a catch. Suppliers can make programs that do not work without TC, and servers can be configured to only serve TC platforms. So even if turning the TC hardware on and off is allowed, users probably will not be able to run all new software without TC activated and there will be an indirect demand on users and companies. Another smart move by the suppliers is separating the OS into a standard mode and secure mode. This allows the users to run both TC application and standard application in the same OS. Without this all old software would stop working when switching to a TC OS (unless multiple OSs is used). This would make the switching a very big step and we doubt that many home users and companies would do it due to huge costs and possible loss of critical existing functionality. Now with the divided OS companies and users can make an easier and less dramatic change to TC.

In the sections describing TC and DRM we mentioned but did not discuss that the performance will be affected by this new security. With all keys and encryption in different processes the performance will be affected. How much can only be speculated in but with computer power continuously improving the load for the processor probably will not be a problem. Hard drive speed on the other hand has not improved significantly over the last decade and is already becoming a bottleneck in today's computers. But we do not think it will be a big problem either since TC will not put much more strain on disks. What we think is going to be affected most is the need for a fast Internet connection. With all overhead being sent for attestation, checking licences etc the strain on the slow Internet connection will be hard. It is hard to speculate on what connection speed that should be required and the suppliers has not gotten far enough in development to specify requirements yet, but we think that using a modem on a TC/DRM platform will not be feasible.

There are ways to hide the new delays for the user. For instance, when a user surf the Internet, the computer can get AIK key in pre-emptive purpose and when the user gets to a page requesting attestation an AIK is ready for usage.

All handling and needed verification of keys with TC and DRM will increase the need for TTP and CAs. Both attestation and the increased need for personal certificates in DRM are causes for this.

There will be costs for users and companies to buy new hardware and new software supporting the new technology. The process of buying new hardware should be done over some time to avoid having lots of working non-TC platforms standing unused and to spread

the cost over a larger period of time. The added complexity of TC and DRM will add costs for the suppliers in both the development process and the support phase.

Will TC stop software piracy, viruses and spam? TC will have no direct affect on spam, piracy and will not stop viruses but TC can improve the techniques trying to prevent them.

Spam is a big problem for companies and users and there are different solutions to this problem even without TC. Examples of these are forcing the computer to perform a meaningless calculation before sending an email or requiring all email to have a sender identity. The idea of forcing the computer to do a calculation before sending an email is a good idea hence it raises the costs for spammers since they must buy much more computers to send as many email as they do today. If the cost increases the amount of spam would decrease since the whole idea with spam is that it cost almost nothing to send an email. The calculation would not affect "normal" users since it does not matter for them if sending an email takes a fraction of a second or 5 seconds. TC could be used to make sure that the calculation actually is done and sending proof of this in the email or use attestation to prove an approved email client/server is used. We think that this would be an excellent solution to the spam problem that bothers the majority of computer users. There is a drawback with this method and that is that the calculation will take longer time on an old computer than a new one, but basing the calculation on the processor speed could solve this. For this method to work it requires that all mail servers are TC platforms and only accepts mail from other TC servers. If spammers would start a non-TC mail server to spread spam, these mails could be discarded or marked as "possible spam" and allow the user to make the selection.

TC cannot stop viruses, but TC allows anti-virus programs to run in secure mode and thereby be protected from viruses trying to disturb the programs efforts. TC also protects applications running in secure mode from viruses. If we look at Microsoft and Intel's way to implement the TC operating system with a standard part and a secure part then as we have mentioned before, the standard part will work as operating system works today, hence malicious code could still run in that part and make your life miserable. This may not be so dangerous if you use your secret data in the secure part where it is protected.

TC does not directly prevent software privacy either but programs using TC can use its functionality to prevent piracy. An example of this is described in section 6.1.1.3.

What effect will it have on companies specializing on security? We start to look on anti-virus programs, Microsoft and Intel's way to implement the TC operating system with a standard part and a secure part will still need protection from different viruses. Therefore TC will not have any direct negative effects on companies selling anti-virus programs.

