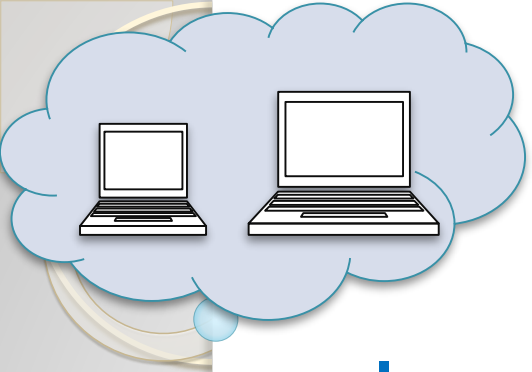


云计算入门

Introduction to Cloud Computing

GESC1001

Philippe Fournier-Viger

Professor

School of Humanities and Social Sciences

philfv8@yahoo.com

Fall 2020

QQ group: 1127433879

Introduction

Last week:

- The **content of this course** (goal, evaluation criteria, rules)
- Introduction to **cloud computing**.

- **Today:**
  - Introduction to cloud computing (part 2)
  - **Chapter 2:** challenges of distributed and parallel systems

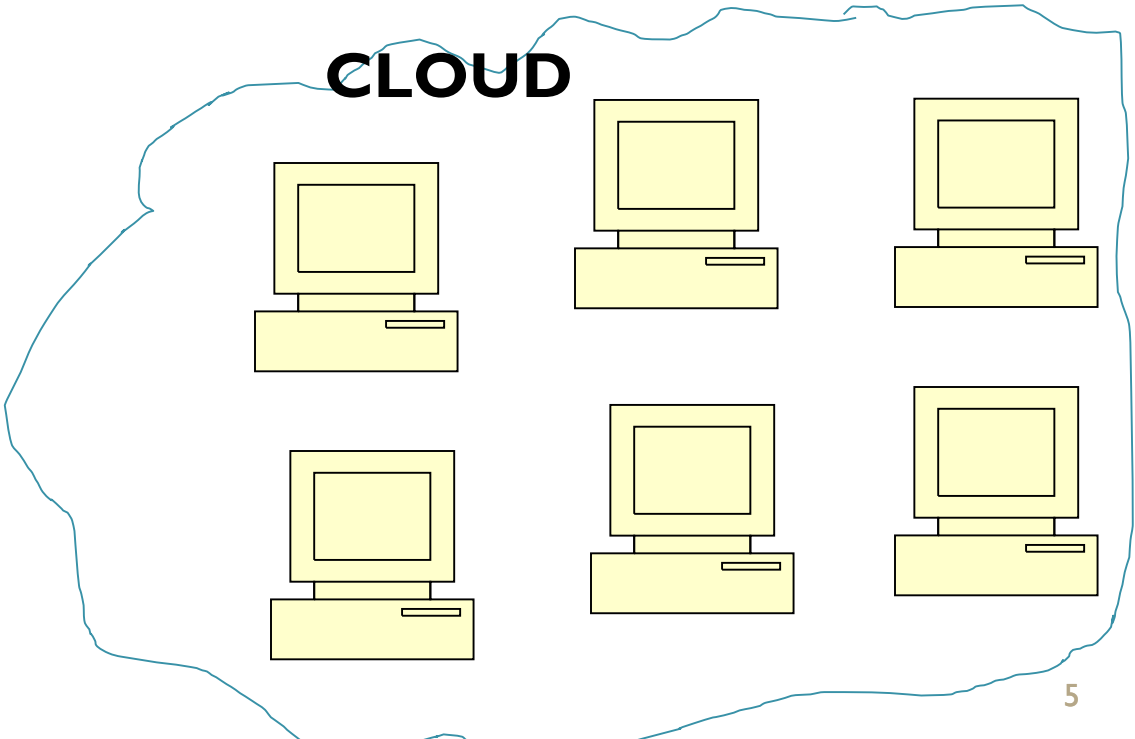
Introduction (cont'd)

The **evaluation for this course** will not be too difficult. But please **study** well.

INTRODUCTION TO CLOUD COMPUTING (PART 2)

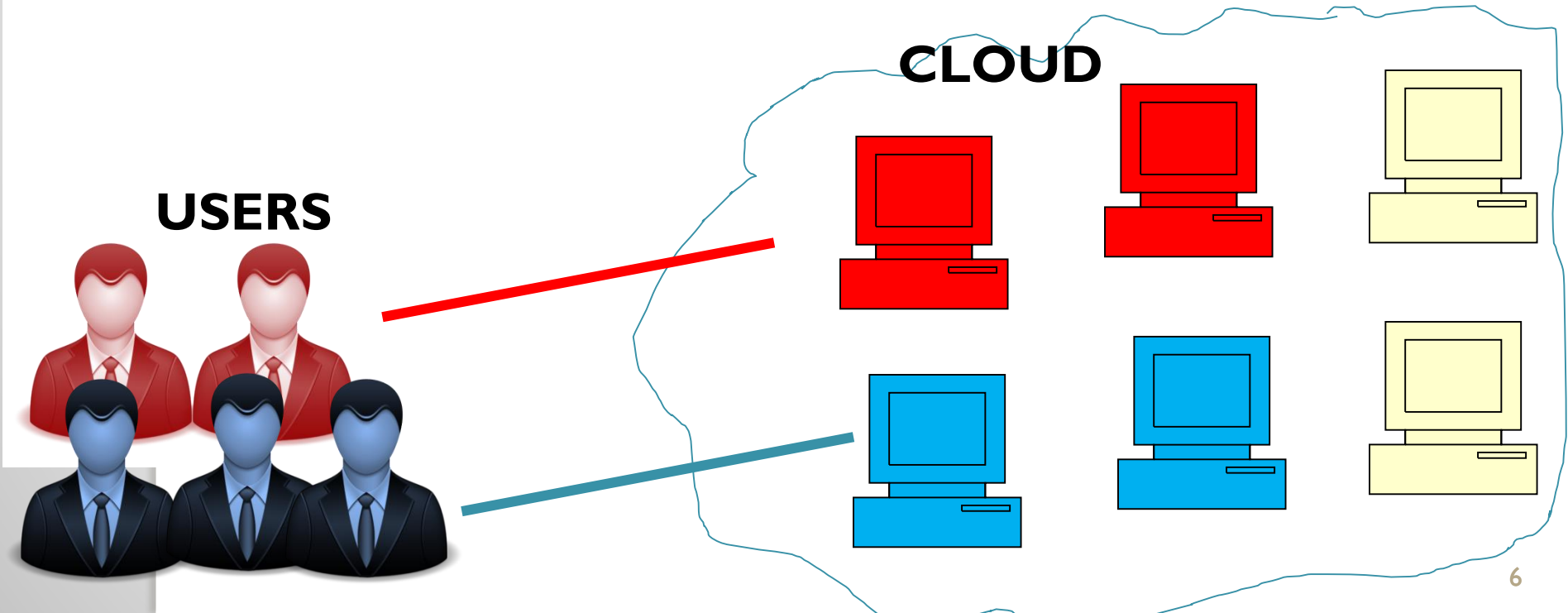
What is cloud computing (云计算)?

Cloud (云): a set of distant (遥远的) computers that provide **computing** or **storage services** to users.



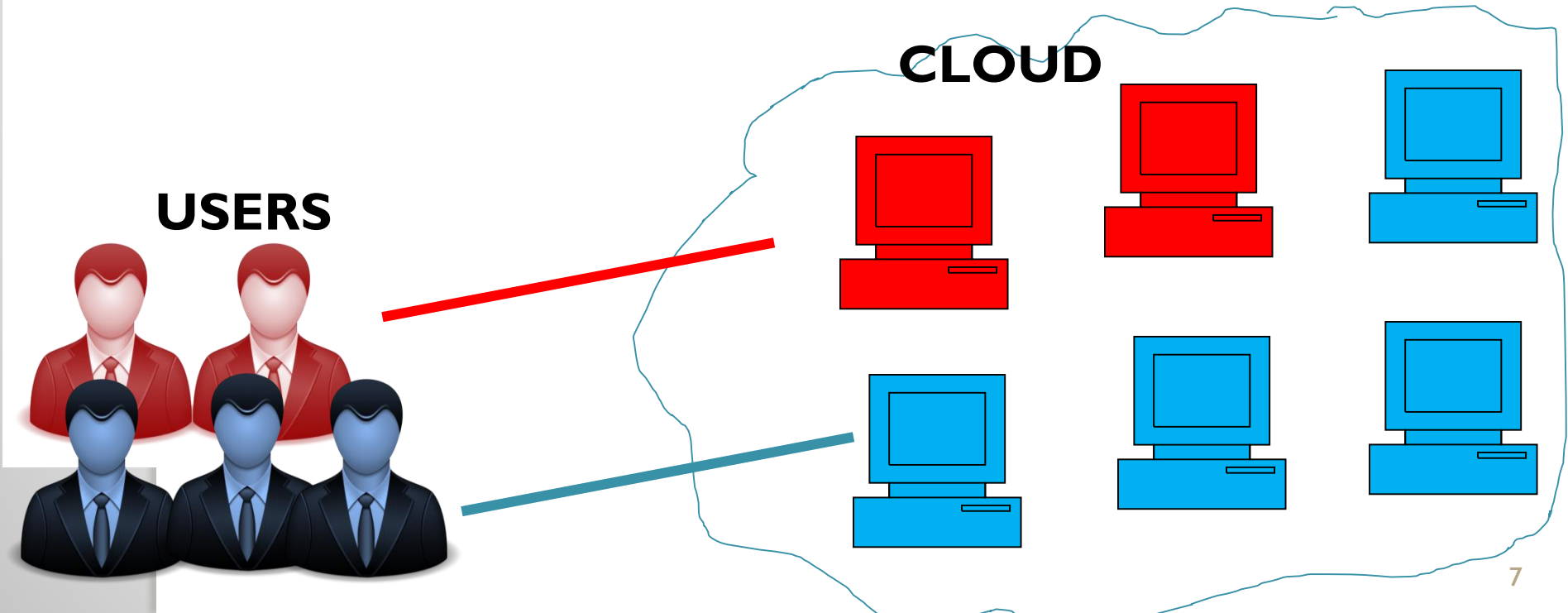
What is cloud computing?

- A **homogenous** (同质) set of computers
- Provides **elastic computing** (弹性计算) capabilities to users
- Each user **pay for what he uses** (元).



What is cloud computing?

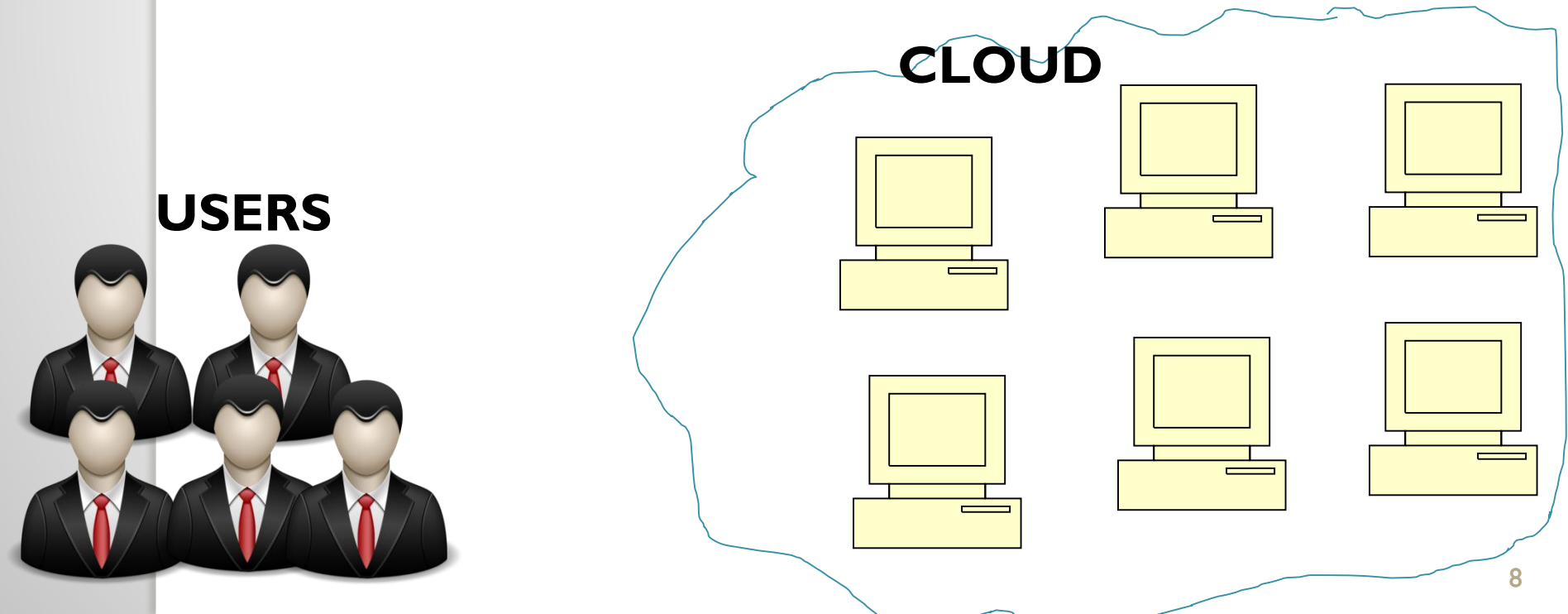
- A **homogenous** (同质) set of computers
- Provides **elastic computing** (弹性计算) capabilities to users
- Each user **pay for what he uses** (元).



