云计算入门

Introduction to Cloud Computing

GESC1001

Philippe Fournier-Viger

Professor

School of Humanities and Social Sciences

philfv8@yahoo.com

Fall 2020

Introduction

Last week:

- The importance of how data is stored in the cloud
- **Cloud storage** (how the data is stored in the cloud).

Today:

- Security in the cloud (云安全)
- The final exam

Week 8 Wednesday 21st October	(PDF / Powerpoint) **Second homework to submit on the 30th October before 23:59 PM at the e-mail address of the teaching assistant as a Word file or PDF file. The e-mail will be announced in the QQ group.** **Each student must do the homework by himself. This is not a team work.**
Week 9 Wednesday 28th October	**Cloud computing security** - chapter 9 (PDF / Powerpoint)
To be announced	**Final Exam (120 minutes)** **Time: _____ Room: _____** **Closed-book exam / Do not forget to bring your student ID.** The final exam will have about **10 questions**. Several questions will require to write a short answer in English (typically a few sentences). Some types of questions that I like to ask in final exams are: - What is the difference between X and Y ? - Explain what is "X". - Explain why X works in a certain way

CHAP 9. CLOUD SECURITY (云安全)

黑客

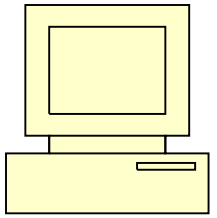
Introduction

- **Security** has **always** been a concern for computer systems.
- We may want to protect:
  - data (数据)
  - software (软件)

Introduction

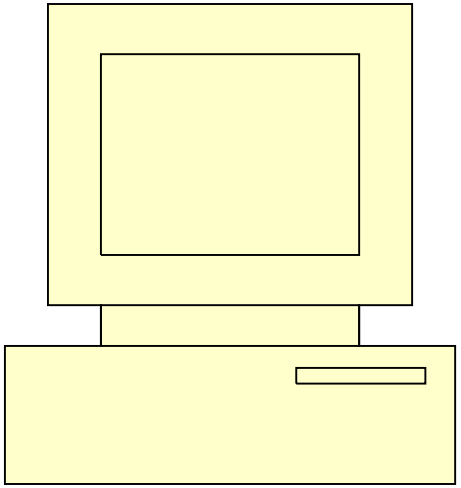
If someone has a **single computer**, s/he could protect the data and software on this computer:

- Using a password (密码) to access the computer.
- Using a password (密码) to open sensitive files (敏感文件) on the computer.
- Data encryption (数据加密)
- …



Introduction

In general, if someone has **physical access** **(物理访问)** to a computer, he can find a way to steal data from the computer.



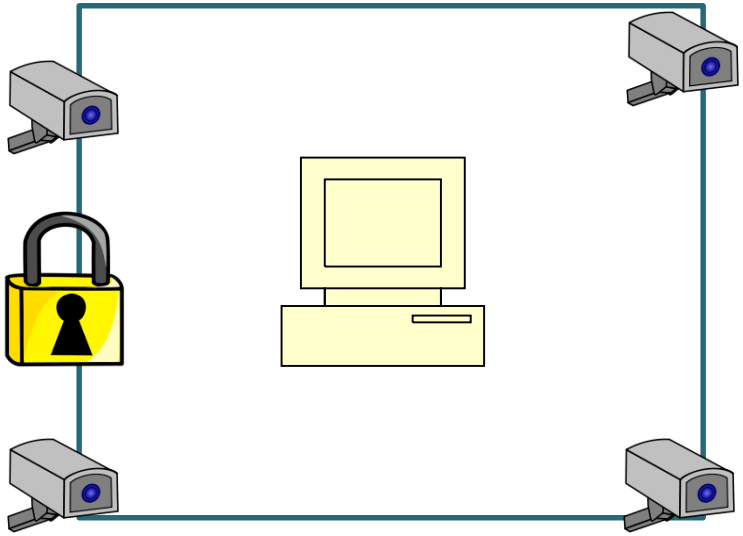
For example:

- Remove the hard drive (硬盘驱动器) from a computer and copy the data.
- Use a **keylogger (键盘记录)**:

Introduction

Thus, we can put the computer in a room with **restricted access** (限制访问) to ensure security.

Room with restricted access



A person must prove his identity to enter the room.

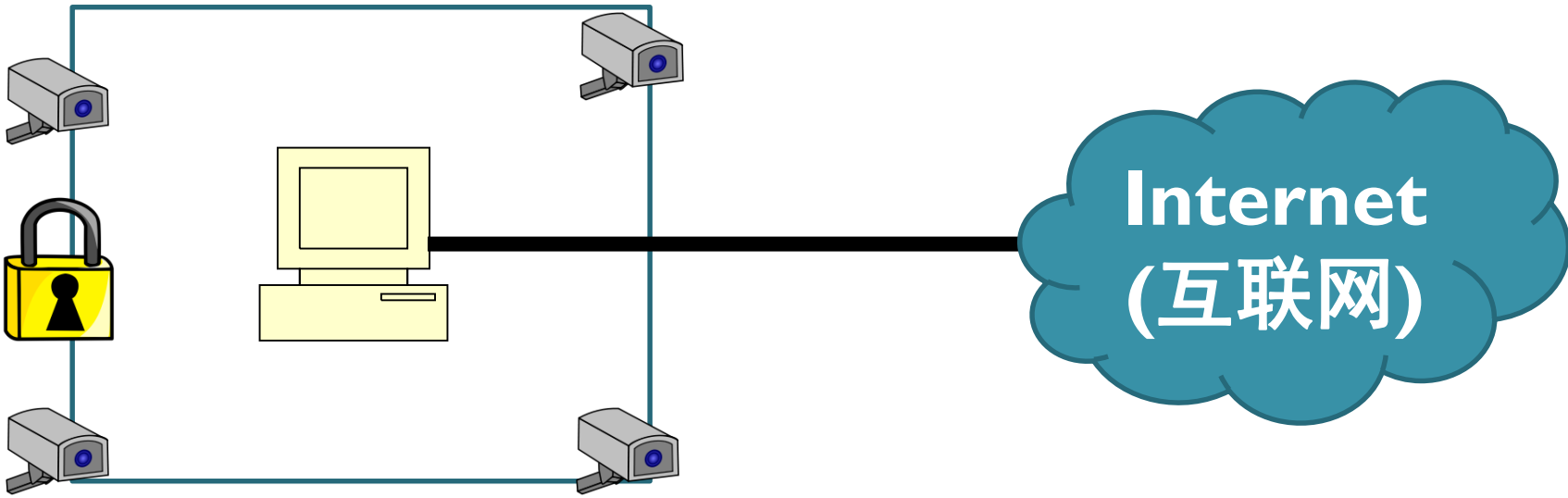
ID card (证件)

Fingerprints (指纹)

Introduction

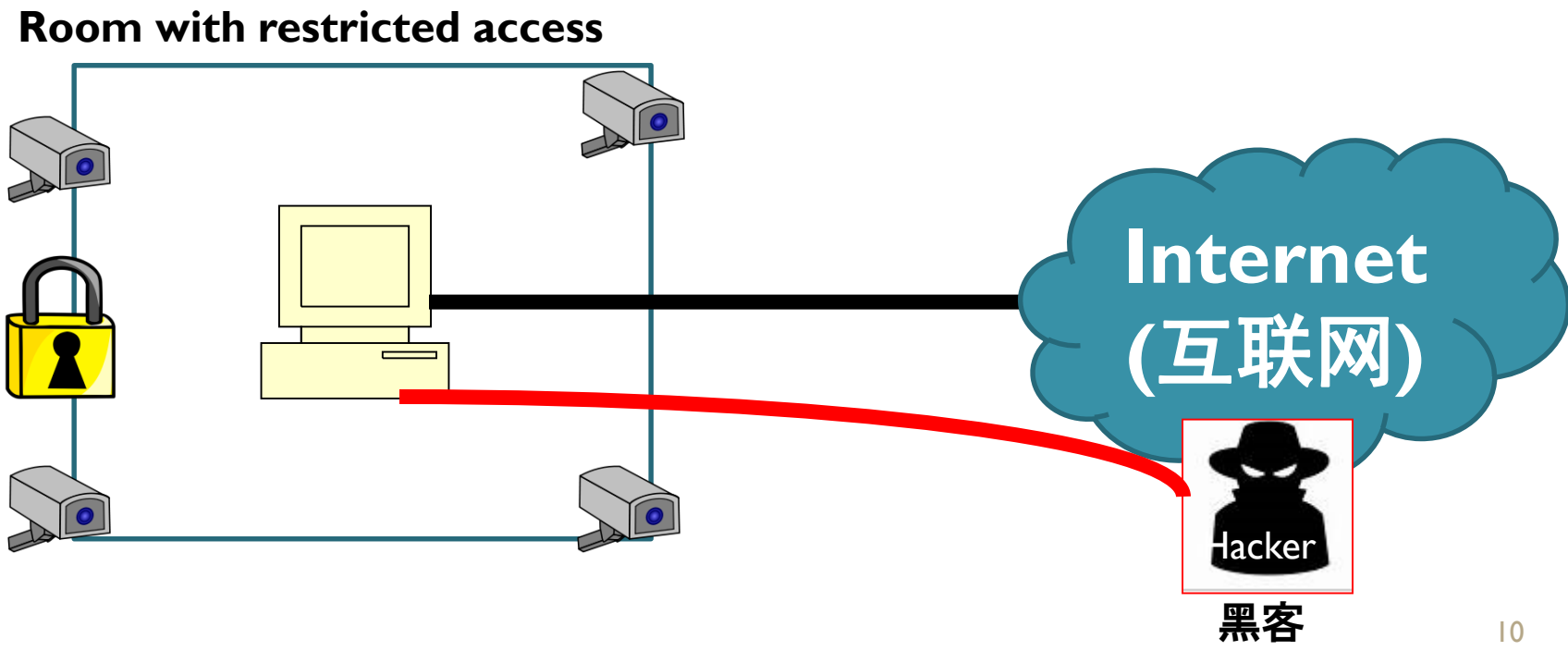
- Nowadays, most computers are connected to networks and the internet.
- Thus, ensuring security is more difficult.