If we take the smart cards vendors, their role in whole this will not be faced out. There will still be a need for smart cards to authenticate users at login time, to store different login keys etc. Microsoft states that NGSCB will even increase the demand for smart cards because it will offer stronger protections and security, this technology will increase the benefit of the smart cards. Smart cards and TC are complementary because they are looking on different authentication problems. TC authenticates the machine and/or the software stack while smart cards are used to authenticate users.

Firewalls are protection mechanism commonly used by both home users and companies to protect their computers and networks from intrusion. As we see it, the effect TC could have on this is that an intrusion would not do as much damage as is could without TC if the secret data is protected. Hence TC will not affect companies specializing in software and hardware firewalls. Firewalls can be seen as a complement to TC instead of a competitor.

As TC is intended to function there will be a big demand for TTPs that can be used in the process of AIK certificate creation. To emphasize how many certificates that would be needed we will make a simple example. If there were one million TC users in Sweden using their computer at home and at work, and every person needs two new AIK certificates each day.

That would end up in fourteen millions certificates per week in Sweden alone. DAA (Direct Anonymously Attestation) could decrease this and companies could have their own TTP instances but we think the need for TTPs will increase dramatically with TC. Companies specializing in TTP structures and similar would certainly get an economic boost out of this.

Will companies and users be "forced" into using TC and DRM? There is an obvious risk for this, if suppliers begin to develop software that only works on TC platform. It is hard to predict which action software developers will take regarding TC support and TC requirements for software. If a software developer only supports TC platform there is a risk that customers will turn to other software developers instead, but a company with a large market share may be able to make demands like that. Suppliers can decide that their software only works on TC platforms and justify it by TC protects their software from viruses and piracy.

Another part of TC that has the potential of forcing companies and users into start using TC is attestation. If many servers on the Internet start requesting attestation, non-TC platform will not be served at all by these servers since they cannot attest them selves. It is the person or organisation owning the server who decides if attestation should be required or not and most parties has little to gain from requesting attestation. For example should not most web server care if you use a TC platform or not. We do not think that the most server owners will start requesting attestation but since the companies behind TC are very big and has big influence it could be enough if they all starts doing it.

There is also a risk that if many companies start using DRM for specifying rights for documents other companies interacting with those companies can feel compelled to start using it too to be able to interact in a smooth way. Microsoft states that documents created with their RMS system can be opened in Internet explorer if a special plug-in is used, options like this may exist for other software developers as well. Options like this plug-in will at least give companies and users without DRM document system the chance to read the documents and therefore decrease the need for buying new DRM software. The plug-ins still needs the underlying PKI structure with certificates etc.

How can TC and DRM violate the integrity and privacy of its users? With TC the use of AIK will give the user good privacy and it is only the TTP used in the AIK creation that sees the users public endorsement key and if this party really is trusted this should not be a problem. To further improve the privacy in the process DAA can be used that allows attestation without the involvement of a TTP.

One aspect with DRM content distribution is if media files are bought by a user then the vendor will probably keep a list of bought licenses over users. How will lists like this be handled? They could be used to start sending directed commercials to customers against their will and could also be sold to other companies interested of users purchase habits. A list like this could violate the user's integrity and privacy. On the other hand, a record over purchased content and licenses should exist so lost content (in the case of a hard drive crash etc) can be downloaded again and used without having to pay another fee. For this reason we think that it is relevant to use records over users and their purchases, but only for this reason and only if the user does not agree on something else.

How safe is TC and DRM? This is hard to speculate on since TC is a huge initiative and there is no implementation ready for testing. There is one procedure that we do question the security of and that is the creation of unique security chip identity key pairs (EK) by the manufacturer of the chip. Much in TC depends on the uniqueness and secrecy of this key pair and giving the responsibility for handling this to the manufacturer does not seem like a good idea. We think that there should be some kind of CA structure for creating end ensuring uniqueness and secrecy of the key pairs.