Resources

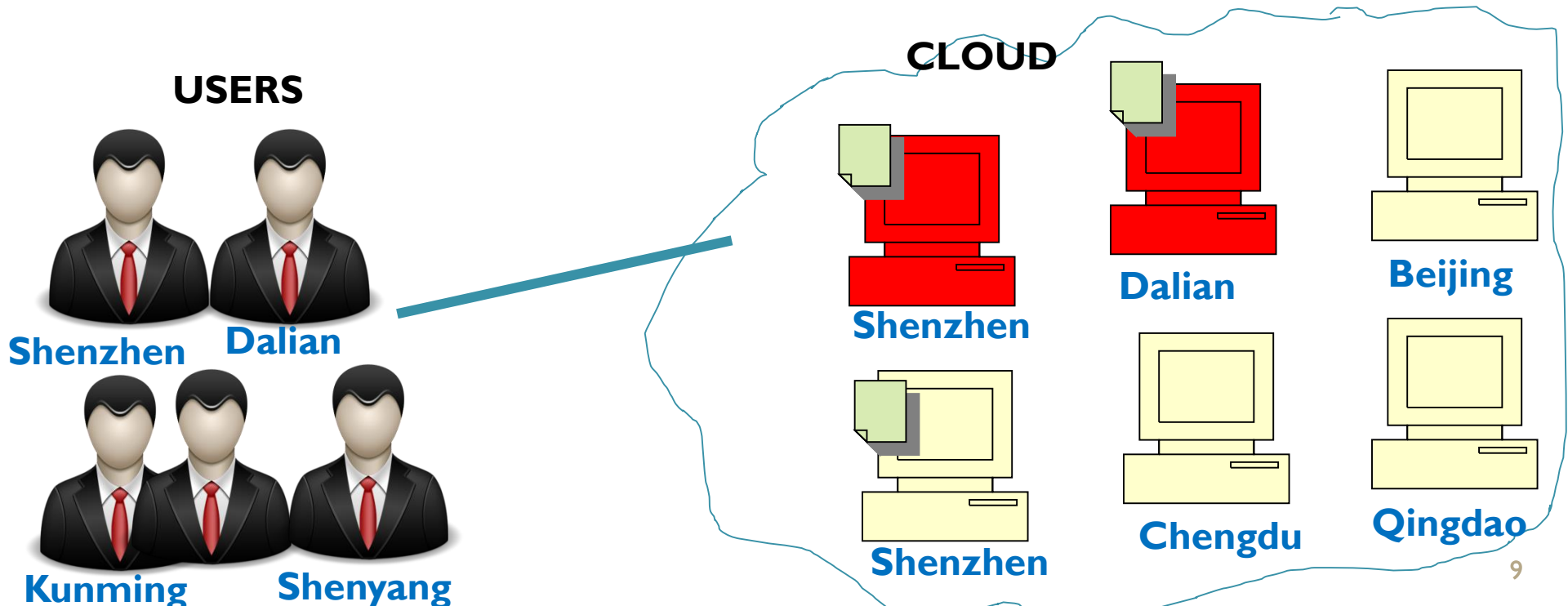
The **cloud** provides 3 types of resources:

- Storage space (存储空间)
- Network infrastructure (网络基础设施)
- Applications and services (软件程序)



What is cloud computing?

- **Public cloud** (公共云) and **Private cloud** (私有云)
- **Data** is stored on **multiple computers**, often **close to the places** where it is used.
- **Multiple copies of the data** can be stored in the cloud.



Cloud provider (云提供商)

- A company or organization that manages a "**cloud**".
- Computers from a cloud are typically located in one or more **data centers (数据中心).**

阿里云计算
Alibaba Cloud Computing

百度云
cloud.baidu.com

What cloud providers do?

- **Manage the cloud**
  - computers, software, etc.
- **Manager customers (客户管理)**
  - billing (发送账单),
  - accounting (会计),
  - call center (呼叫中心),
  - technical support (技术支持),
  - backup, management…

What cloud providers do?

- **Manage security**
  - authentication (认证),
  - certification,
  - intrusion prevention (入侵防御),
  - virus protection (病毒防护),
  - cryptography (密码术),
  - physical security,
  - access control (访问控制),…

Why using the cloud?

- Data is accessible anywhere, from multiple devices (e.g. phone, laptop)
- Easy to share data with other people
- No need to
  - buy equipment.
  - hire employees to manage the computers.
- The cloud is "**elastic**": the user can ask for more resources when needed.
- Generally, cheap (元).

Three ways of using the cloud (delivery models - 云交付模型)

1. **Software-as-a-service**
   - The user stores his **data** in the cloud or use cloud applications (云应用) (e.g. 百度云，use the 163 e-mail service)
2. **Platform-as-a-service**
   - The user may use his own **applications** in the cloud (e.g. install an application to manage customer data)
3. **Infrastructure-as-a-service**
   - The user may install his own **operating system** (操作系统 - e.g. Linux, Windows) and his own **applications,** and may have control on the **network**.

Ethical issues (伦理问题)

If you are using a **public cloud**, **you are giving control** of your data, applications and network **to someone else.**

Risks:

- **Hackers** (黑客) may access your data,
- Your **data** may be **lost** or **corrupted** (损坏的数据),
- **Infrastructure failures** (基础设施故障)
- Service may become unavailable (不可用的服务)
- Hard to find the source of a problem and fix it.

Some solutions…

- Each country should have **rules and regulations** to ensure **responsibility** (负责任) and **accountability** (问责).

- **Cloud providers** should keep a **log** (访问日志) :
  - Who uses cloud data and applications?
  - Who accesses the network?

Cloud vulnerabilities (云的漏洞)

A cloud may become unavailable due to:

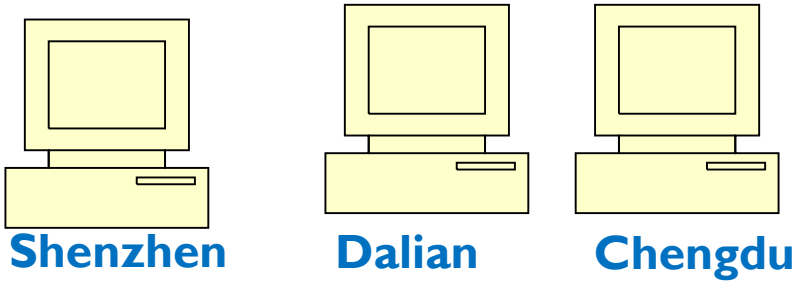
- **malicious attacks (**恶意攻击)
  - **2004: Yahoo** (雅虎) was inaccessible after an attack on **Akamai** in 2004.
- **infrastructure failures** (基础设施故障)
  - **2012:** some **Amazon** (亚马逊) servers went unavailable after lightning strikes. Also caused by hardware and software bugs/deficiencies.
- Because many computers and applications run in the cloud, their **complex interactions** may lead to **unexpected problems**.

How to avoid these problems ? →

Data replication (数据复制)

Storing copies of resources in multiple geographic locations.

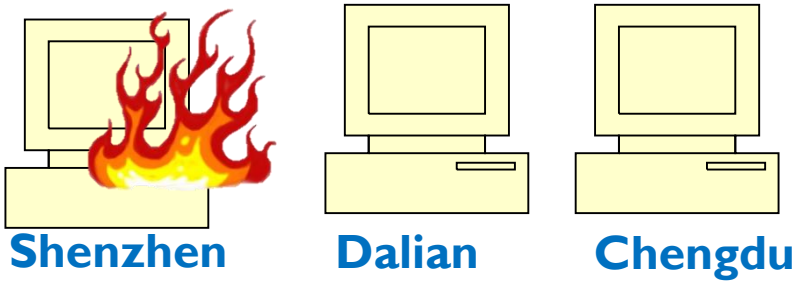
- Reduces the risk of losses,



Data replication (数据复制)

Storing copies of resources in multiple geographic locations.

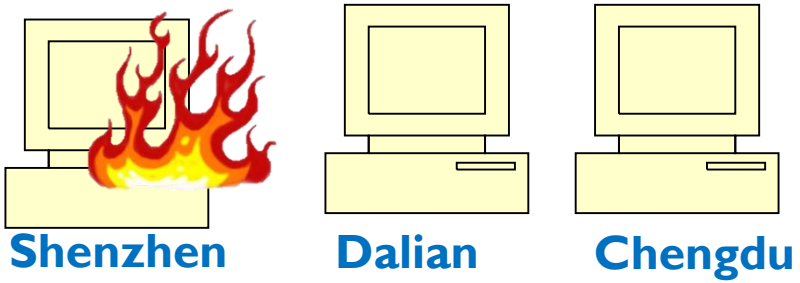
- Reduces the risk of losses,



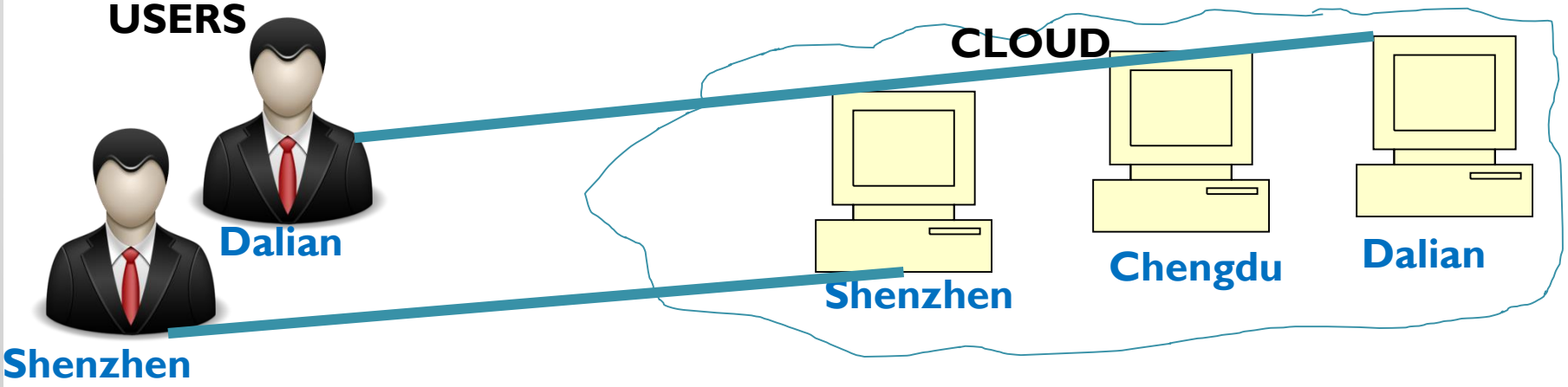
Data replication (数据复制)

Storing copies of resources in multiple geographic locations.

- Reduces the risk of losses,



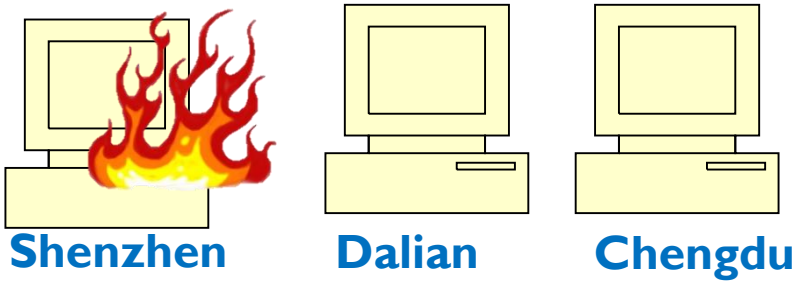
- Reduces communication traffic (网络通信),



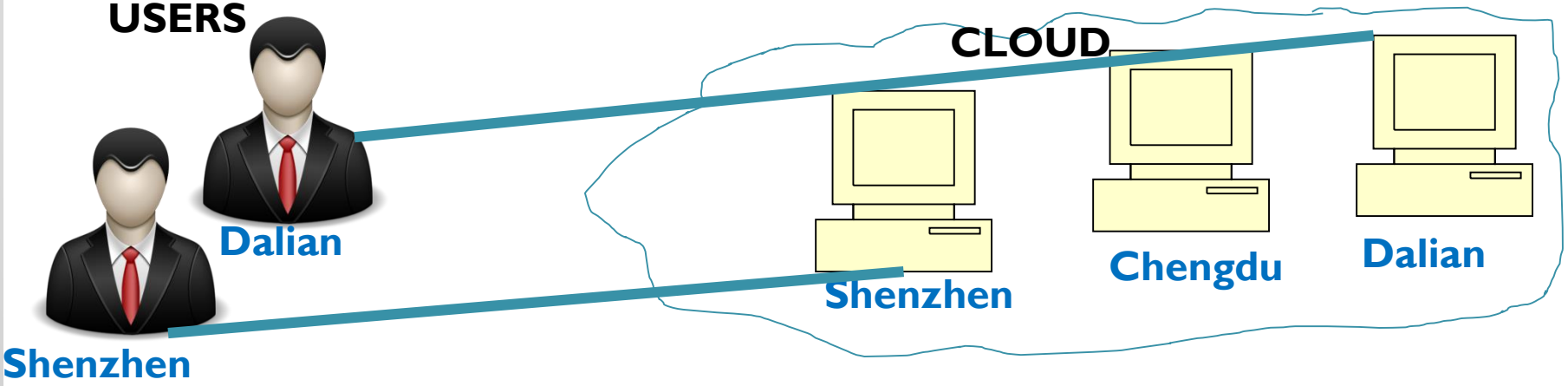
Data replication (数据复制)

Storing copies of resources in multiple geographic locations.

- Reduces the risk of losses,



- Reduces communication traffic (网络通信),



- Reduces energy consumption (能源消费)
  (by using computers in locations where electricity is cheap)

Data replication (数据复制)

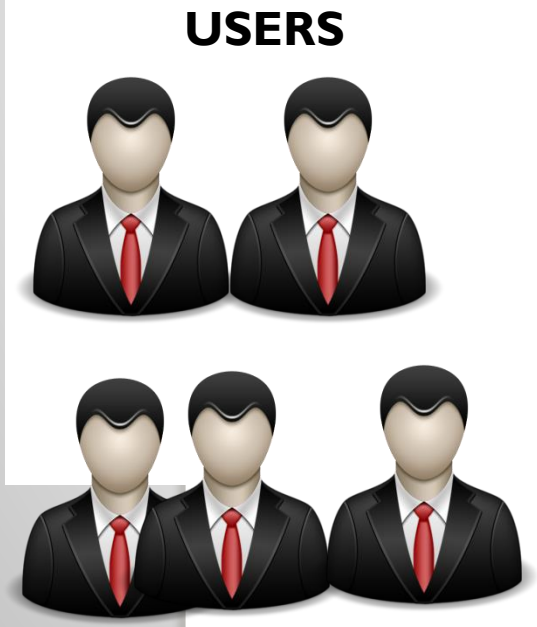
How many copies of the data?