Room with restricted access



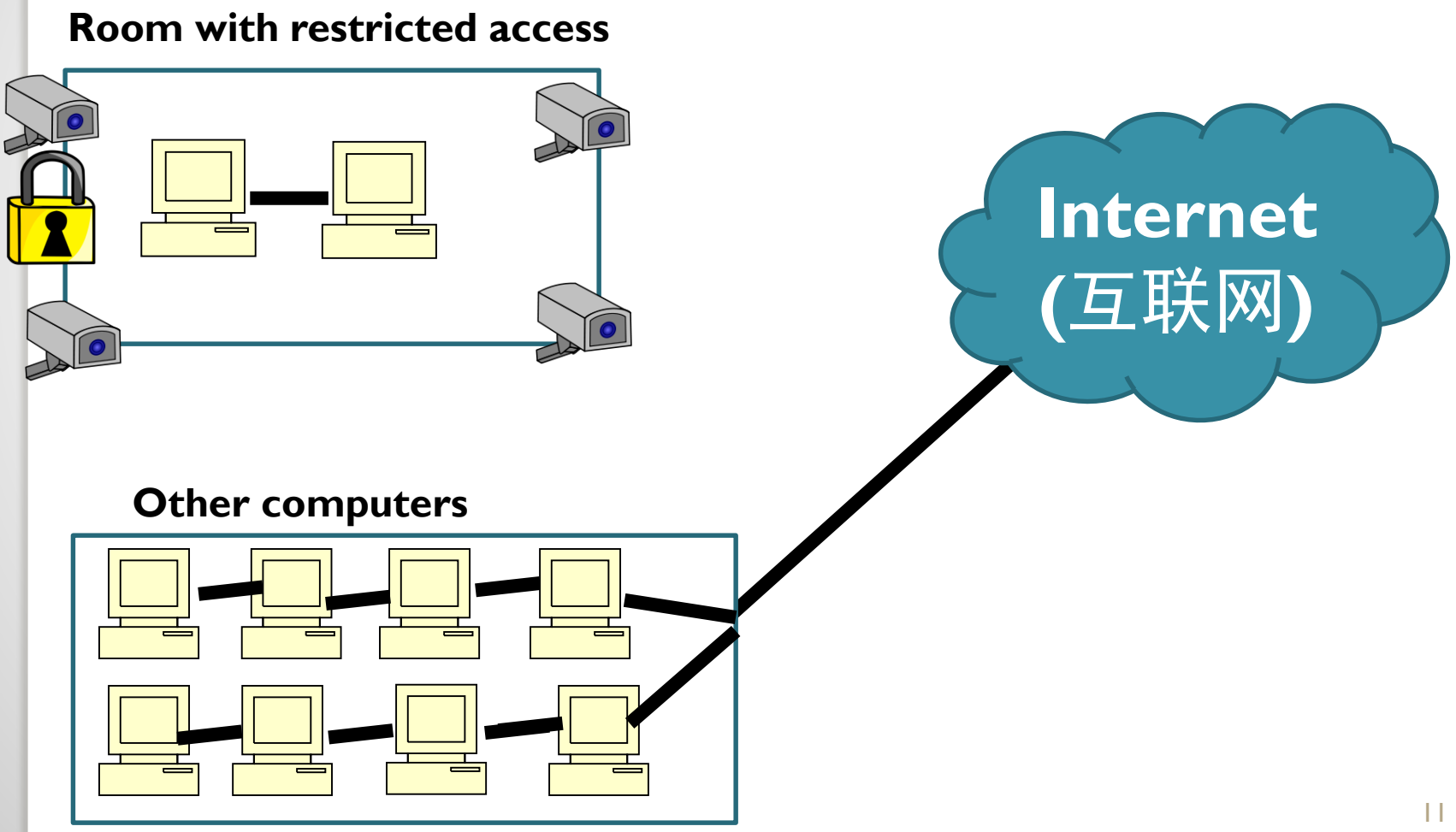
Introduction

- Nowadays, most computers are connected to networks and the internet.
- Thus, ensuring security is more difficult.
- A **hacker** (黑客) may attack using the internet or the network.



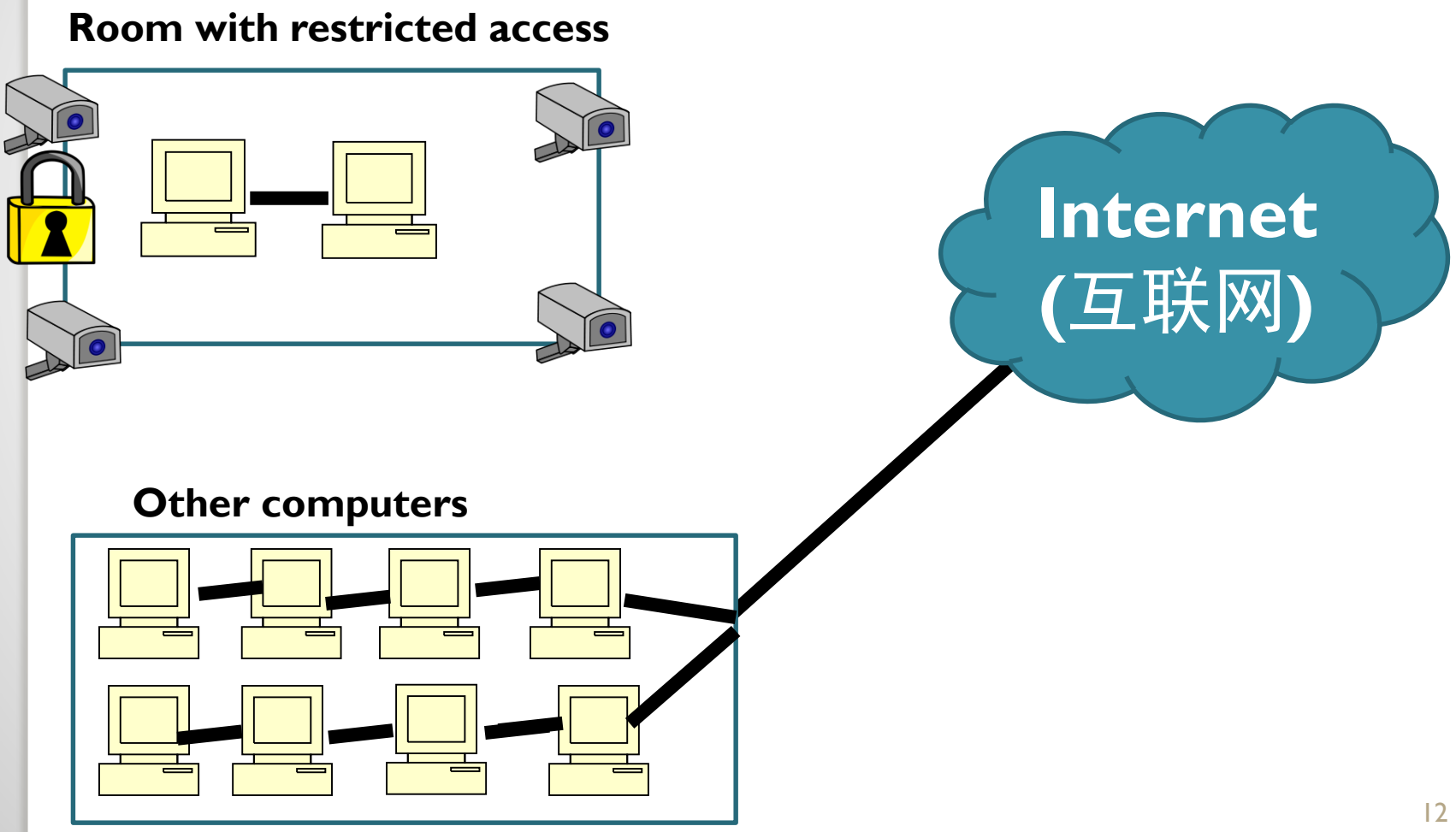
Introduction

Thus, big organizations often store **sensitive data** (敏感数据) on computer(s) that are not connected to the internet or networks.