Much of the security in TC is based on hardware components, hence putting TC security on a higher level compared with software based security. Hardware components are much

harder to ”hack” and often require the hacker to be physically present. We think that TC (TPM, secure graphic/ I/O, protected memory and protected storage) will push the security on computers to a new level, but we do not think that it will be flawless and careful testing and maintenance should be performed to fix any flaws and security holes. TC gives the user a computer with much integrated security; hence the users/companies do not need add that many separate security components (e.g. Pointsec, PGP, etc) to get a secure platform/system.

DRM cannot be considered completely secure without TC to support it. Without TC, DRM clients execute in the shared memory space and do not have any tamper protection. Therefore it would be relatively easy to modify or monitor a DRM client and thereby get hold of keys and unprotected media. Even if it is possible to hack DRM without TC, DRM will raise the security level and will be used despite of its flaws. DRM can be considered moderately secure without TC, but with TC backing it up it will take the security to a new level. The DRM clients will be tamper-protected, run in curtained memory and have the I/O channels and graphic channels protected, hence stealing keys and media from the client will be very hard.

In appendix 1, some comparisons between current solutions and TC solutions are made.

What kind of operating system support TC? In the present there is only one implementation underway and that is Microsoft NGSCB. This implementation has its own specifications and ideas with some parts of it protected by patent. This implementation makes use of TCG’s TPM v1.2, which is free for anyone to use in the implementations of TC. We can expect to see NGSCB integrated in an operating system no sooner than in Microsoft Longhorn version 2 which probably will have its first release earliest during 2006.

TCG’s TC specifications are free to use for every operating system developer. But there are no implementations that make use of it yet, as far as we know.

LaGrande is a template for how an operating system developer should implement a TC operating system based on Intel’s hardware. This template is not tailored for a specific developer and can hence be used by anyone. That is a bit weird because it seems like NGSCB and LaGrande pretty much goes hand in hand but Microsoft has protected some parts by patent but LaGrande is free to use for any developer. Microsoft and Intel have also made agreements on new hardware for the TC platform. Microsoft has also stated that the NGSCB hardware will have an open specification, and that Linux could be written for it, but so far we have not seen any non-discriminatory licensing of the patents made. So, we think that Microsoft easily could prevent Linux from using the hardware features. Is this legal? If Linux will implement support for TC with similar features as Microsoft then there will be a big dispute. We think that, if they really want to, they could come around Microsoft’s patents and develop Linux with TC functionality. We think that we can expect Linux with TC on the market not long after Microsoft’s first release depending on the markets demand.

We think that many OS developers at least have an idea about an implementation of an OS with TC support. If this technology hit the market then they cannot afford to not be a part of it. TC has been an issue for many years, so everybody knows that it is important and that we have to do something about it.

Which software and hardware will be “trusted”, how and by whom will it be approved? There will not be any agencies that have to approve any applications or approve any software companies to let them use the TC API. This is not necessary anyway because of the strong process isolation offered by TC. Due to the process isolation developers are not able to develop applications running in the secure mode that can compromise other applications data or data stored by the secure part. The application have to be approved by the user to run in the secure mode and the secure mode have to approve the application as well based on the comparison of the integrity metrics made by the platform and the metrics in the manifest. Software vendors, not necessarily for Microsoft’s TC platform may have to upgrade their SDKs (Software Developing Kits) to be able to develop a “trusted” application and they have

to make use of the API for that particular TC platform to be able to run in the secure mode. The program also needs to be connected to a manifest that the secure part needs for measurement of that program.

For the TC hardware to be trusted, a chain of credentials created by the manufacturer will follow with TC hardware. These credentials will tell who made the TPM and is signed by the manufacturer. The credentials are evaluated before an AIK is signed by a TTP.