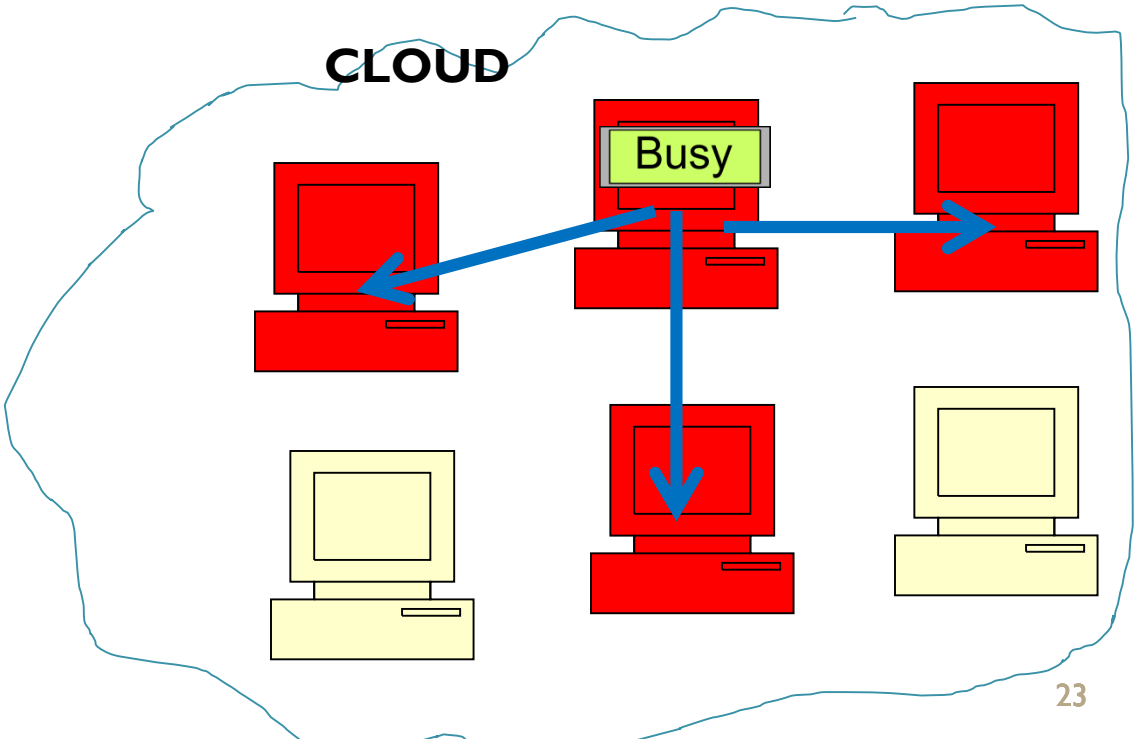
- **Many copies**: increases **users' satisfaction**
  quick response time (响应时间), maximum availability (可利用性), low cost, more reliability (可靠性)

- **Few copies:** increases the **cloud provider's satisfaction**
  financial benefits (元),
  smaller resource utilization (较小的资源利用), …

Load-balancing (负载平衡)

- In the cloud, **if a computer is too "busy", it can share its tasks** with other computers.
- This will reduce the **workload** (工作负荷) of each computer,
- This will reduce the **response time** (响应时间) for users.





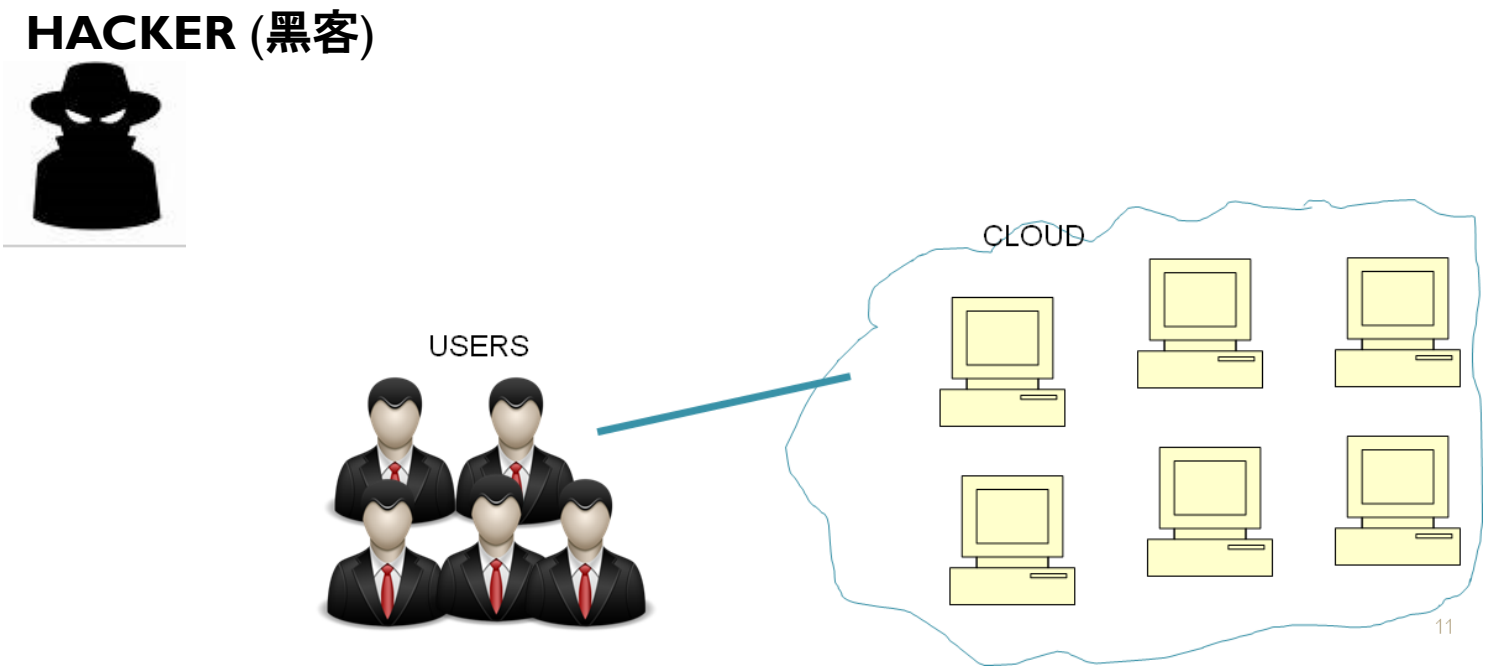
Security in the cloud

Security is the biggest challenge (最大的挑战) **for cloud computing!**

- A cloud provider must gain the **trust** of users (用户的信赖).
- Sometimes, a **public cloud** should **not be used** (medical data, military data).
- **Private clouds** may provide better security and improve performance for real-time applications (实时应用).

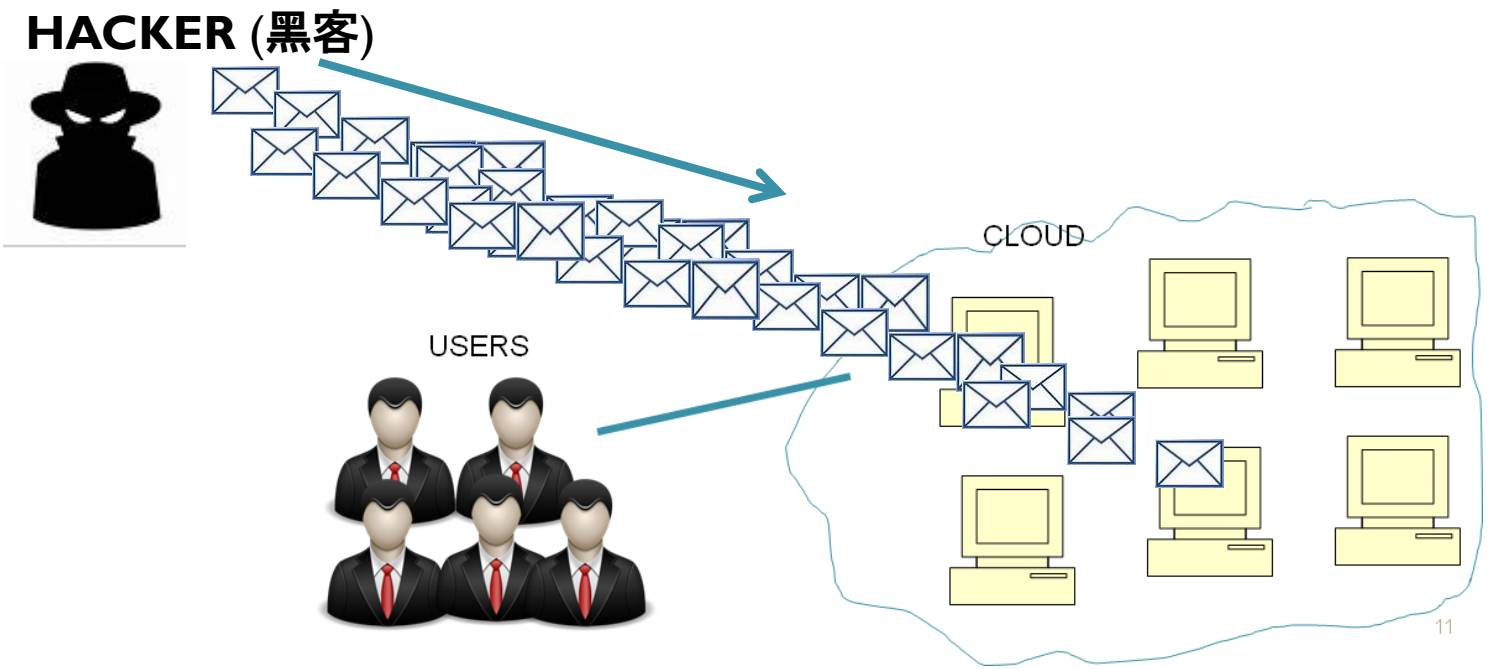
Attacks against the cloud

An important challenge for cloud computing is to **protect the cloud from hackers (黑客).**



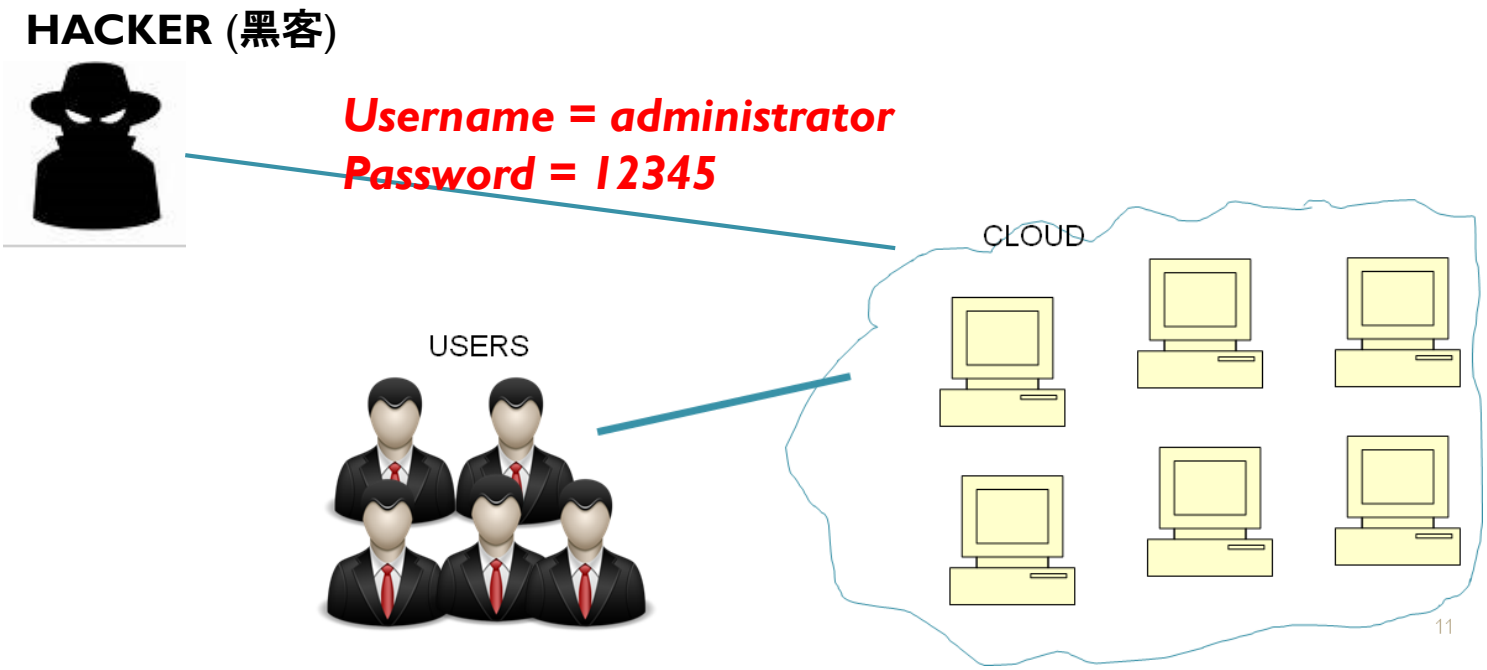
Denial-of-Service attack (DoS 拒绝服务)

- A **hacker sends millions of messages** to the cloud.
- The cloud spends all its computing power to process these fake messages.
- Hence, the cloud becomes unavailable to the real users.