Introduction

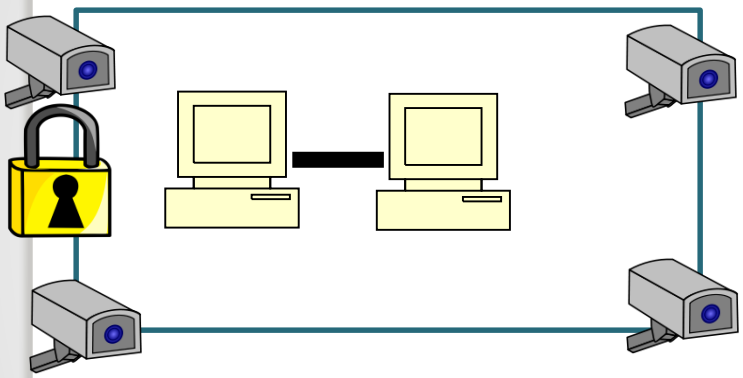
A person who enter the restricted access room may not be allowed to bring **USB flash drives** (USB闪存驱动器), cellphone (手机), etc.



Introduction

Is this a perfect solution?

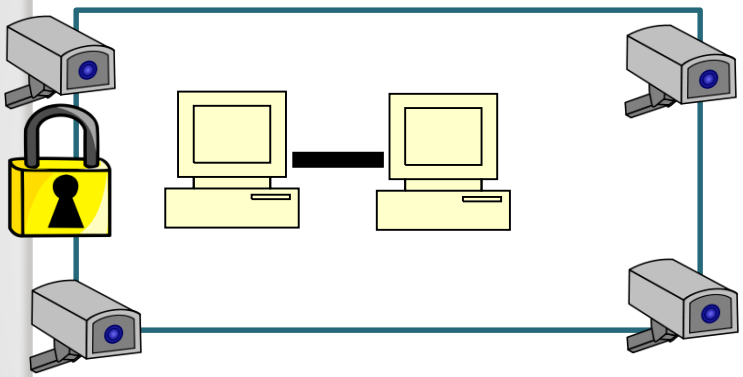
Room with restricted access



Introduction

Is this a perfect solution?

Room with restricted access



No!

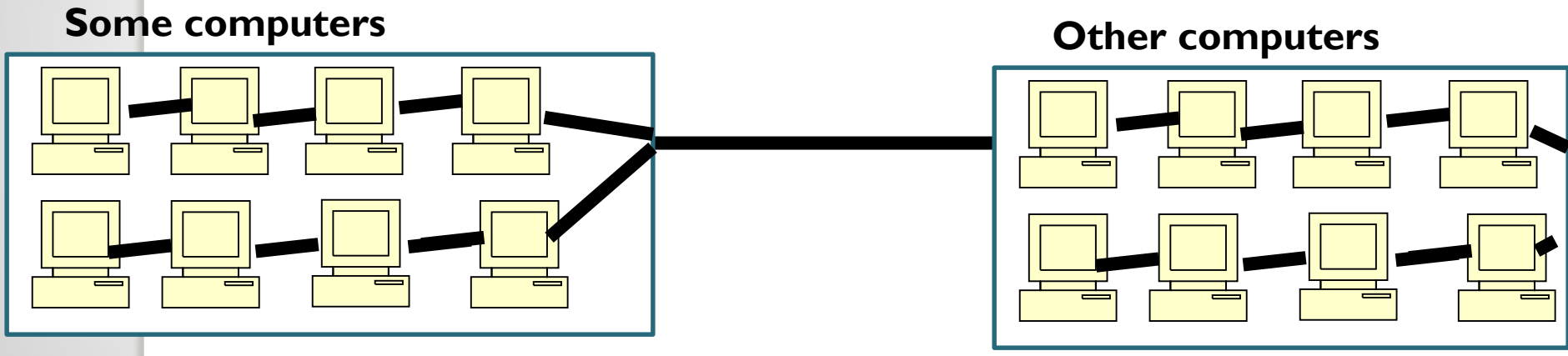
There are many other risks.

For example:
hackers may listen to **electromagnetic radiations** (电磁辐射) to obtain sensitive data
https://www.techrepublic.com/article/air-gapped-computers-are-no-longer-secure/

Introduction

We also need to protect not just the computers but also the **network equipment.**

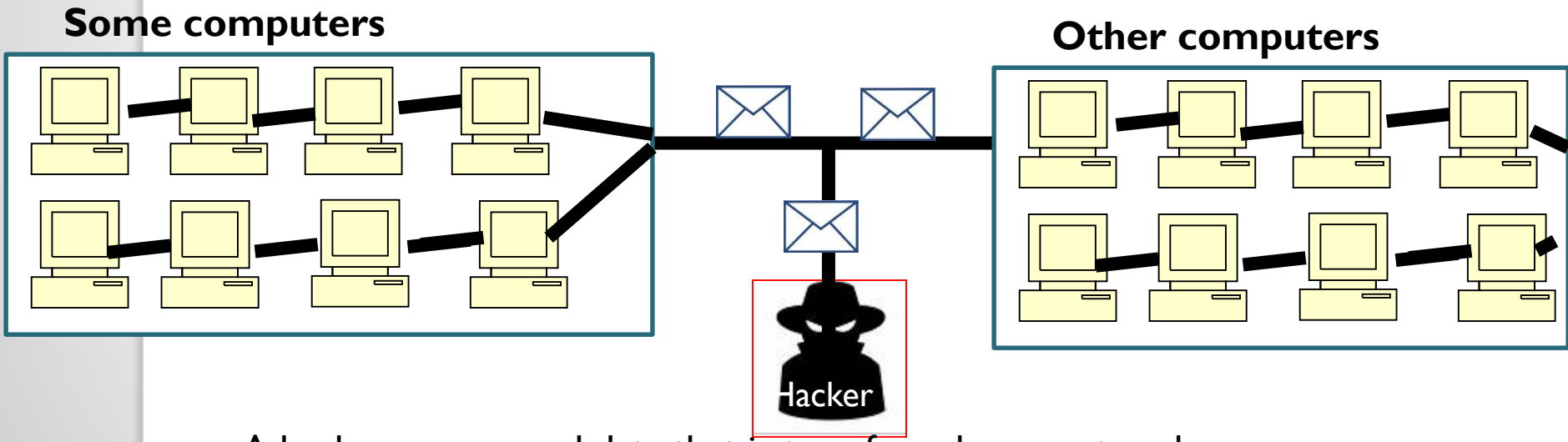
- Network cables (网络电缆)
- Routers (路由器), switches (网络交换机)…



Introduction

We also need to protect not just the computers but also the **network equipment.**

- Network cables (网络电缆)
- Routers (路由器), switches (网络交换机)…

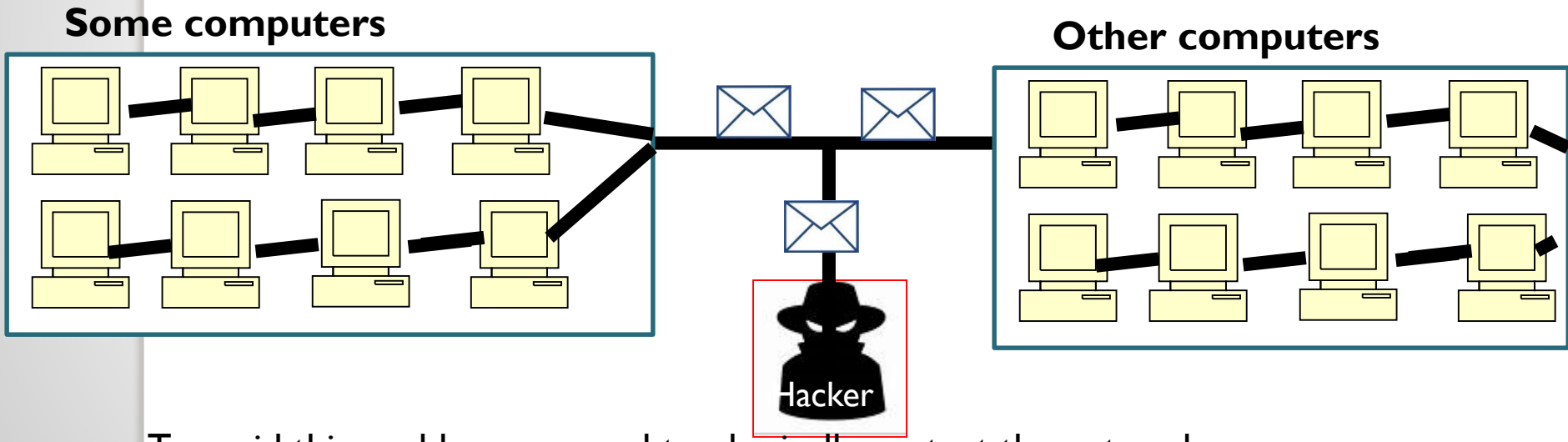


A hacker may record data that is transferred on a network.