Will governments get special access to criminal's data secured by a trusted platform and/or DRM? Criminals as everybody else can take advantage of TC and DRM, hence protect secret data in away that makes it hard to compromise. This data could be really valuable to governments that solve crimes. The question then arise, if there will be a backdoor in TC that allows governments to access TC protected data. To this Microsoft replied with the following statement: "Microsoft will never voluntarily place a backdoor in any of its products and would fiercely resist any attempt to require backdoors in products". Note the word "voluntarily"[42], there is a risk that government agencies demand a backdoor for example to "fight terrorism" and as the situation is in the world today, this is a possibility that should be taken into account. If a backdoor were to be introduced it would be possible for government agencies to access secret information. This is one disturbing scenario and will hopefully not occur.

We have to consider the security aspect here as well, if there is going to be a backdoor, this backdoor will definitely be a desirable subject for hacking. If a hacker succeeds to find this backdoor and hack it, the hacker would have access to lots of secret information. We think that there should definitely not be a backdoor, due to security and privacy.

What happens with open source programs? As long as there is a standard mode in the TC OS, programmers can continue writing free programs as they do today. If we look at the possible scenario ten to fifteen years from now where the OS only supports a secure mode powered by TC the situation would be different. In this scenario the programmers writing free software will have to learn the API to program against the TC-OS and development tool would probably cost more which could decrease the amount of free software. We think it is possible that the standard mode will disappear in the future since this would give the suppliers even more control.

How will TC and DRM affect companies, users and suppliers? By answering the sub queries above and reviewing the effects in section 6, we feel that we have answered this question to a fairly high extent. There are still some unclear aspects but we think that these will become apparent when the TC initiative draws closer to completion.

8 Conclusion

The TC initiative is still in the development stage and there are currently many aspects left unresolved. The companies behind TC OSs (mainly Microsoft and Intel) are reluctant to reveal detailed information on their TC initiatives and there is much secrecy involved. We hope that more players (e.g. Linux, MacOS) come into the TS-OS market to give users more options.

A TC platform can be seen as an ordinary computer with a security package included in the hardware, hence the user get a good base of security that can be enhanced with further components. TC can in most cases be seen as a complement to other security solutions.

There are many different undergoing implementations of DRM and most of them have similar structures. DRM allows owners and creators of digital content to specify how it should be used and at the same time makes sure that specified rights are followed. There has not been any working way of doing this before and DRM shows much promise on enforcing rights for digital content.

We think that DRM needs an establishment period to get higher interoperability and to standardize the rights specification vocabulary in some manner. We have no doubts that DRM will be a success with or without TC to back it up. There are already working DRM solutions on the market today and more will come.

TC and DRM will bring the security of PC to a higher level and brings many benefits and some drawbacks to its users.

The main positive aspects for the home users are the secure storage of data, the secure execution environment for applications and a wider range of cheaper digital media. The main drawbacks for the home users are that they can be controlled with attestation, that applications can prevent other applications to open their files and that users get less control over their digital media.

The main positive aspects for companies are the right handling of digital content. As well as the secure storage of data, secure execution environment and the use of attestation to control the users of their servers. Attestation is also a drawback for companies since servers beyond their control can in turn control the company. In the same way as for users, it is a problem for companies that suppliers can prevent other applications from using their files.

The main positive aspects for suppliers are that they gain more control over their software products and they can decrease or stop piracy on their products. The main drawback for the suppliers is the cost for developing TC/DRM and the added complexity to applications using TC/DRM.

Since TC is a very big initiative, we had to make some abstractions and did not have time to go into detail of every aspect; hence there may be flaws in TC that we did not discover. We did discover one flaw that we think is significant for the trust in TC. This flaw is in the creation and handling of unique key pairs for the security chip. As this process is currently planed, the manufacturer of the chip is responsible for creating the unique key pair for the chips. We do not think this is good enough since much security is based on this key pair and therefore should be created and handled by a trusted organisation or agency.

As we see it, the benefits with TC and DRM outweigh the drawbacks and we therefore think that TC and DRM should be fulfilled and taken into production.

This standpoint is based on the information available today and there may be new or changed information as TC draws closer to release, which may give rise to the need for a reconsideration of this standpoint.