Unauthorized access (未经授权的访问)

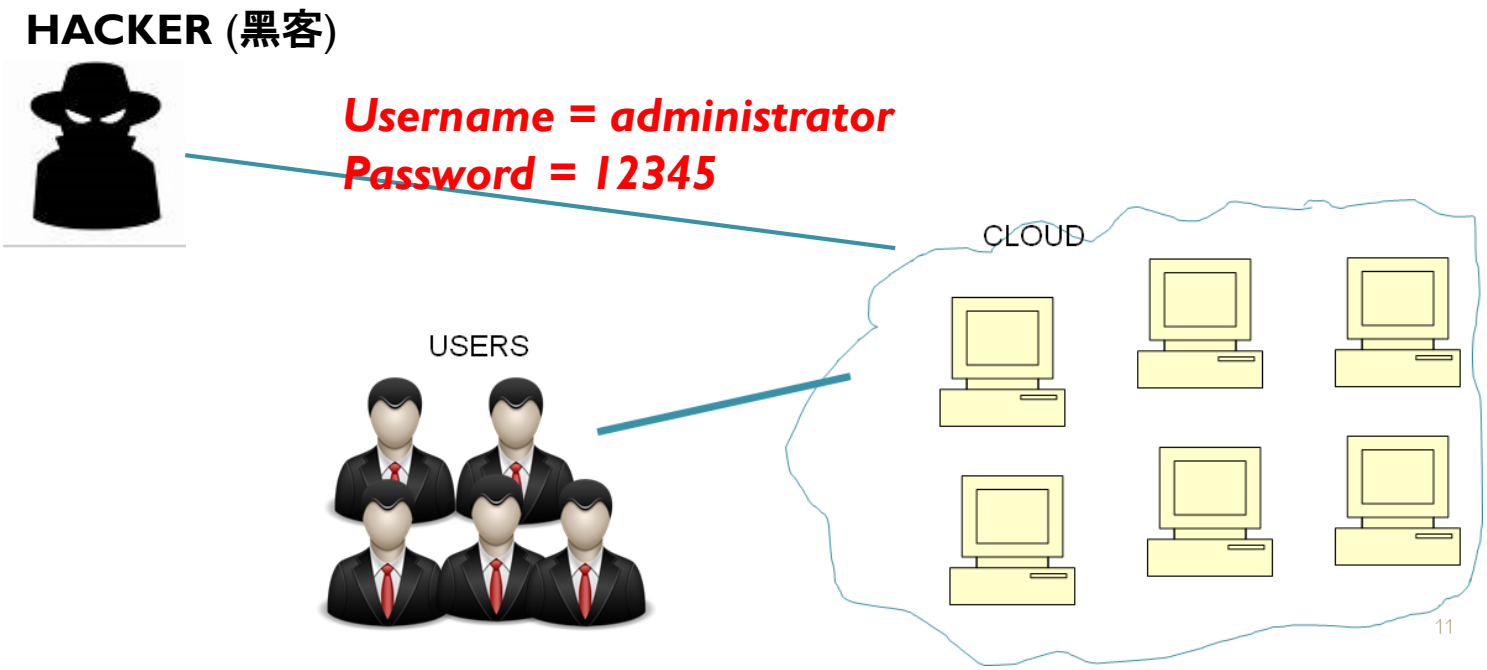
If a cloud is accessible from the Internet, a **hacker** may steal a **username (用户名)** and **password** (密码) to access and/or modify data.



Unauthorized access (未经授权的访问)

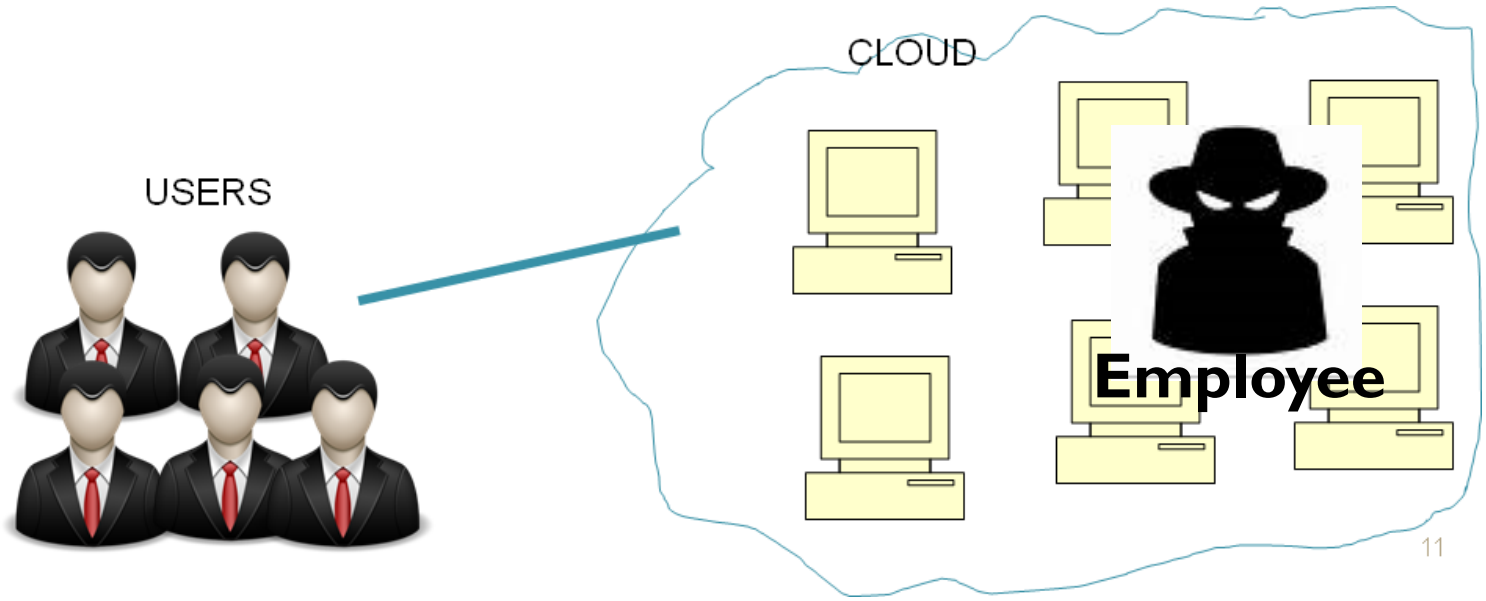
If a cloud is accessible from the Internet, a **hacker** may steal a **username (用户名)** and **password** (密码) to access and/or modify data.

Solutions: using strong passwords, changing the passwords often, etc.



Insider threat (内部威胁)

An employee or someone who has **physical access (物理访问) to the cloud** steals or modifies data.



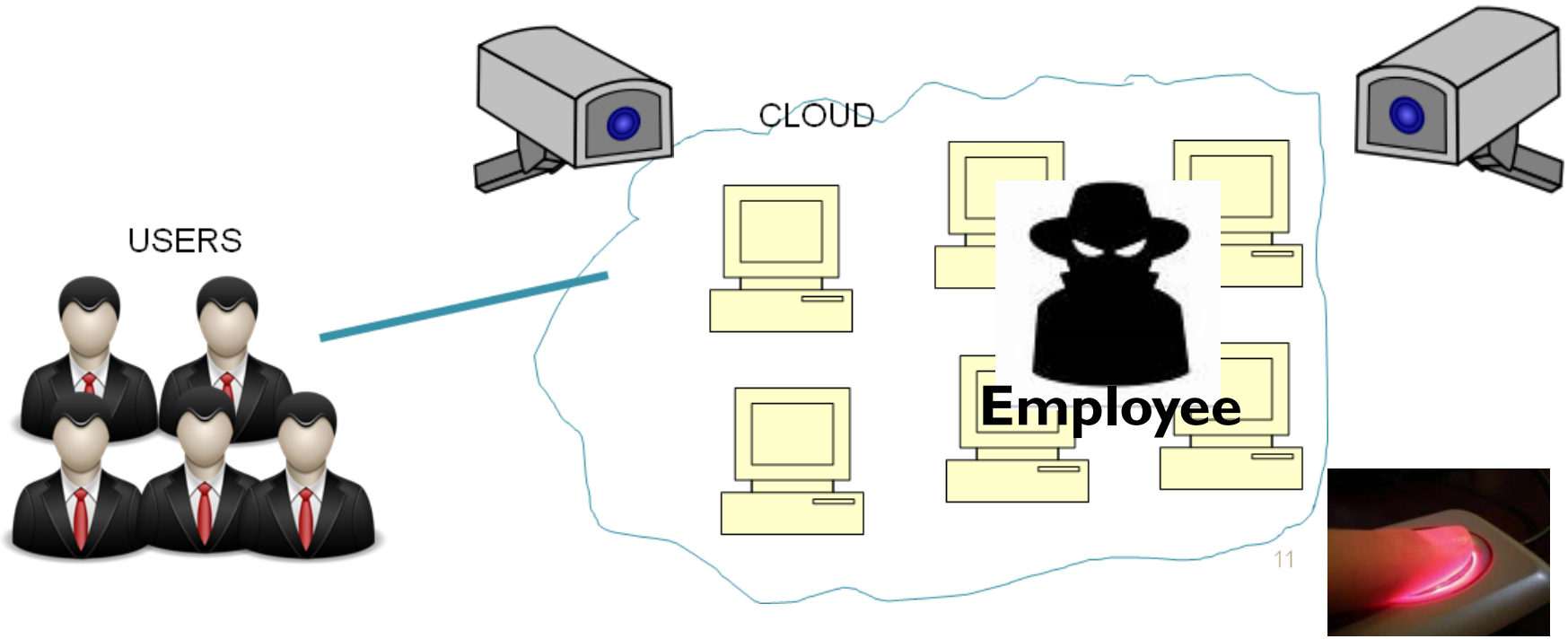
Insider threat (内部威胁)

Solution: increasing security

security cameras (安全摄像机),

locking doors (锁门),

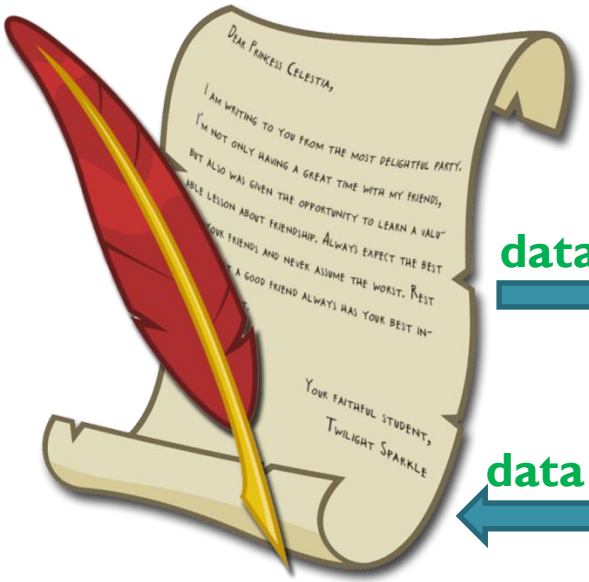
background check for employees (背景调查) …



Insider threat (内部威胁)

Solution: **data encryption 数据加密**

- Transforming the data to so that the data cannot be understood by a thief (小偷).
- However, data need to be decrypted by cloud applications. So a risk remains.



数据加密
data encryption

数据解密
data decryption

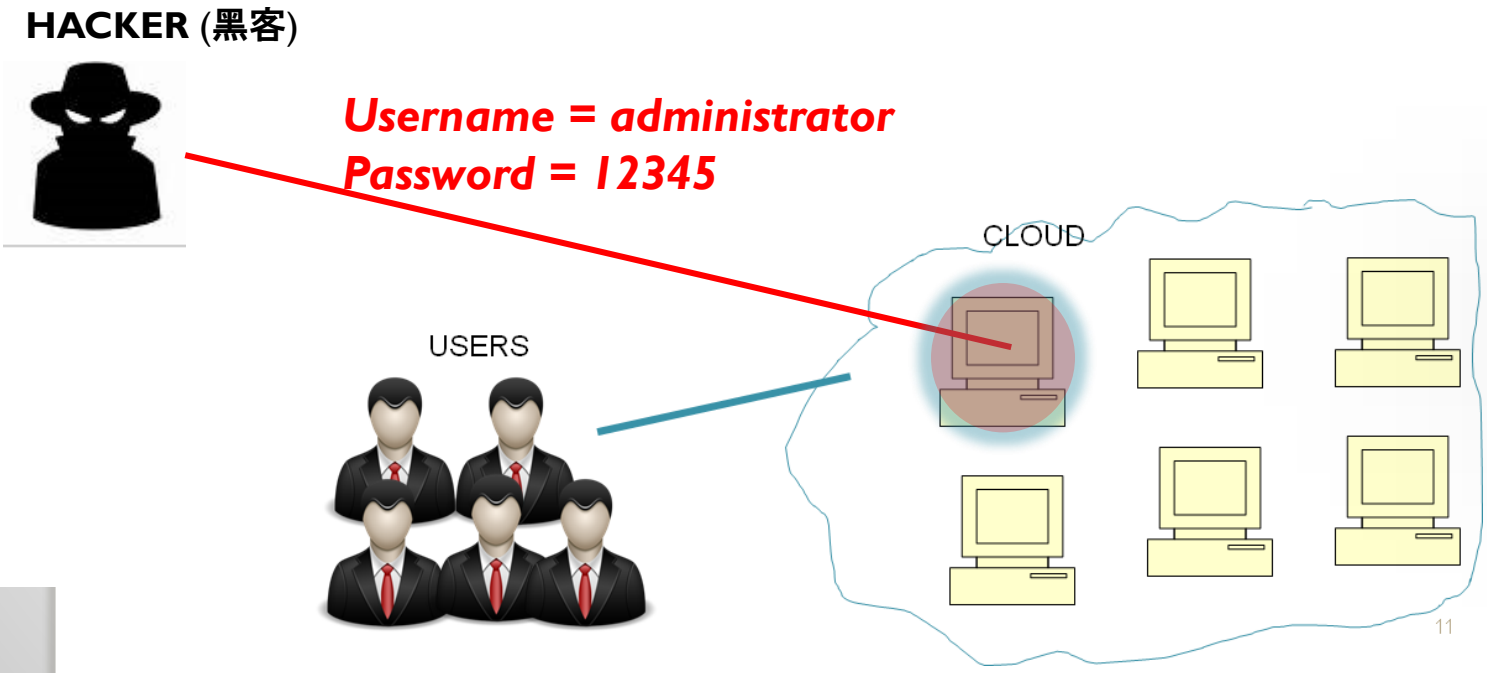
From shutterstock

Hackers using the cloud to attack

This attack targets the "**infrastructure-as-service**" model.