Introduction

We also need to protect not just the computers but also the **network equipment.**

- Network cables (网络电缆)
- Routers (路由器), switches (网络交换机)…

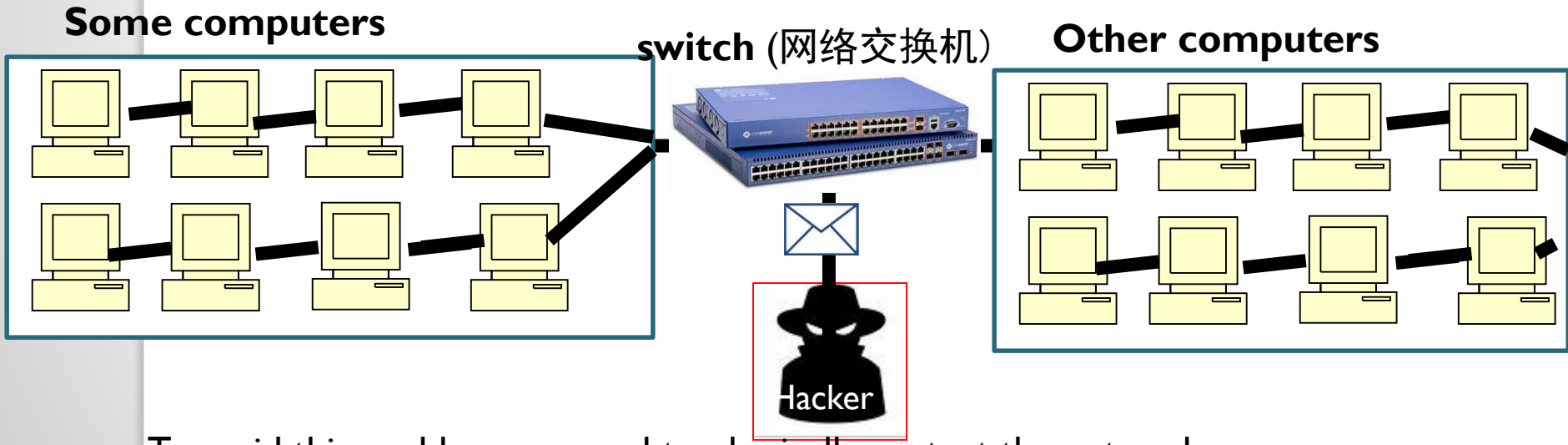


To avoid this problem, we need to physically protect the network.
Moreover, we can use data encryption (数据加密) and other security measures.

Introduction

We also need to protect not just the computers but also the **network equipment.**

- Network cables (网络电缆)
- Routers (路由器), switches (网络交换机)…

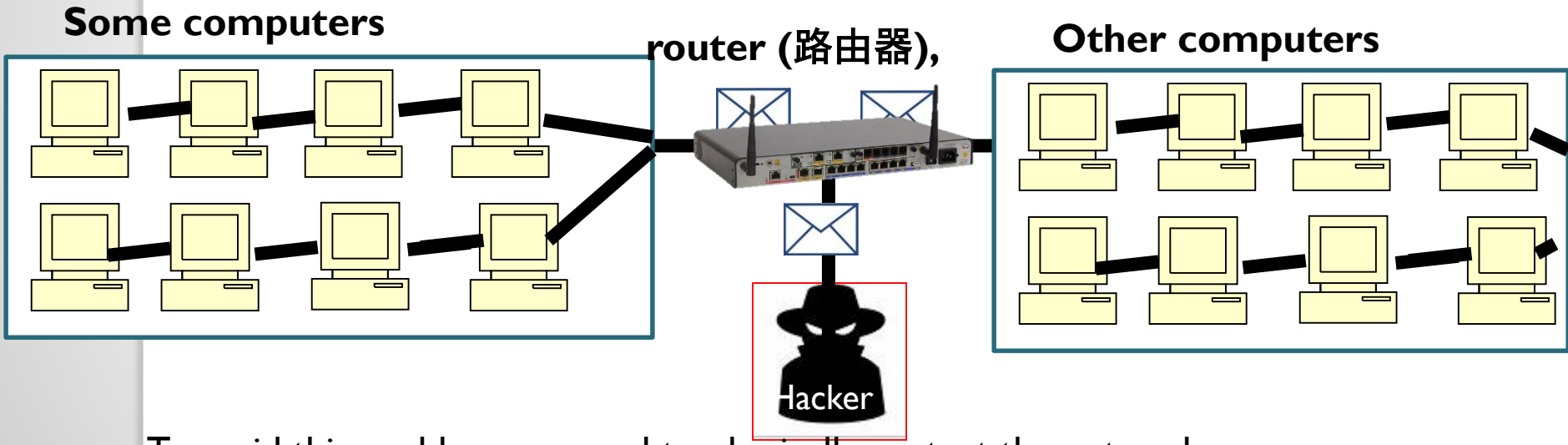


To avoid this problem, we need to physically protect the network.
Moreover, we can use data encryption (数据加密) and other security measures.

Introduction

We also need to protect not just the computers but also the **network equipment.**

- Network cables (网络电缆)
- Routers (路由器), switches (网络交换机)…

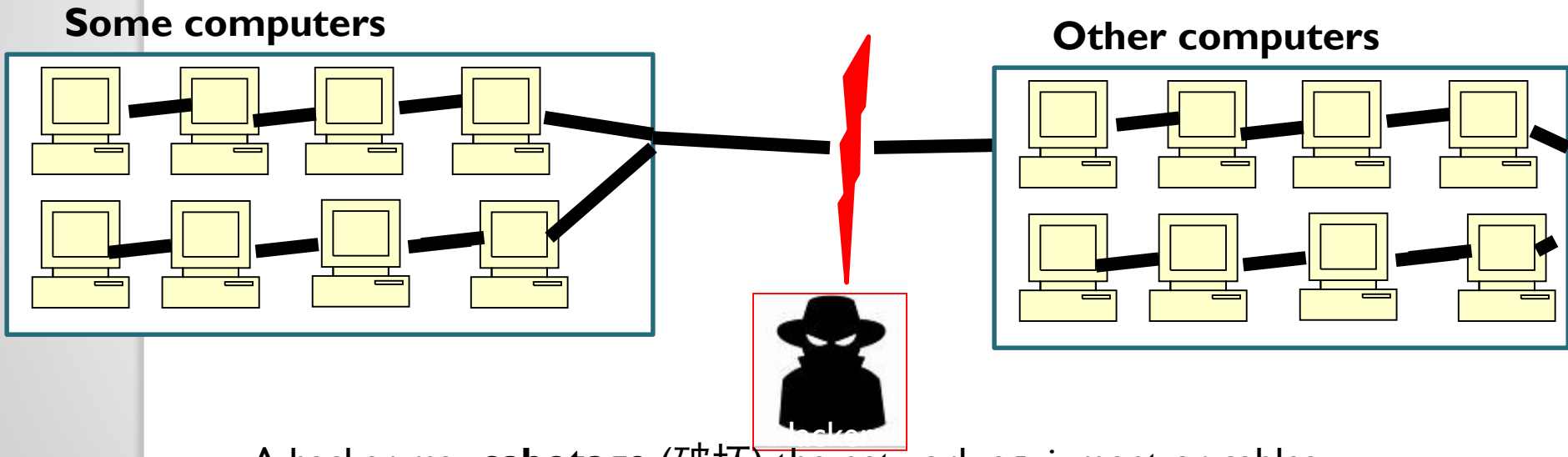


To avoid this problem, we need to physically protect the network.
Moreover, we can use data encryption (数据加密) and other security measures.

Introduction

We also need to protect not just the computers but also the **network equipment.**

- Network cables (网络电缆)
- Routers (路由器), switches (网络交换机)…



A hacker may **sabotage** (破坏) the network equipment or cables

What are the security threats (安全威胁)?

Many types of **security threats** (安全威胁):

- **malware** (恶意软件): software designed with a malicious intent (恶意目的) to destroy data, spy on the user, etc.
- **viruses** (病毒): a computer program that makes copy of itself and can infect other computers.
- **ransomware** (勒索软件)
- **hackers** (黑客),
- **cyberwarfare** (网络战): a country that attack another country's computer network to cause damage, disruption or steal information.

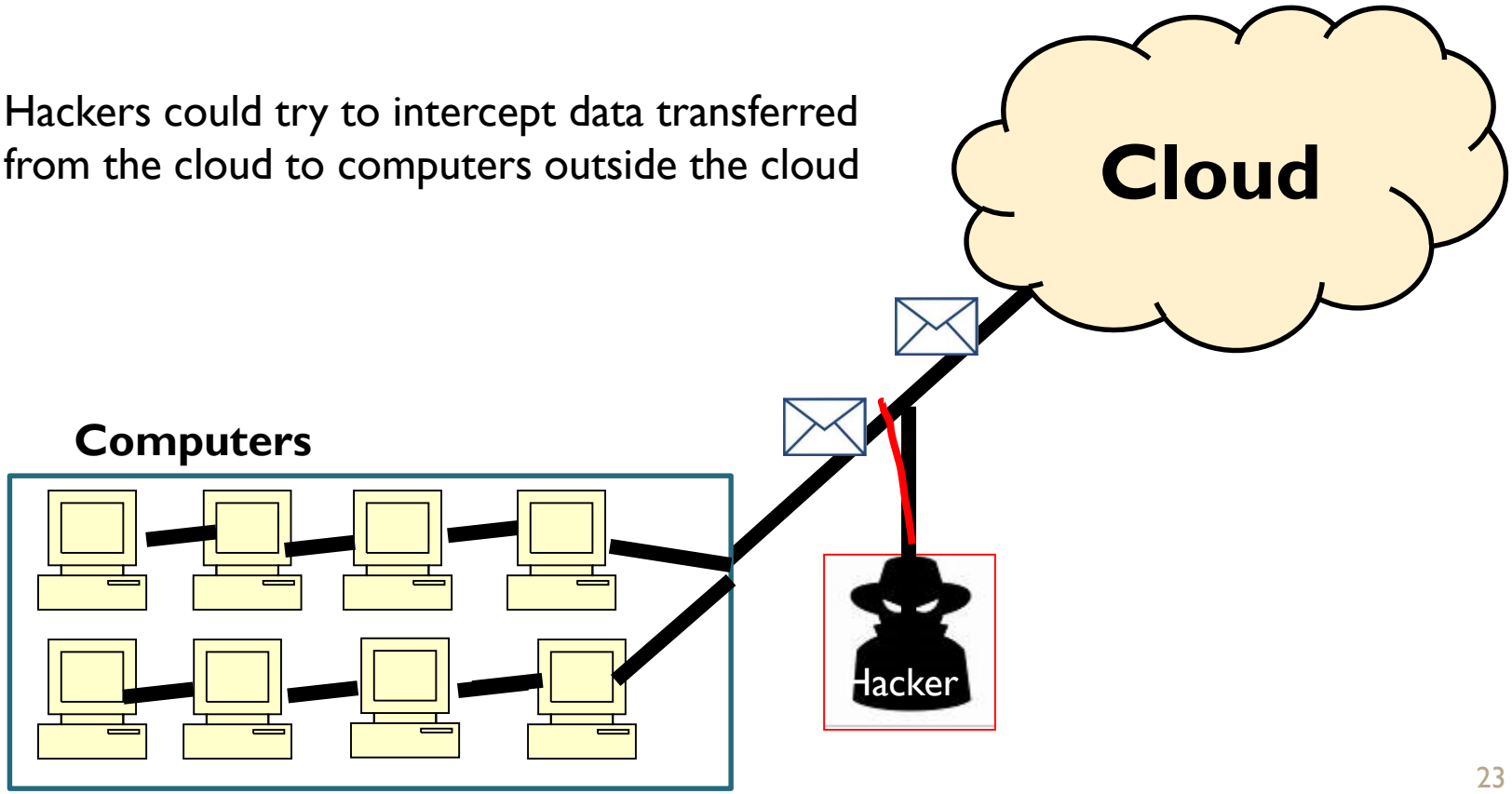
Security in the cloud (云安全)