TC and DRM provide a good base of security and it is up to the developers to use this functionality in a good and responsible way.

References

[1] How trustworthy is trusted computing?, Computer Volume: 36 Issue: 3, Vaughan-Nichols, S.J, March 2003

[2] http://www.trustedcomputing.org/home, TCPA, April 2004

[3] https://www.trustedcomputinggroup.org/about/, TCG, April 2004

[4] Understanding trusted computing: will its benefits outweigh its drawbacks?, Security & Privacy Magazine, IEEE , Volume: 1 Issue: 3, Felten, E.W , May-June 2003

[5] Privacy and trusted computing, Database and Expert Systems Applications, Reid, J, Gonzalez Nieto, J.M. Dawson, E. Okamoto, E, Sept 2003

[6] https://www.trustedcomputinggroup.org/downloads/Main_TCG_Architecture_v1_1b.zip, TCG, May 2004

[7] http://www.webpk.de/02072003_bmwa_trustworthycomputing/download/1_TCG%20-%20Strategie%20und%20Konzept%20-%20Michael%20Waidner.pdf, Waidner Dr. Michael, IBM Research Division, May 2004

[8] http://www.giac.org/practical/GSEC/Chris_Hageman_GSEC.pdf, Hageman Chris, GIAC, May 2004

[9] ftp://download.intel.com/technology/security/downloads/scms19_direct_proof.pdf, Brickell Ernie, Intel Corporation, May 2004

[10] https://www.secure.trusted-site.de/certuvit/pdf/9205BE.pdf, Bödiker Dr. Patrick, Infineon Technologies, May 2004

[11] http://www.intel.com/design/mobile/platform/downloads/Trusted_Platform_Module_ White_Paper.pdf, Bajikar Sundeep, Intel Corporation, May 2004

[12] http://www.intel.com/idf/us/fall2003/presentations/F03USSCMS25_OS.pdf, Manolopoulos Zorba, Intel Corporation, May 2004

[13] https://www.trustedcomputinggroup.org/downloads/TCG_Backgrounder.pdf, TCG, May 2004

[14] https://www.trustedcomputinggroup.org/downloads/tpmwg-mainrev62_Part1_Design_Principles.pdf, TCG, May 2004

[15] http://www.netproject.com/presentations/TCPA/charles_powell.pdf, Powell Charles Scott, Infineon Technologies, May 2004

[16] http://www.cse.msu.edu/~cse891/Sect002/lectures/tcpa.pdf, MSU, May 2004

[17] https://www.trustedcomputinggroup.org/downloads/TCG_1_0_Architecture _Overview.pdf, TCG, May 2004

[18] https://www.trustedcomputinggroup.org/downloads/TSS_Version__1.1.pdf, TCG, April 2004

[19] http://www.infoworld.com/article/03/01/29/hntcpa_1.html, Krill Paul, InfoWorld, May 2004

[20]http://www.hyperorg.com/blogger/mtarchive/001162.html, Weinberger D, Hyperorg, May 2004

[21] ftp://download.intel.com/technology/security/downloads/scms18-lt_arch.pdf, Intel Corporation, May 2004

[22] http://www.microsoft.com/resources/ngscb/documents/ngscb_security_model.doc, Microsoft, May 2004

[23] http://www.rsaconference.com/rsa2003/europe/tracks/pdfs/appliedsecurity_ w15_lamacchia.pdf, RSA, April 2004

[24] http://www.microsoft.com/resources/ngscb/four_features.mspx, Microsoft, May 2004

[25] http://www.ece.cmu.edu/~rfrenz/papers/rfrenzNGSCB.pdf], Frenz Ryan W, May 2004

[26] http://sewpsc.sewp.nasa.gov/documents/SANS04_Night.pdf, SEWP, May 2004

[27] http://www.blackhat.com/presentations/win-usa-04/bh-win-04-blight/bh-win-04-blight.pdf, BlackHat, April 2004

[28] http://www.cs.cornell.edu/People/egs/syslunch-spring04/palladium.pdf, England Paul, May 2004