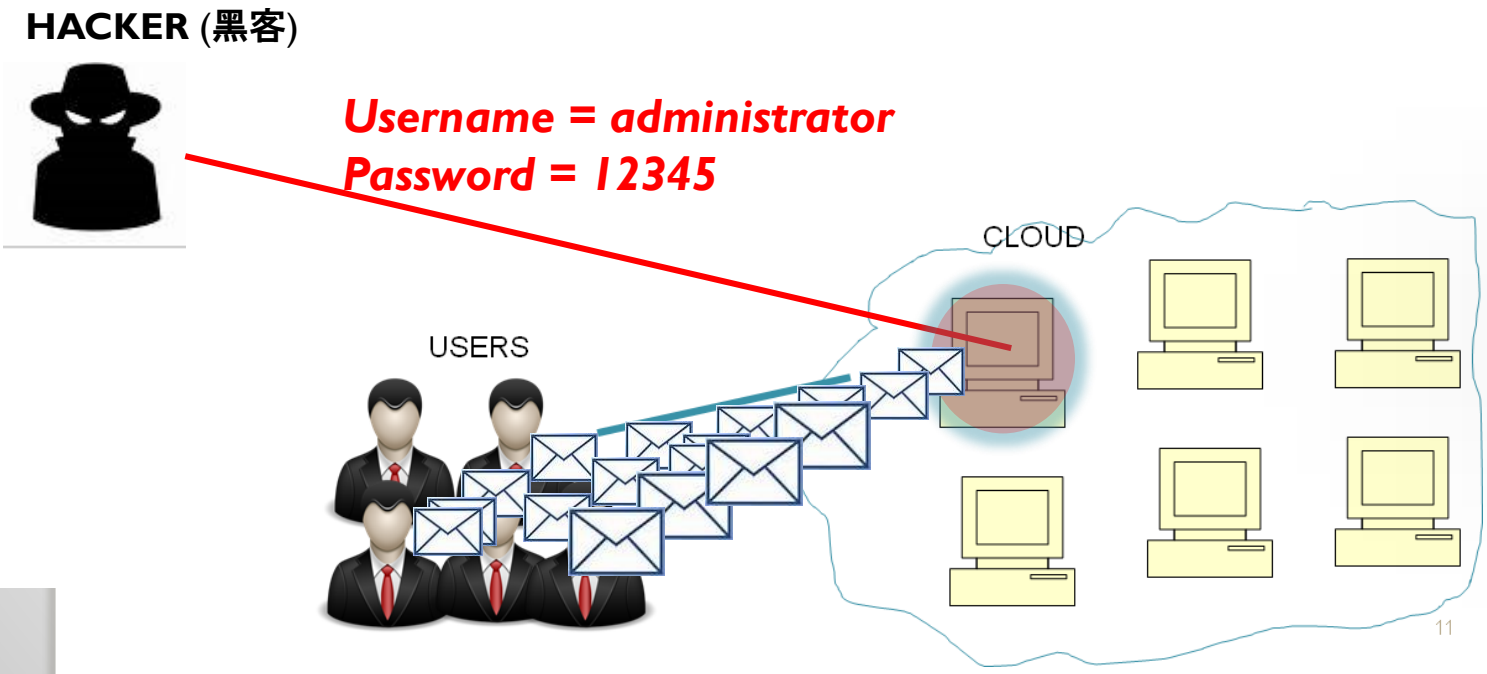
For example:

1) a **hacker** gain access to the cloud by using the **username** and **password** of the administrator.



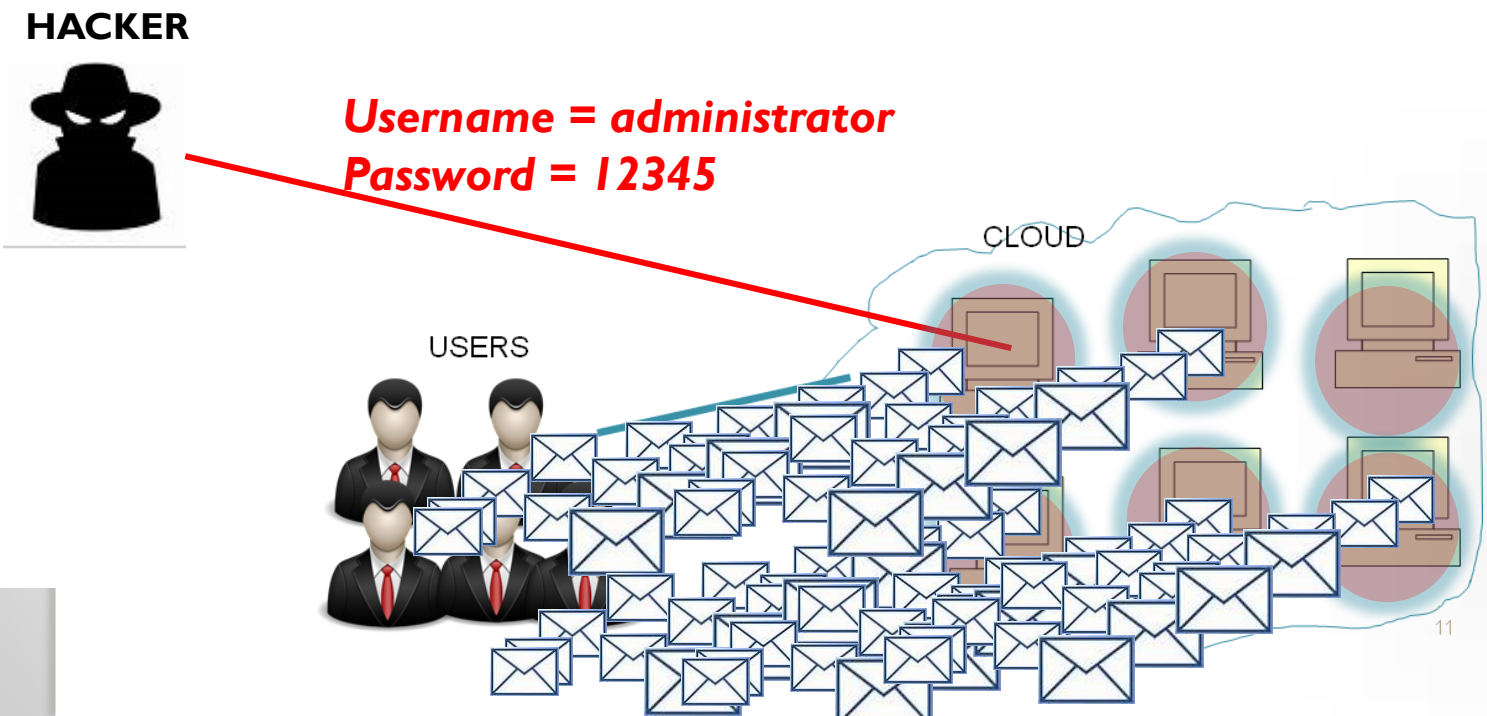
Hackers using the cloud to attack

2) The **hacker** uses computers from the cloud to send millions of messages to users or other websites to take them down.



Hackers using the cloud to attack

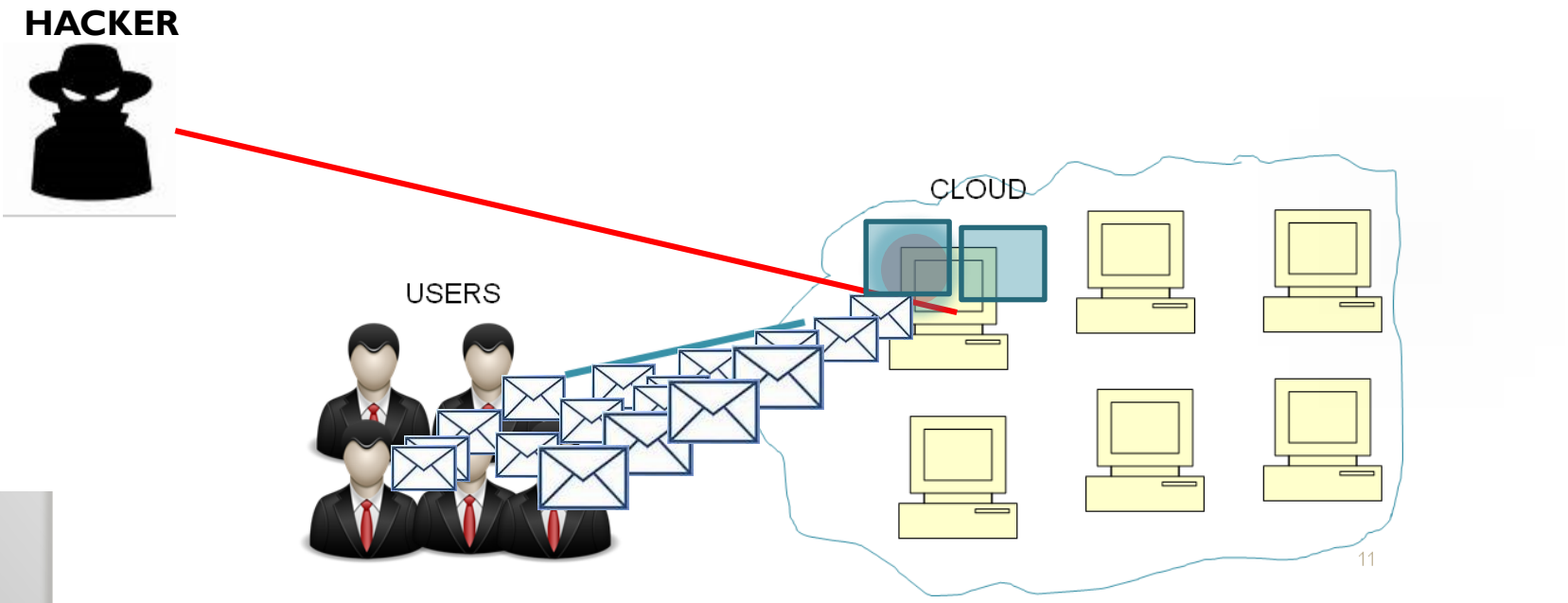
A hacker can potentially use **<u>all</u> the resources** provided by the cloud to perform big attacks.



Hackers using the cloud to attack

Solutions:

- advanced security measures.
- **virtualization (虚拟化)**: each user cannot use the whole cloud. A user can only run applications in a **virtual machine** (虚拟机 - a kind of virtual computer) in the cloud.



What cloud providers must also do?

- **Capacity allocation** (容量分配)
- **Load balancing** (负载平衡)
- **Energy optimization** (能量优化): try to reduce energy consumption
- Provide **Quality-of-Service** (服务质量) guarantees
- …

Challenge:

- It requires to know the **current state** of the **cloud**.
- Knowing the **state** of the cloud is difficult because there are many computers and their states change.

What cloud providers must also do?

A solution:

- **Self-management (自我管理)** and **self-organization (自组织)**: the cloud **automatically manages** and **organizes itself**.
- Good.
- But it may become difficult to find the causes of security breaches (安全漏洞) or other problems. The cloud is a dynamic system (动态系统) with perhaps thousands of computers.

Interoperability (互操作性)

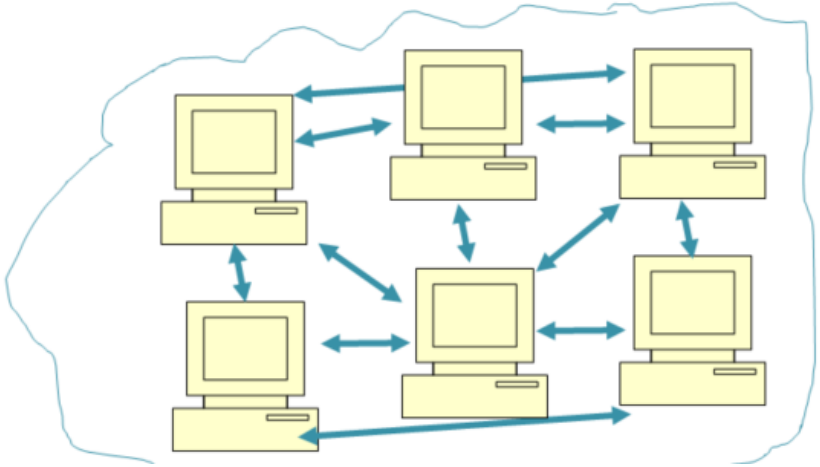
- It is desirable to avoid "Vendor lock-in" (厂商锁定)(that the user can easily change cloud provider)
- Not always easy!
- **Solution**: develop **standards** (标准) for **cloud computing**.

CHAPTER 2 – PARALLEL SYSTEMS (并联系统) AND DISTRIBUTED SYSTEMS (分布式系统)

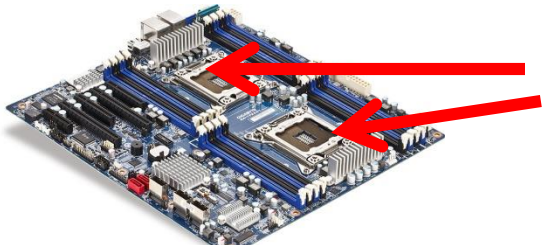
Parallel system (并联系统)

A set of components (**e.g.** computers, processors, threads) **that perform tasks in parallel** (在并行).

- **e.g.** a **cloud** communicating via a network.



e.g. a **multi-processor computer (多处理器电脑)** where processors communicate via shared memory (共用存储器).

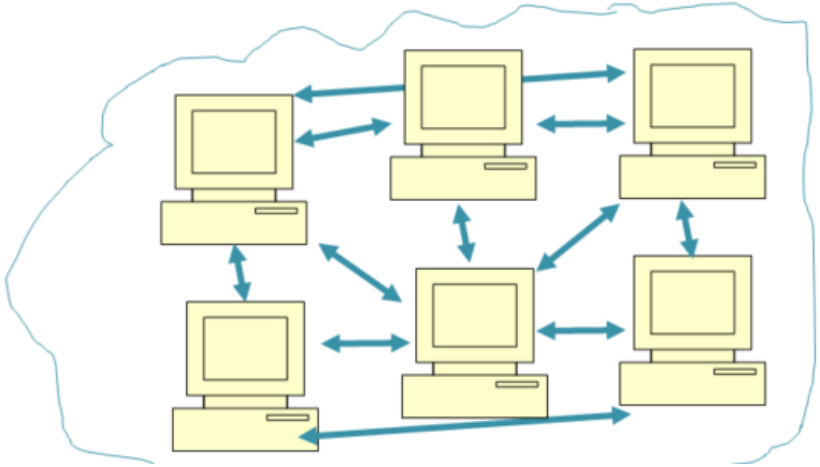


2 processors

Distributed system (分布式系统)

A set of computers that **communicate through a network** by **sending/receiving messages**.