- Computers/data are in the cloud.
- Using the cloud **reduces the risk** that someone from an organization has **physical access** to the data or a computer system (**insider threat**).
- **Insider threat** (内部威胁): someone from an organization attacks the organization.
- **Is the cloud safe?**
  - Big cloud companies, certainly.
  - Smaller companies, maybe not.

Security in the cloud (云安全)

The cloud also creates new security and privacy concerns.

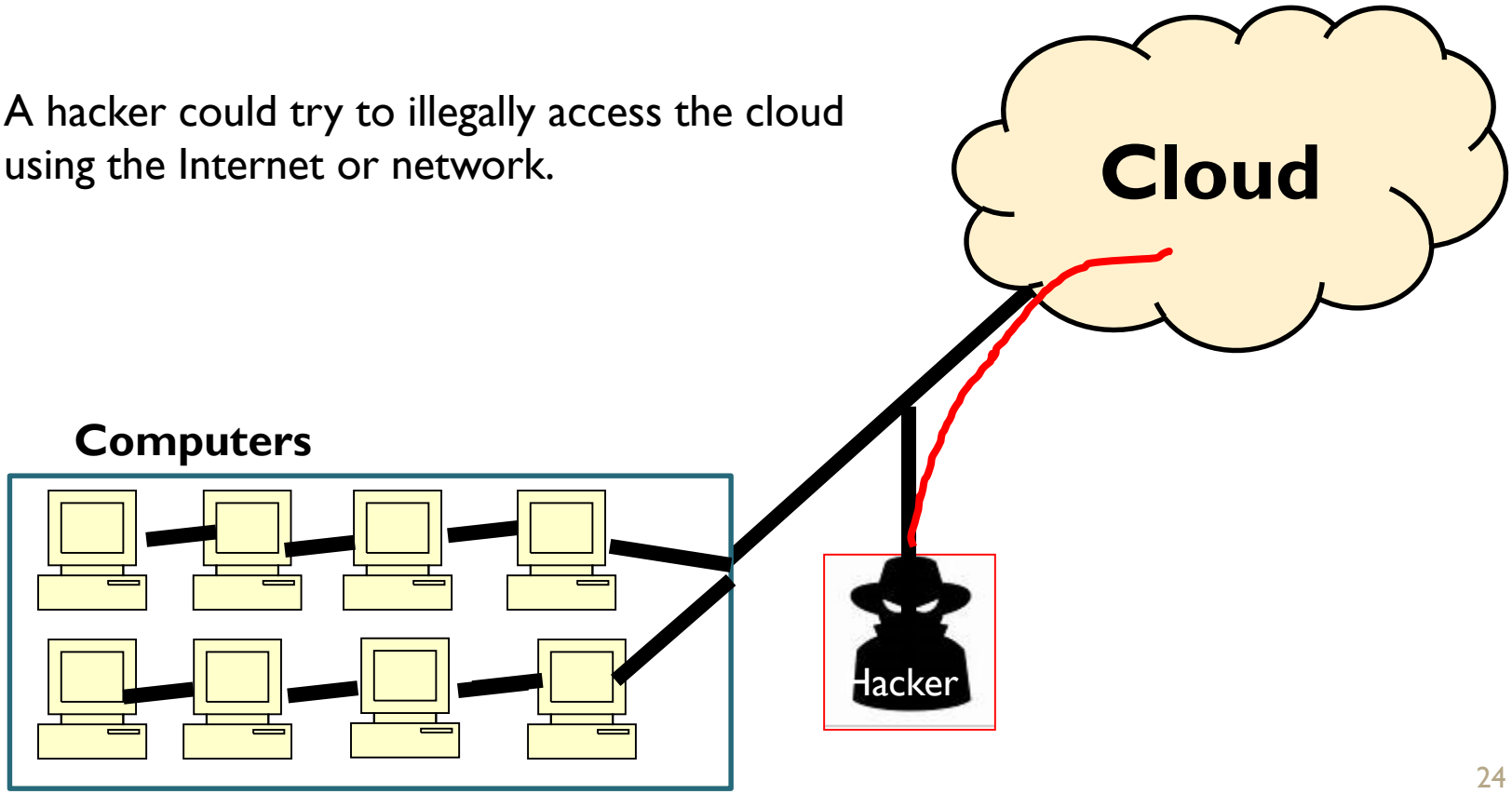
Hackers could try to intercept data transferred from the cloud to computers outside the cloud



Security in the cloud (云安全)

The cloud also creates new security and privacy concerns.

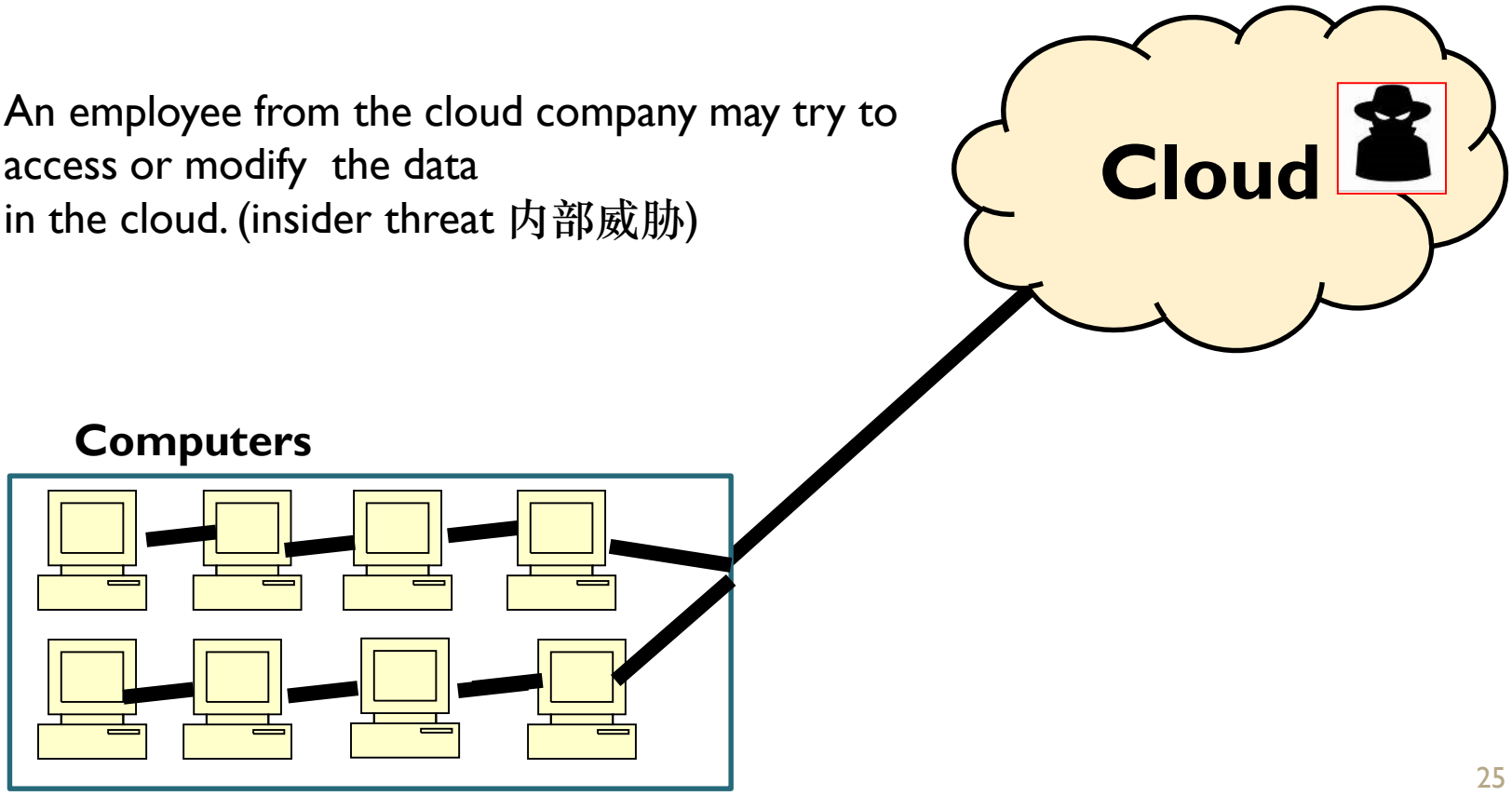
A hacker could try to illegally access the cloud using the Internet or network.