[29] http://www.intel.com/technology/security/index.htm, Intel Corporation, May 2004

[30] http://www.intel.com/technology/security/downloads/LT_Arch_Overview.pdf Intel Corporation, May 2004

[31] http://www.intel.com/technology/security/downloads/LT_policy_statement_0_8.pdf

[32] http://www.computerworld.com/governmenttopics/government/legalissues/story/0,1080182372,00.html?from=story_picks, Computer World, May 2004

[33] Microsoft tar kontrollen över dina document, Microdatorn 6, Per Lövgren, 2004

[34] http://www.microsoft.com/windowsserver2003/techinfo/overview/rmsoverview.mspx, Microsoft, April 2004

[35] http://www.eurescom.de/~pub/deliverables/documents/P1200-series/P1207/TI/P1207-TI.doc, Wegner Susan, Eurescom, May 2004

[36] http://xml.coverpages.org/AgnewREL200209.pdf, Cover Pages (OASIS), May 2004

[37] eXtensible rights Markup Language Specification 2.0 , ContentGuard, May 2004

[38] http://xml.coverpages.org/XrMLTechnicalOverview21-DRAFT.pdf, Cover Pages, June 2004

[39] Olofsson Kent, Pirat kopiera på jobbet lever en juridisk gråzon, Nätverk & Kommunikation, 2003 edition 12

[40] http://www.newyorkmetro.com/nymetro/movies/columns/hollywood/n_8677/, Thompson Anne, June 2004

[41] http://www.technewsworld.com/story/32982.html, TechNewsWorld, June 2004

[42] http://www.microsoft.com/technet/Security/news/ngscb.mspx, Microsoft, June 2004

[43] On research methods, Pertti Järvinen, 2001

Appendix 1 - Comparisons: TC versus current solutions

Here we make a few comparisons between TC and existing technologies to see similarities and differences.

TC protected storage versus traditional encryption (PGP)

When reading the thesis one could ask oneself the question; what is the difference between traditional encryption like PGP (Pretty Good Privacy) and TC protected storage. The two works pretty much in the same way with a few exceptions. With PGP the symmetric key is encrypted by the user's public key, hence only the user's private key can decrypt it. The private key can be stored in different ways where the easiest and most insecure is to place it on the hard drive protected by a password. Storing it on the hard drive only protected by a password is not very secure and it is better to keep the private key on a floppy disk or protect it with a smart card key. In TC the symmetric key is protected by a secret key protected inside the TPM, this simpler and gives high security. The other improvement of TC compared to PGP is the key generation that is hardware generated in TC and software based in PGP. This also improves the security since the keys is the foundation of the security.

If we just focus on the encryption service, TC is probably the better of the two but PGP with a smart card is also a good solution. For a user only after the encryption service, PGP with a smart card would be the best choice if we look at what the user get for the money, but TC includes much more than just encryption. Therefore it is a balance on how much security that is needed.

TC protected storage versus hard drive protection (Pointsec)

The most obvious difference hard drive protection such as Pointsec and TC's protected storage, is that Pointsec encrypts the entire hard drive were protected storage only encrypts the secret data on the computer. Having the entire hard drive encrypted has a negative performance impact and here TC is the faster of the two.

Pointsec keeps its keys on the hard drive but has a password protection preventing the computer from booting without correct password. The user selects this password and it is important to choose a good one to keep a high security. The TPM protect the key for the protected storage and this process can be considered secure.

A TC platform can authenticate a user with password too but it is recommended that a smart card and/or some biometric method be used.

Attestation versus thin clients

On first look on attestation from the company perspective it may look as the same thing as thin clients but that is not the case. A thin client only has the choice to run programs that are installed on the company server, no others. With attestation on the other hand the user can run any program and it only limits what program the employee can use when contacting a TC server.

Växjö
universitet

Matematiska och systemtekniska institutionen
SE-351 95 Växjö

tel 0470-70 80 00, fax 0470-840 04
www.msi.vxu.se