- **e.g.** a **cloud** communicating via a network.



e.g. a **messaging applications** like **Wechat** (微信)

…

Computers communicate and exchange data to work together

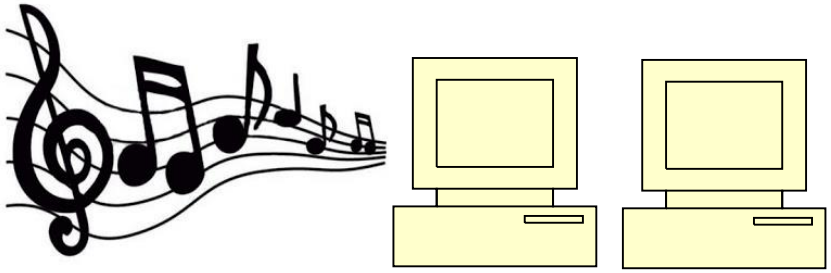
Which tasks can be done in parallel (在并行)?

- Counting the number of students in the classroom

- Grading the assignments (作业) of students
- …

Which tasks cannot be done in parallel?

- Playing a song. Using more computers will not play the song faster.



- "Pregnancy (怀孕). More women will not reduce the length of pregnancy".

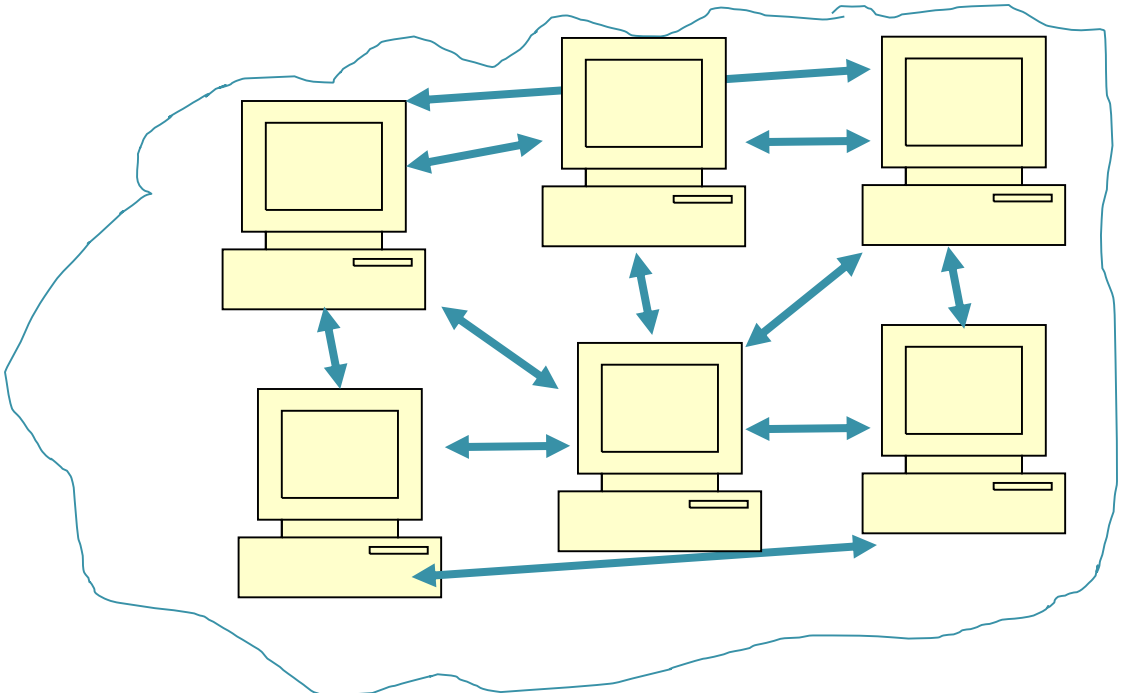
- Cooking （做饭): washing vegetables, cutting a green onion, etc. Some steps cannot be done in parallel.

Is it faster to do a task in parallel?

Sometimes. Why?

Because time must be spent for coordination (协调) and communication.

Using more computers is not always better!



Challenges of **parallel computing**

- Parallel system
- There are **several challenges** related to coordinating a set of computers that work in parallel.
- An **overview** of the **main challenges** →

Failures (故障) may occur

After a failure, the state of a parallel system may become incorrect.

Example: consider that a person **A** wants to transfer **100** 元 to a person **B**

A 600元

B 200元

Example

Transfer_money(A, B, amount){

read(*A*);

A := A – amount;

write(*A*);

read(*B*);

B := B + amount;

write(*B*);

A 600元

B 200元

Example

Transfer_money(A, B, amount){

read(*A*);

A := A – amount;

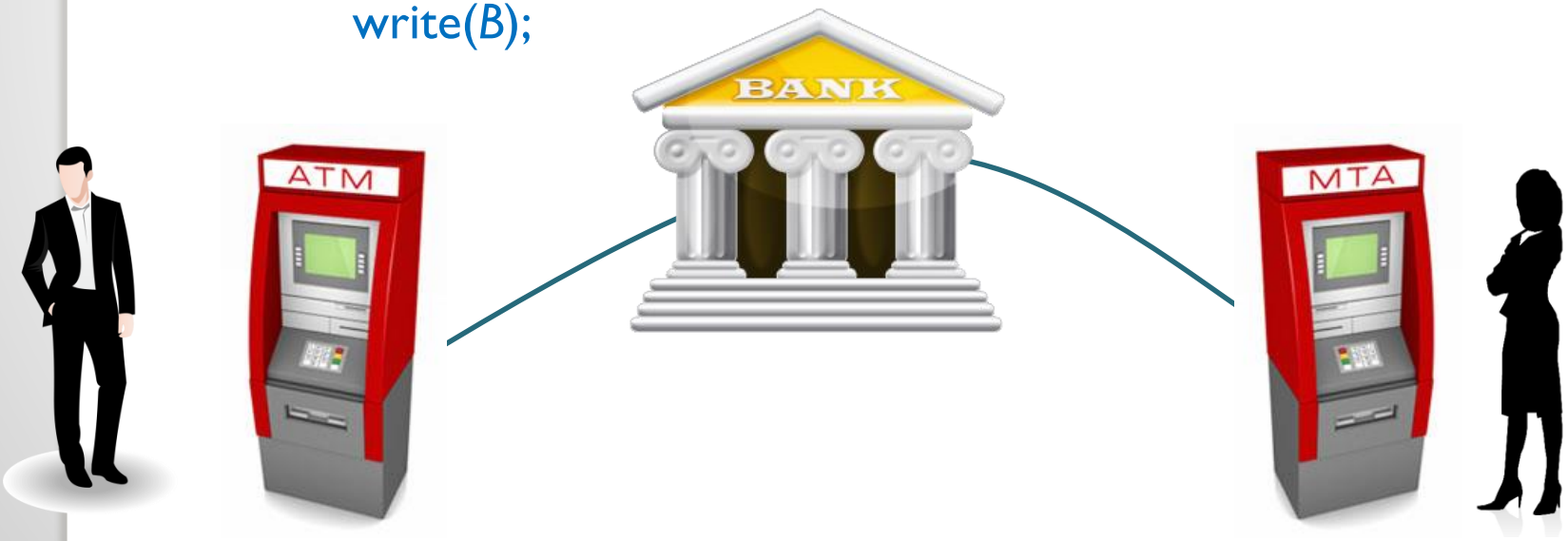
write(*A*);

read(*B*);

B := B + amount;

write(*B*);

100 元 is withdrawn from the account of **A**



A 500元

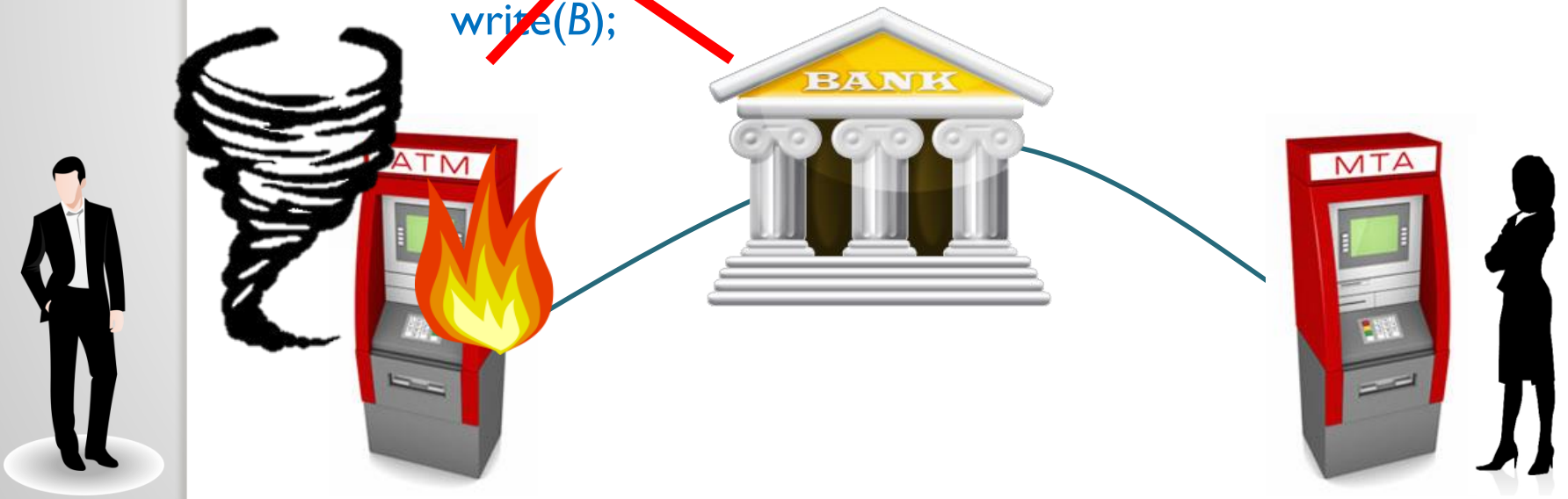
B 200元

Example

```
Transfer_money(A, B, amount){
        read(A);
        A := A – amount;
        write(A);
        read(B);
        B := B + amount;
        write(B);
```

A failure occurs before the money is transferred to person B.

The money is lost!



A 500元

B 200元

A solution

- Each computer should keep the history of the operations that it performs on data.
  (a **log** -访问日志)
- If there is a failure, a computer can **undo** the operations to return to a correct state.

Concurrent accesses to data (并发访问)

Some resources must not be accessed by two computers at the same time. Otherwise, problems may occur!

Example: Person A wants to transfer 50 元 to person B
Person B wants to transfer 500 元 to person A

A 600元

B 1000元

Example

Transfer_money(A, B, 50元)

read(A);

A := A – **50元** ;

write(A);

read(B);

B := B + **50元** ;

write(*B*);

Transfer_money(B, A, 500元)

read(B);

B := B – **500元**;

write(B);

read(A);

A := A + **500元**;

write(*A*);

A 600元

B 1000元

Example

Transfer_money(A, B, 50元)

read(A);

A := A – **50元** ; **A = 550元**

write(A);

read(B);

B := B + **50元** ;

write(*B*);

Transfer_money(B, A, 500元)

read(B);

B := B – **500元**;

write(B);

read(A);

A := A + **500元**;

write(*A*);

A 600元

B 1000元

Example

Transfer_money(A, B, 50元)

read(A);

A := A – **50元** ; **A = 550元**

write(A);

read(B);

B := B + **50元** ;

write(*B*);

Transfer_money(B, A, 500元)

read(B);

B := B – **500元**;

B= 500元

write(B);

read(A);

A := A + **500元**;

write(*A*);

A 600元

B 500元

Example

Transfer_money(A, B, 50元)

read(A);

A := A – **50元** ; **A = 550元**

write(A);

read(B);

B := B + **50元** ;

write(*B*);

Transfer_money(B, A, 500元)

read(B);

B := B – **500元**;

write(B);

read(A);

A := A + **500元**;

write(*A*);

A 1100元

B 500元

Example

Transfer_money(A, B, 50元)

read(A);

A := A – **50元** ;

write(A);

read(B);

B := B + **50元** ;

write(*B*);

Transfer_money(B, A, 500元)

read(B);

B := B – **500元**;

write(B);

read(A);

A := A + **500元**;

write(*A*);

A 550元 **← Some money is lost!!** **B** 500元

1600 → 1100

Example

Transfer_money(A, B, 50元)

read(A);

A := A – **50元** ;

write(A);

read(B);

B := B + **50元** ;

write(*B*);

Transfer_money(B, A, 500元)

read(B);

B := B – **500元**;

write(B);

read(A);

A := A + **500元**;

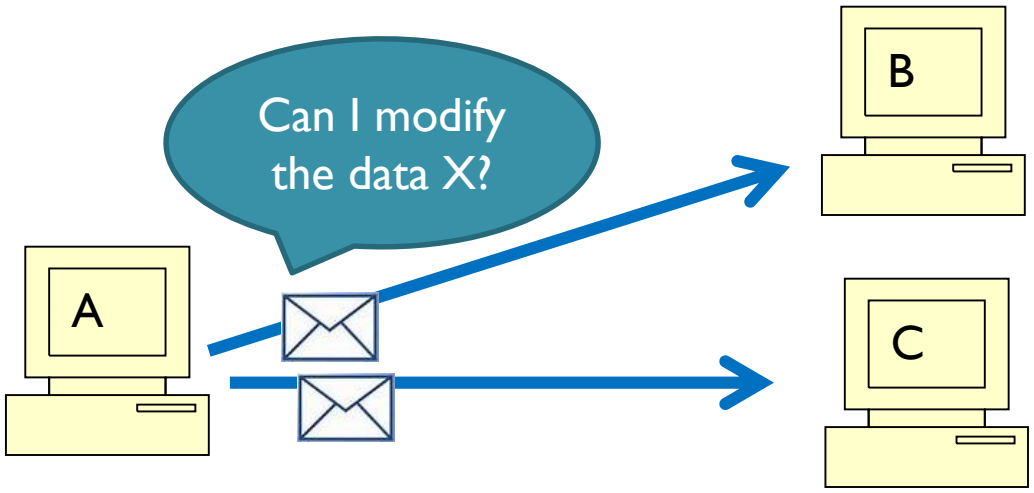
write(*A*);