Security in the cloud (云安全)

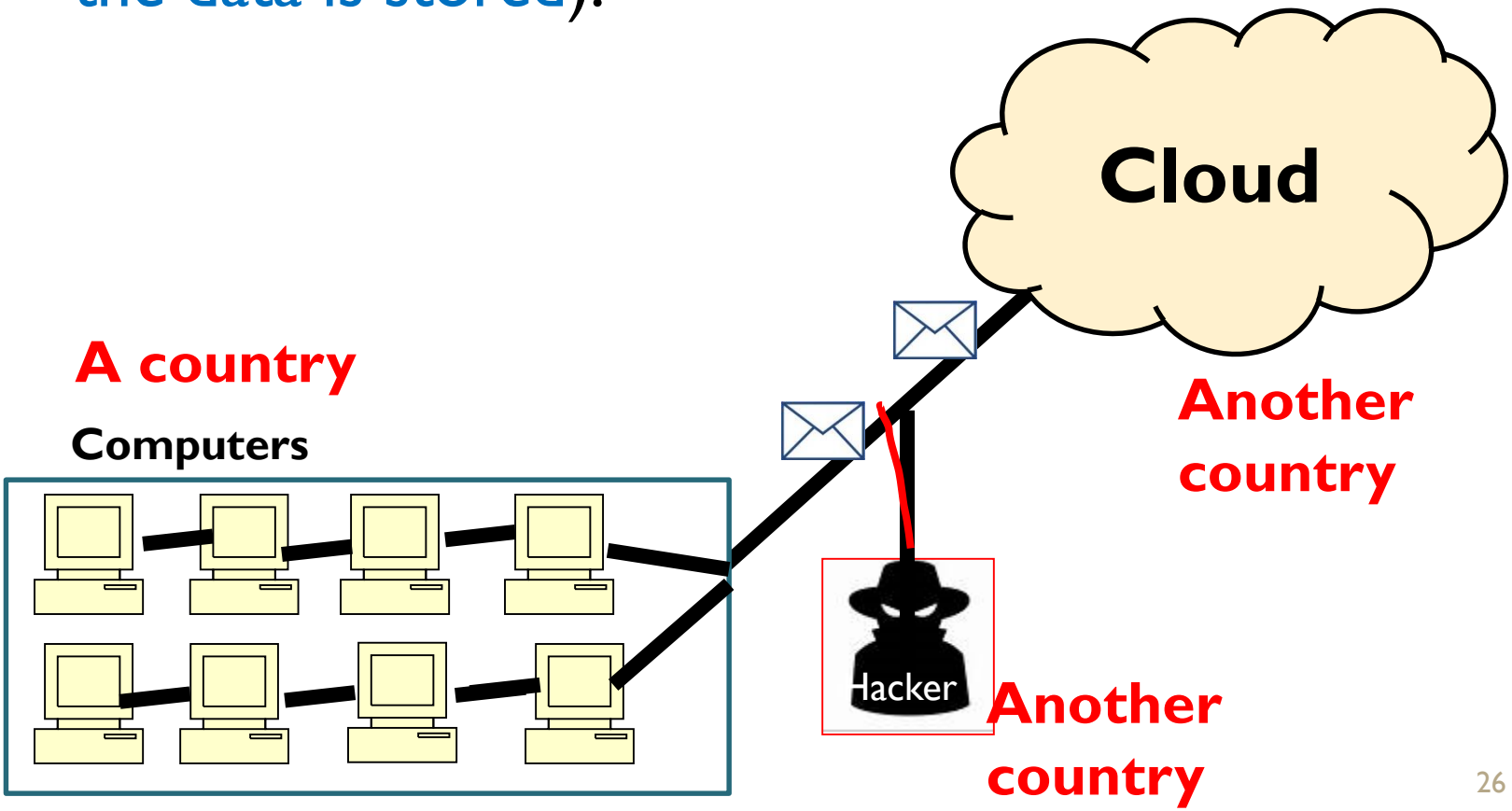
The cloud also creates new security and privacy concerns.

An employee from the cloud company may try to access or modify the data in the cloud. (insider threat 内部威胁)



Security in the cloud（云安全)

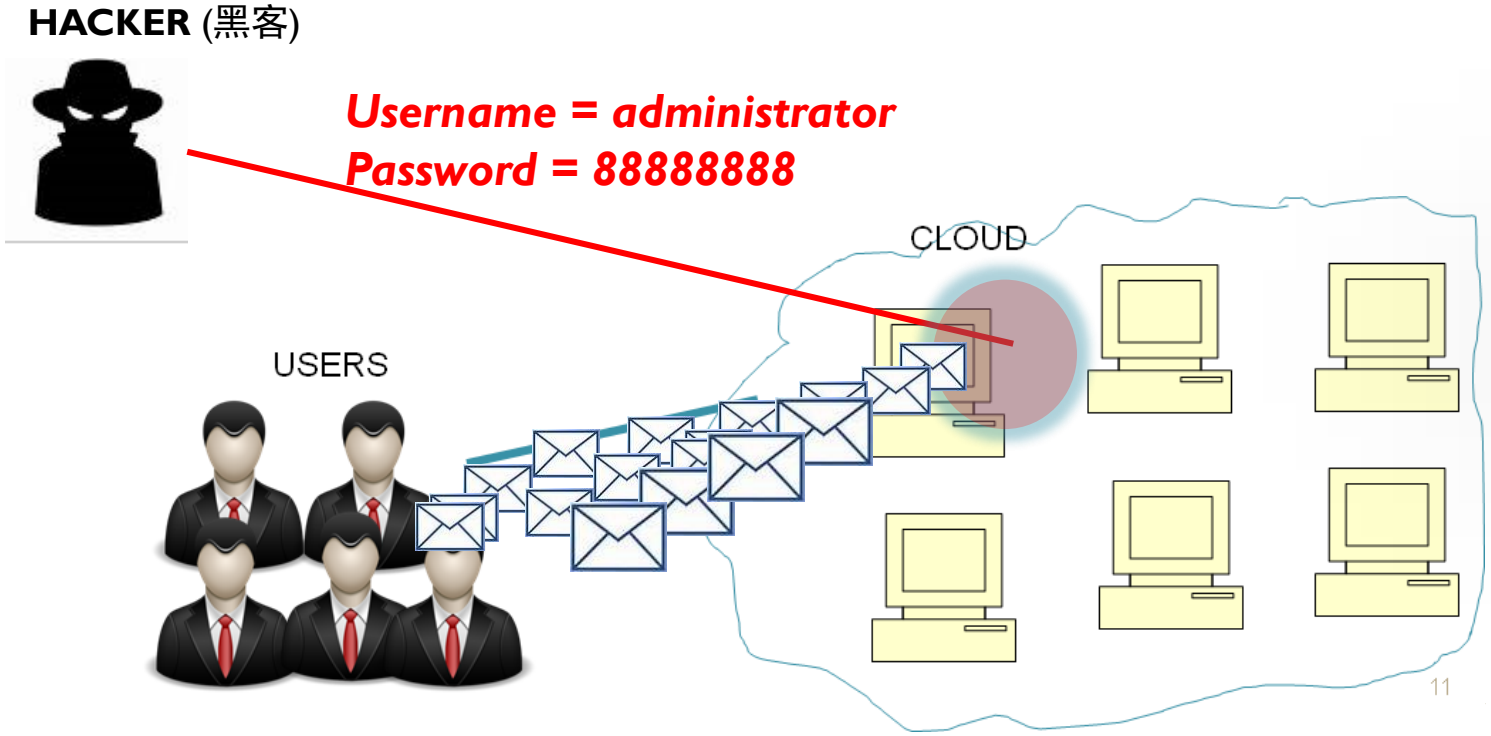
Using the cloud also raise **legal concerns** (法律的问题) (**e.g.** depending on the countries where the data is stored).



Cloud security risks

It is easy to start using the cloud without understanding security risks.

- Security risks for users
- Risk that the cloud is used to launch large attacks,



Traditional threats (传统威胁)

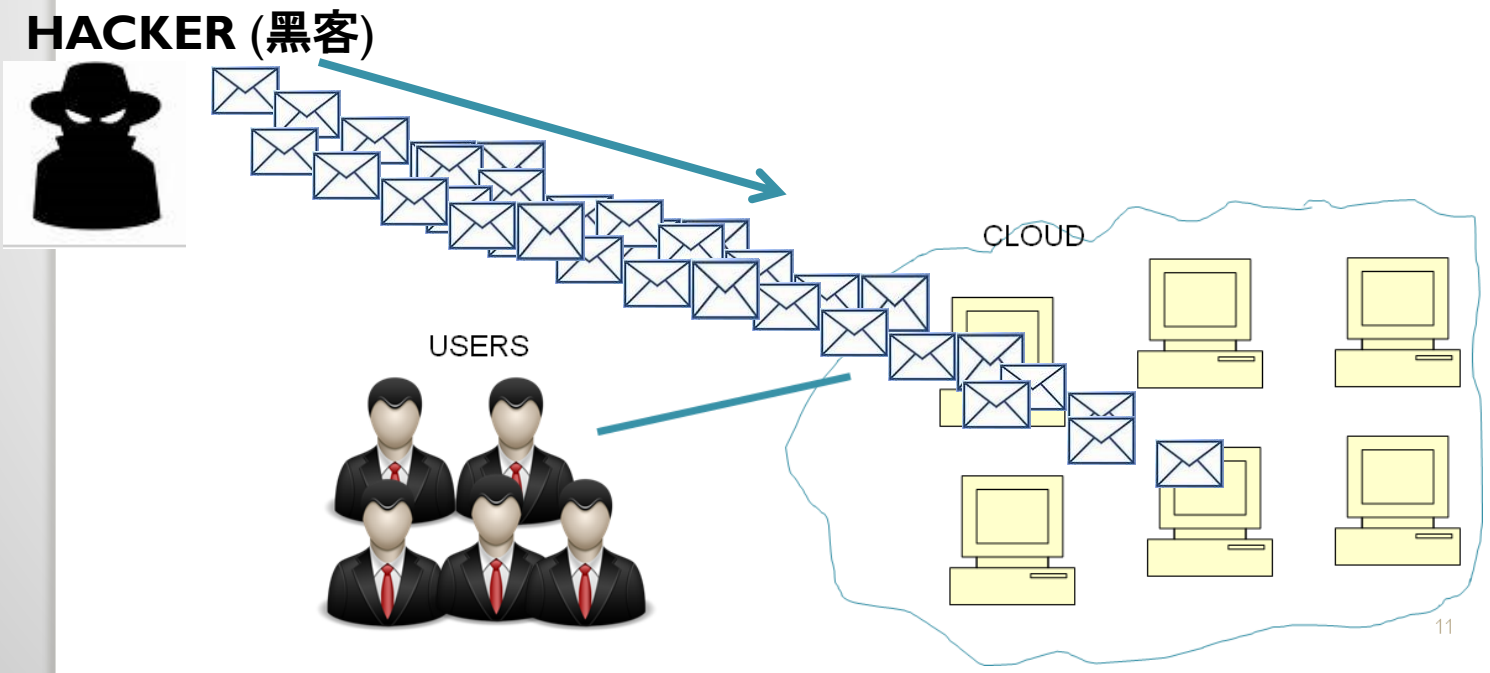
- Any computer system connected to the internet faces some security threats.
- Threats are greater in the cloud, because there is:
  - more users
  - more resources are available.
- The user must:
  - protect the infrastructure used to connect to the cloud,
  - and interact with cloud applications.
- This is difficult because some parts of the computer system are outside the organization (when using a public cloud)

Threats related to authentication (认证) and authorization (授权)

- An organization may have many users who access the same cloud applications.
- Different **levels of privileges** should be assigned to different persons based on their roles in the organization.
  - **e.g.** the accounting department of a university should not have access to modify student grades
- But it is not easy to adapt **security policies** of an organization to the cloud.

Some popular types of traditional attacks, used on the cloud

Denial-of-service (拒绝服务 DDOS): sending thousands of messages to the cloud so that it becomes very busy and cannot provide service to its users.



Some popular types of traditional attacks, used on the cloud

- **Phishing (网络钓鱼):** a bad person creates a fake website to ask users to enter their personal information (credit card number, etc.)
- **Cross-site scripting (跨站点脚本):** a hacker inserts scripts into a webpage to bypass access control or collect information.

Finding who the attacker is?

- It is more difficult to find where an attack come from in the cloud.
- Why?
  - the cloud is a complex system with multiple virtual machines interacting with each other,
  - resources are shared by many users.

Availability (可用性) of cloud services

- The cloud **should** always be **available**.
- Potential threats:
  - system failures (系统故障),
  - power outages (停电),
  - catastrophic events (灾变事件)
    (flood, earthquake, fire, etc.)
- As a result, cloud services could be shut down for long periods of time.
- This could prevent an organization from functioning properly.
- This could damage the reputation of an organization, and cause the loss of sales

The issue of third-party control

- **Cloud providers may lack transparency** (透明度).
- A **cloud provider** may **subcontract** (外包) some resources from a third party (第三方) who should maybe not be trusted.
- Some subcontractors (分包商):
  - may fail to keep customer data.
  - may use poor quality storage devices.
  - may not keep enough copies of your data.

Insider threat

- Usually cloud providers do not tell what are their **hiring standards** (招聘标准).
  - background check? (背景调查)
- Someone working for the cloud provider may try to access your data.

Not taking responsibility

- Some cloud providers could access your data
- Usually, **cloud providers** take **no responsibilities for data security**.
- The **user** is **responsible** of data security.
- For example, the **terms of service of Amazon**: ""We ... will **not be liable** to you for any direct, indirect, incidental ... damages ... nor ... be responsible for any compensation, reimbursement, arising in connection with: (A) your inability to use the services ... (B) the cost of procurement of substitute goods or services ... or (D) any unauthorized access to, alteration of, or deletion, destruction, damage, loss or failure to store any of your content or other data."

Proving responsibility

If may be very difficult to prove that a service provider has deleted your data.

Cloud may not be secure

- The method provided by the cloud provider for accessing the cloud may not be secure.
- Some hackers may steal the passwords and usernames of users using the cloud to gain access.

Cloud delivery models (review)

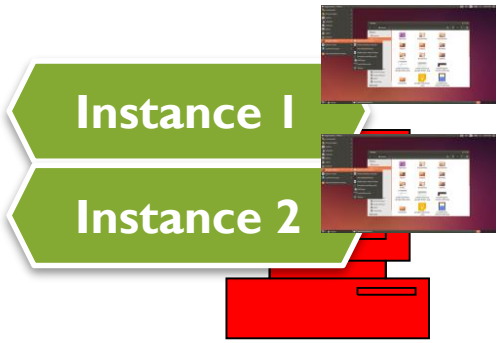
1. **Software-as-a-service**
   - The user pay to store his **data** in the cloud or use an applications provided by the cloud provider (e.g. use an e-mail service)
2. **Platform-as-a-service**
   - The user may install his own **applications** in the cloud (e.g. install an application to manage customer relationships)
3. **Infrastructure-as-a-service**
   - The user pay to use resources, and may install his own **operating system** (e.g. Linux, Windows) and his own **applications,** and may control the network to some extent.

Security for the cloud models

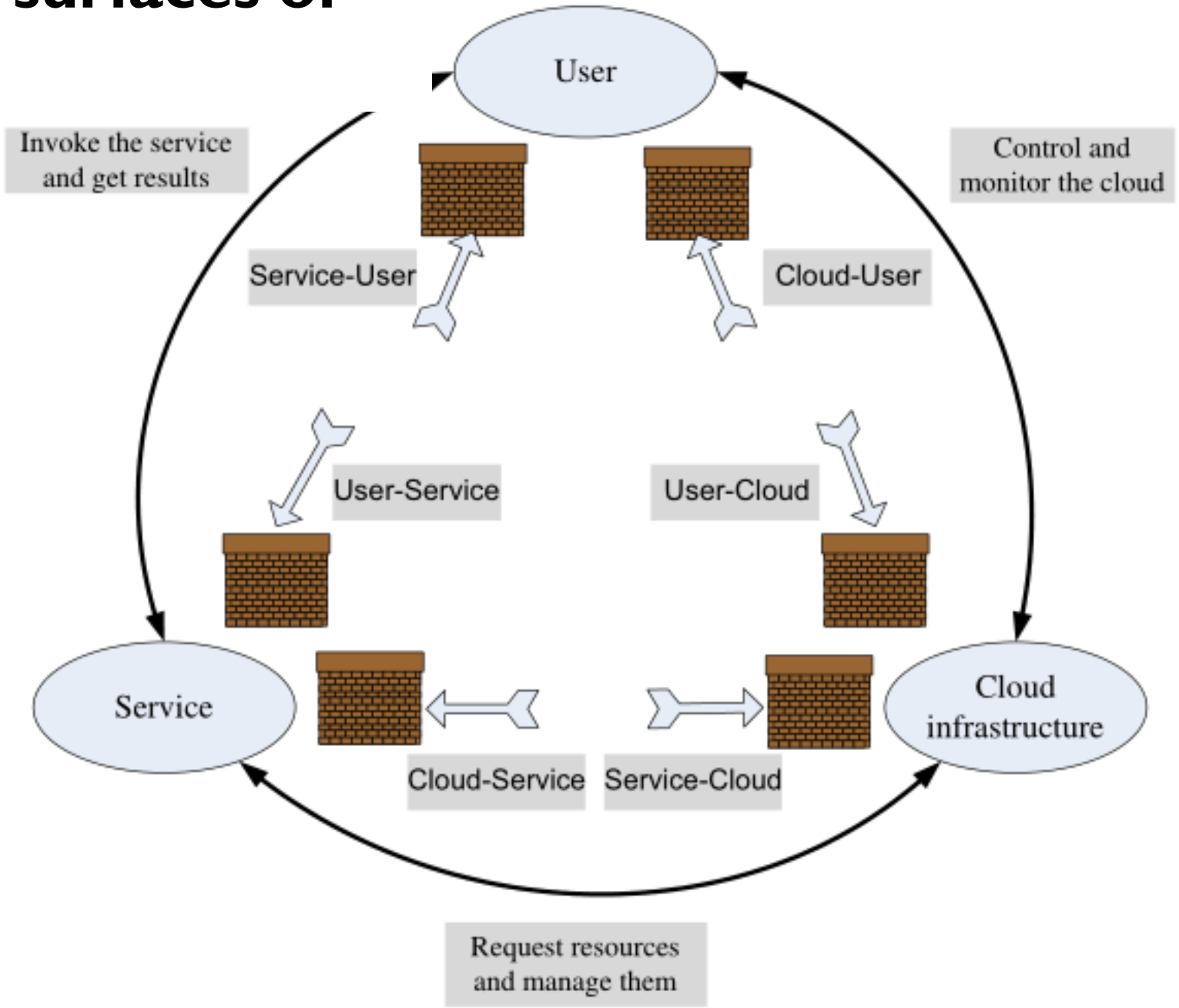
- ***Software as a service (SaaS):***
  - the most secure
  - the cloud cannot be used to launch attacks and no threats from shared technology
- ***Platform as a service (SaaS):***
  - more secure than laaS,
  - no threats from shared technology
- ***Infrastructure as a service (laas):***
  - the most vulnerable.
  - the cloud may be used to launch attacks.
  - threats resulting from shared technology