A 550元

B 550元

A solution

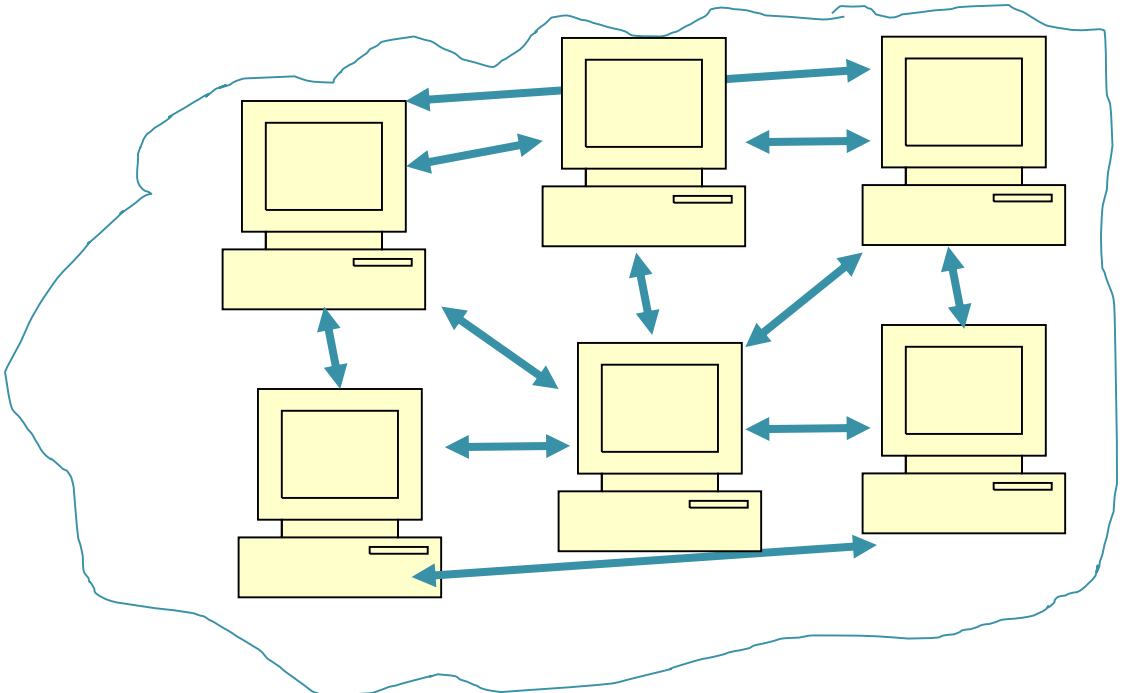
- A **computer** should always **request the permission to modify data** that is **shared** with other computers.
- This ensure that no other computer will modify the data at the same time.
- This is called **mutual exclusion** (互斥).



The **computer A** requests to modify data **X**

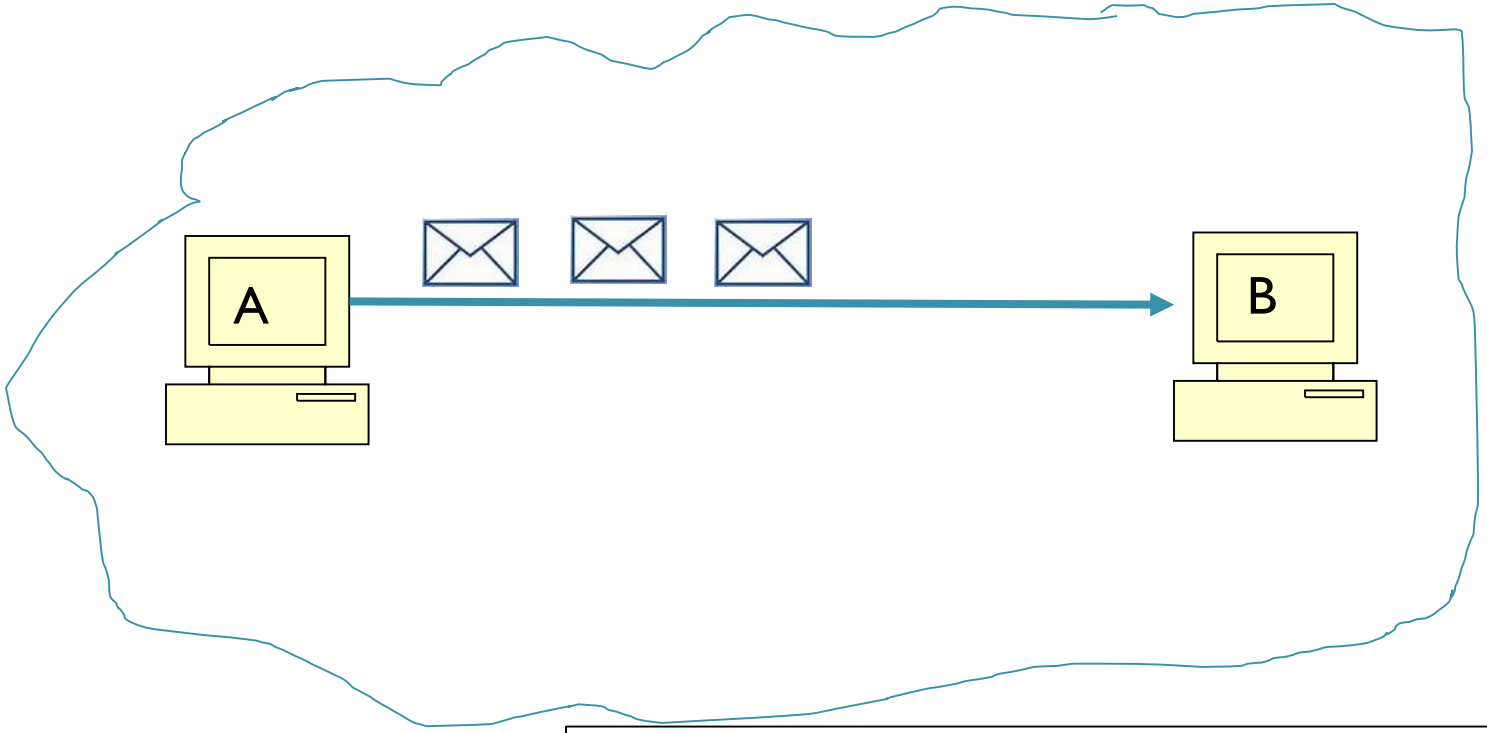
Challenges of **distributed computing**

- Distributed system
- An **overview** of the **main challenges** →



How computers communicate?

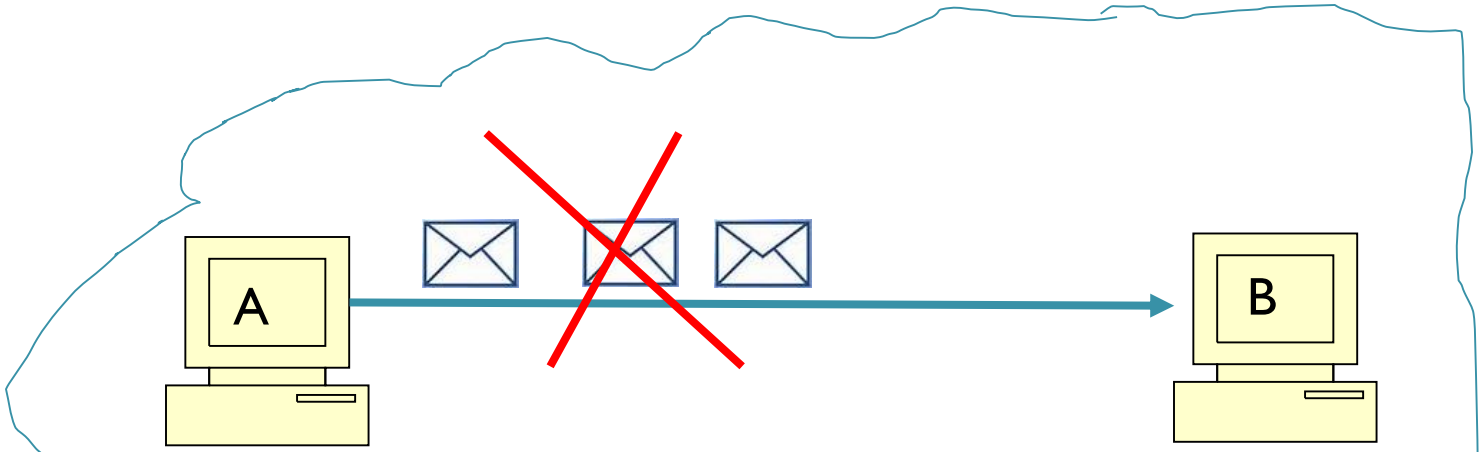
- **Computers send messages** to other computers through a network such as the Internet.
- **Protocol (通信协议)**: a set of rules that some computers follow to communicate with each other



Several problems may occur when sending messages over a network →

Problem 1: Message loss (消息丢失)

Some messages may be lost due to an unreliable network, or other problems occurring on a network.

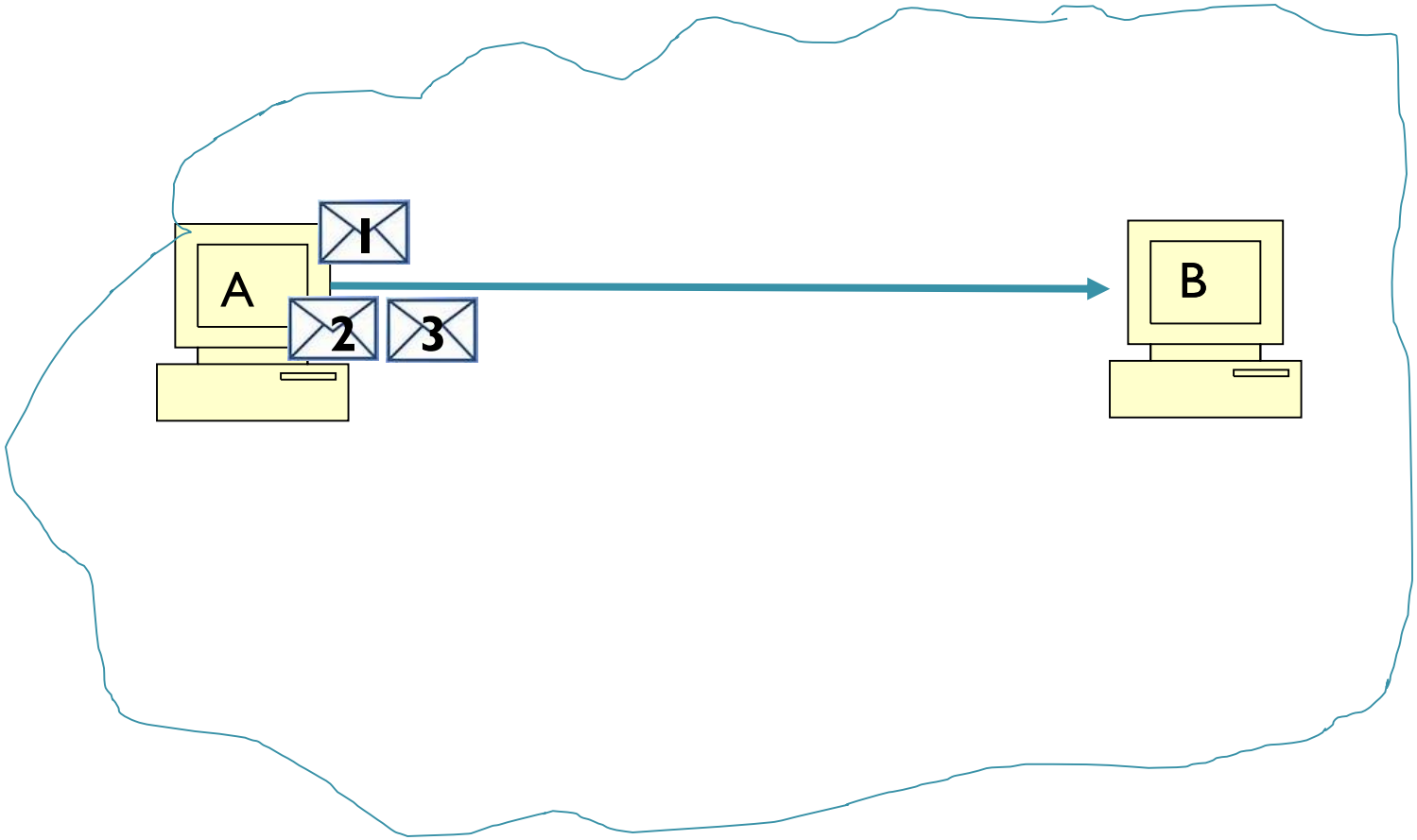


Computer A send three messages to **computer B**.
A message is lost.
The **computer B** does not receive the **second message**.