Shared technology threat

- A computer in the cloud may run multiple virtual computers (instances).
- Some **virtual machine managers** may have security flaws, which may affect the security of the computer and the other virtual machines.



Several surfaces of attack



Top concern for user (最关心的问题): security

- **Surveys** report that **security is the top concern** for cloud users.
- Users must **trust** the cloud service provider to **benefit from the economical advantages** provided by the cloud.
- **Major concerns**: unauthorized access to confidential information and data theft.
- **Data is more vulnerable in storage** than while it is being processed because it is processed for short periods of time.
- Close attention should be paid to the security of **storage servers** and to **data in transit.**

Another concern: cycle of data

- The user cannot make sure that data that should have been deleted is deleted.
- Even if deleted, another user may be able to recover the data, if it is not deleted properly.

Protecting personal information

- For users, it is important to protect **personal personal information** from malicious persons such as: names, birthdates, IDs, credit card numbers…
- If the cloud provider is in another country, the laws may be different.
- It may be difficult for a user to understand how the laws of another country apply to his data.

Some solutions

- A company may decide to **avoid processing sensitive data on the cloud** to reduce security risks.
- Another solution is to **encrypt** the **data** before storing it in the cloud
  (using legal software; and it may require to apply for a permit).
- Another solution is to use a **private cloud.**

Trust (信任)

- Similar to the problem of trust on the internet (online banking 网上银行, etc.).
- What is trust?
  - **A perceived risk** (**e.g.** probability of losing data, probability of hackers stealing data)
  - **Interdependence (相互依存)**: the interest of an entity cannot be achieved without reliance on the other entity. (**e.g.** benefits of using the cloud)
- Types of trust →

Types of trust

- **Deterrence-based trust** (基于威慑的信任): the penalty for breaching trust exceeds any potential benefits that one could get from breaching trust.

- **Calculus-based trust**: it is believed that it is in the other party best interest to not breach trust.

- **Relational trust** (关系信任): two entities work together for a long time and both rely on each other. Thus, they know they can rely and depend on each other.

Online trust

- Unlike traditional trust, when using a cloud service, there might be **no person to person interaction** involved in trust.
- The Internet offers individuals the ability to **change** or **hide their identities**.
- Because of this, traditional ways of building trust are not applicable in an online environment.
- In real-life, trust is often based on **accountability** (问责), which requires to know the identity of a person.

Some solutions

- **Digital signatures** (数字签名), **digital certificates** (数字证书) to prove that some entity is who it claimed to be.
- Digital certificates (数字证书) are provided by trusted organizations.
- Evaluating the **reputation of a cloud provider** (long history, other users…).

Operating system security

- An **operating system** (操作系统) allows multiple applications to share the hardware resources of a computer.
- An **operating system** may offer several security features:
  - user authentication (用户认证),
  - different levels of **privileges**（安全特权 ） for users, and for applications,
  - Cryptography (密码), …
- Applications should just have enough privileges to perform their task.

Related issues

- Should an operating system be able to protect the user from malicious software such as viruses, etc. ?
- Some hardware like **ATM** (自动取款机), **cellphones**, **webcams** (摄像头), **video game consoles** (电子游戏机), may have **security flaws** (安全漏洞) that cannot be fixed.
- In general, operating systems provide weak security. Thus, application security is more important.

Virtual machine security (虚拟机安全)

- Using **many virtual machines** (虚拟机) is safer than running multiple programs in the same operating system.
- However, **virtual machines** running in an **operating system** can be attacked using security vulnerabilities (安全漏洞) of the operating system.
- A virtual **machine manager** (虚拟机管理器) will provide some security features.
- **Virtual machine managers** are often much less complex than traditional operating systems. Thus, it is easier to ensure security.

Virtual machine security

- It is possible to make a copy of a virtual machine and use it to test for malicious behavior.

- Using virtual machines is **safer**…. but it **increases costs** (it requires more expensive hardware, more CPU time, memory, network bandwidth…)

Some virtual machine attacks

1. A virtual machine bypasses the resource limits and uses all the resources available by the VMM and starves other virtual machines.
2. *Buffer overflow:* a virtual machine use a special type of attack to access the memory of another virtual machine.
3. …

ADAPTIVE CYBER-DEFENSE

Based on the book « Adversarial and uncertain reasoning for adaptive cyber-defense » (2019)

Introduction

- Many companies focus on **cost-effectiveness** and **performance** of software or hardware rather than **security**.
- Companies that consider security often focus on:
  - developing better **software** and **hardware** to increase security.
  - Having many layers of security (encryption, access control, firewall, virus scanner, etc.).
- The market often prefers **homogeneity**, **standardization** and **predictability**…

Introduction (continued)

- Thus, most **cyber-defense** (technologies to protect a computer system) are **static (do not change)**, nowadays.
- If a company is attacked, it will usually **slowly** react and then make changes to its cyber-defense.
- Attackers can take advantage of this.
  - They have time to plan their attacks.
  - They have time to look for vulnerabilities.
  - After hacking a computer system, hackers can sometimes use it for a long time.

A solution

- Some security researchers work on designing security systems that are less homogeneous and predictable.
- For example (advanced):
  - Diversity,
  - Randomization of the network address space,
  - Randomization when compiling software,
  - Randomization of memory space and instruction sets,
  - Systems that dynamically changes over time…
  - …

But…

- Complex security solutions may decrease performance,
- Managing systems that are less homogeneous and are dynamic is more complicated and costly.

- Security of a system can be modelled using *game-theory* or *control theory*.

Conclusion about this part

- Prepare yourself well for the final exam.

http://philippe-fournier-viger.com/COURSES/CLOUD/

FINAL EXAM

Final exam

- **120 minutes**.
- It is a **closed-book exam.**
- Questions will be approximately evenly distributed between the different topics that we have discussed.

  8 lectures = 8 to 10 questions

Final exam

- Answers must be written in English.
- Some typical questions in my exams:
  - What is the advantages/disadvantages of using X instead of Y ?
  - When X should be used?
  - How X works ? or why X is designed like that?
  - There could be 1 or 2 questions that are similar to the assignments.

Final exam

- **If you are not sure about the meaning of a question in the final exam** because of English, you may raise your hand to ask me.

- **No electronic devices** are allowed.
- A pen/pencil/eraser can be used during the exam.
- Bring your **student ID card**.

Question 1

What are the advantages and disadvantages of using a **private cloud** (instead of using a **public cloud**)?

Answer 1

- **Advantages**:
  - more secure, better performance for real-time applications,
- **Disadvantages**:
  - requires an investment in infrastructures, need to be operated by the organization,
  - data may be centralized in one location (data loss, transfer speed to outside, etc)

Question 2

What are the benefits of using data replication in cloud computing?

Answer 2

- Reduces the risk of data loss,
- Reduces communication traffic / increase speed,
- Reduces energy consumption (by dispatching computation to area where electricity is cheaper,

Question 3

- In the **Map Reduce model,** what is the role of the "**master instance**" (what does it do?)

Answer 3

- It split the data into parts.
- It gives a data block to each instance.
- Its starts the reducing instances
- It monitor the task completion and the state of the other computers

Question 4

- What is big data? Give a definition.

Answer 4

- **Volume**: a large amount of data
- **Velocity**: the data is arriving at a very high speed (e.g. streaming data)
- **Variety**: the data is of different types such as text, images, audio, video, graphs,….
- **Veracity**: the data may be of poor quality (inaccurate) or not trustworthy.
- **Value:** it is important to try to use big data to obtain some "business value", that is to extract useful knowledge from data.

Question 5

Why is it important for a cloud provider to **monitor performance** in the cloud?

Explain the reasons for monitoring performance.

Answer 5

- “pay for what you use” (each user pay for what he uses),
- to perform load balancing and increase performance
- …

References

- Chaptre 9. D. C. Marinescu. Cloud Computing Theory and Practice, Morgan Kaufmann, 2013.