How to avoid this problem? →

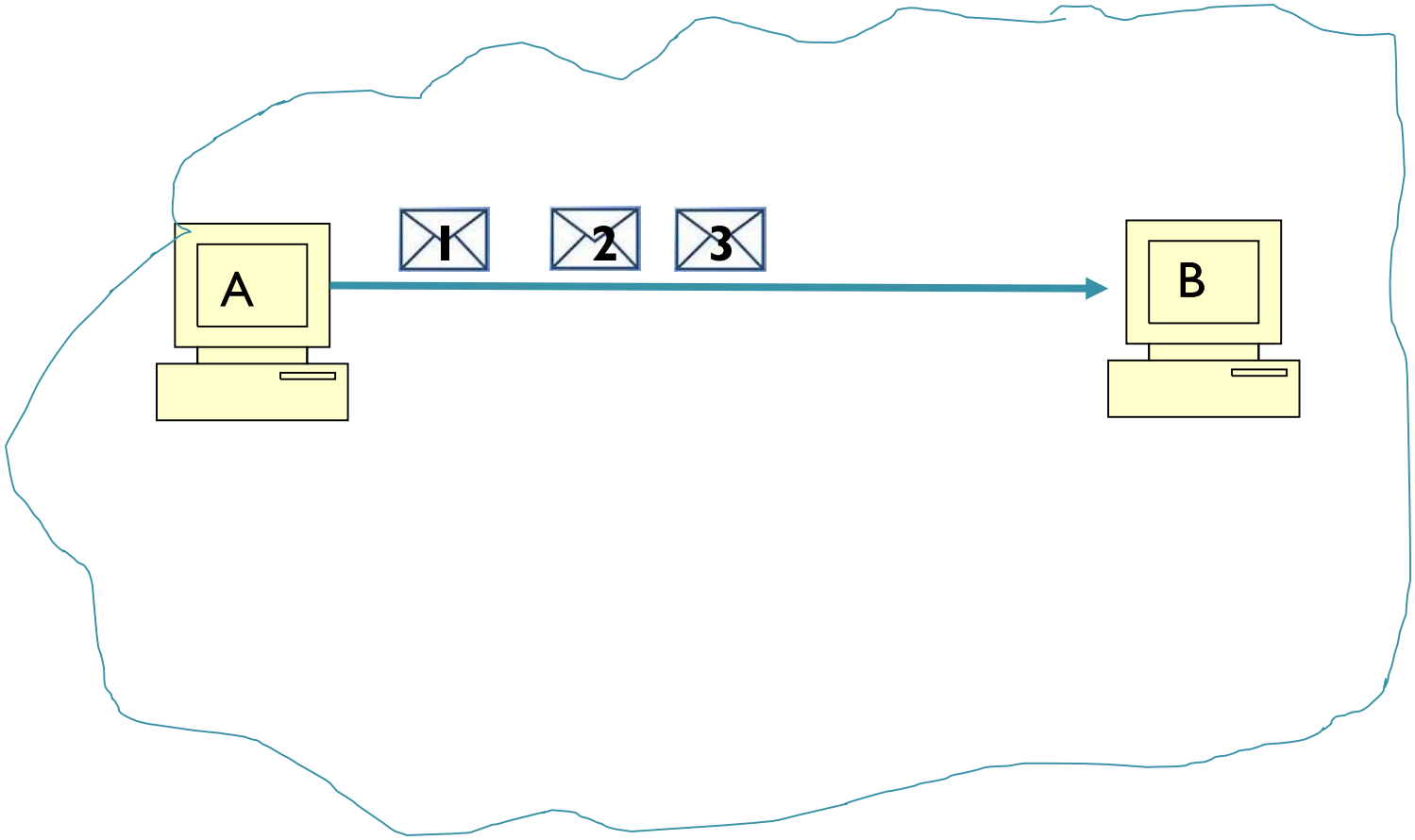
A solution

Messages are numbered using
sequence numbers (序列号) 1,2,3,…. before being sent.



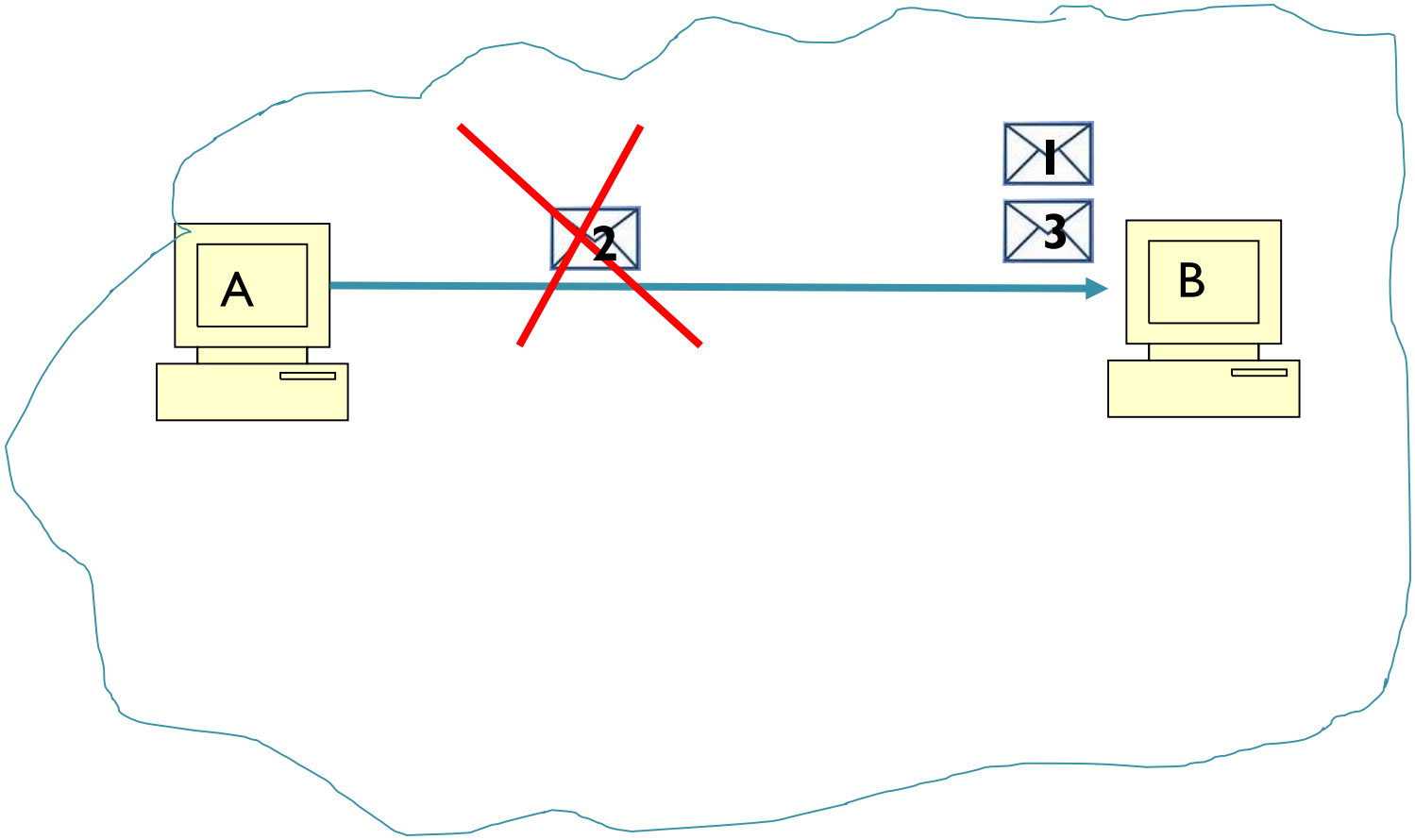
A solution

Computer A sends the three messages.



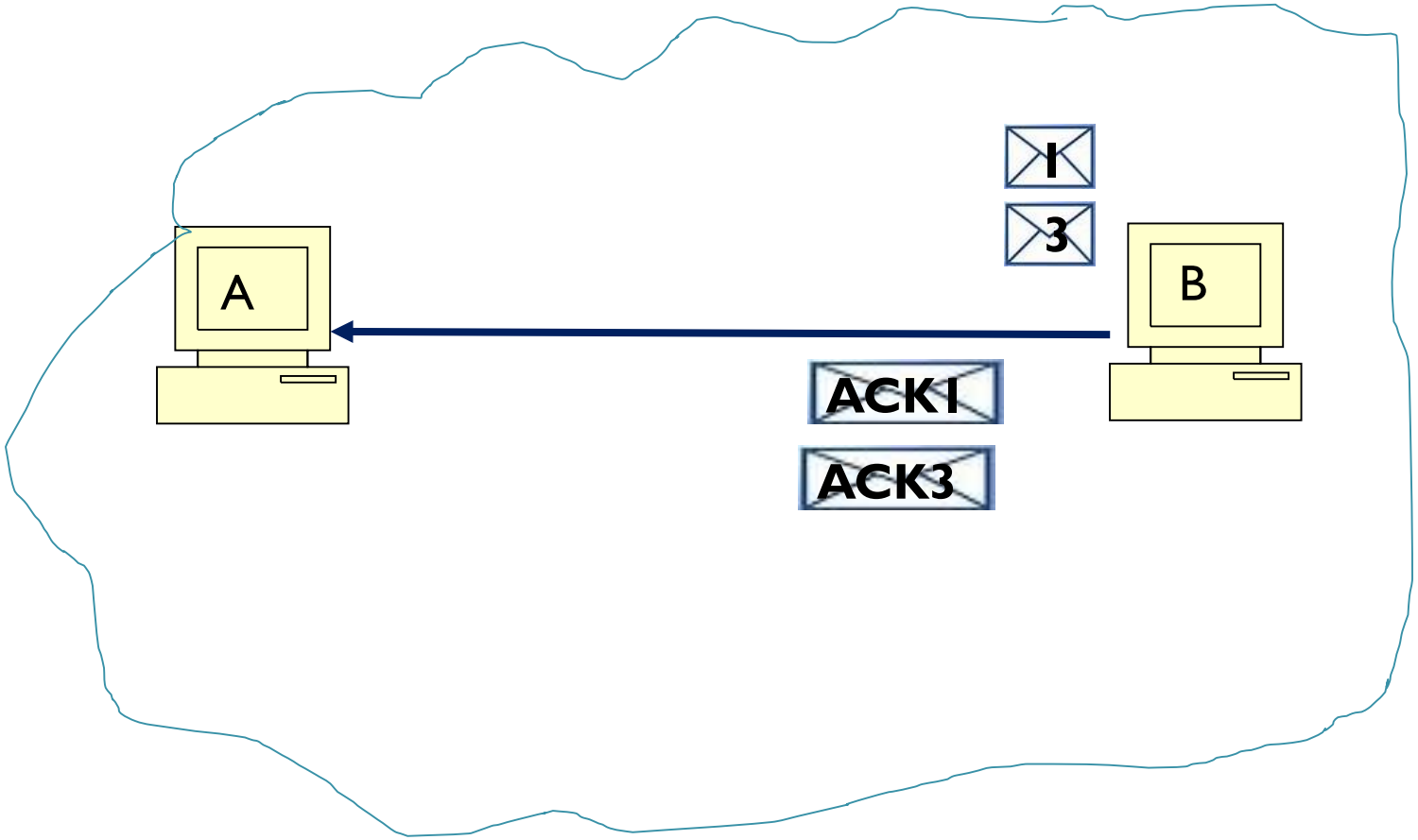
A solution

Message 2 is lost.
Computer B only receives **messages 1** and **3**



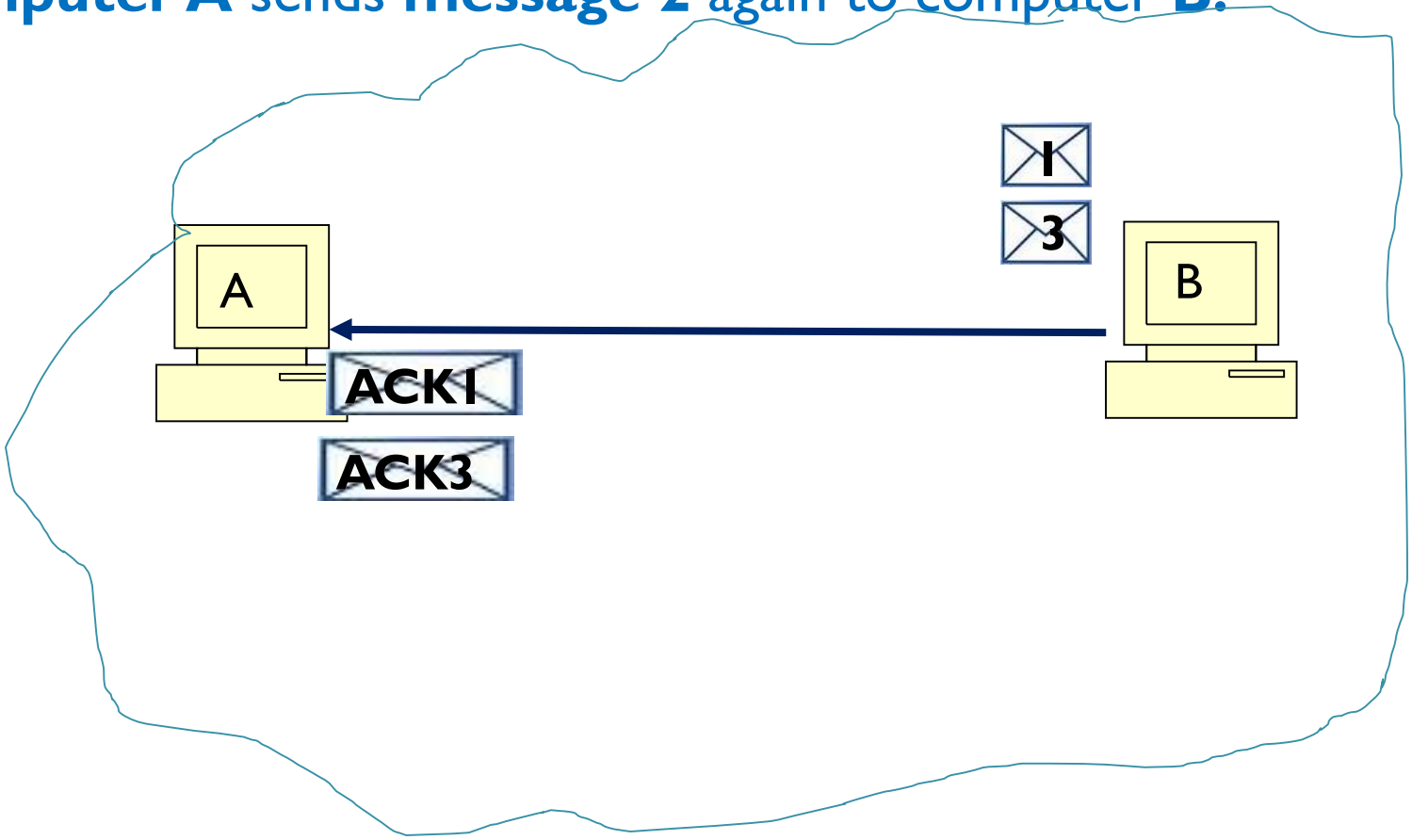
A solution

Computer B sends acknowledgements (确认) to inform **computer A** that it has received **messages 1** and **3**



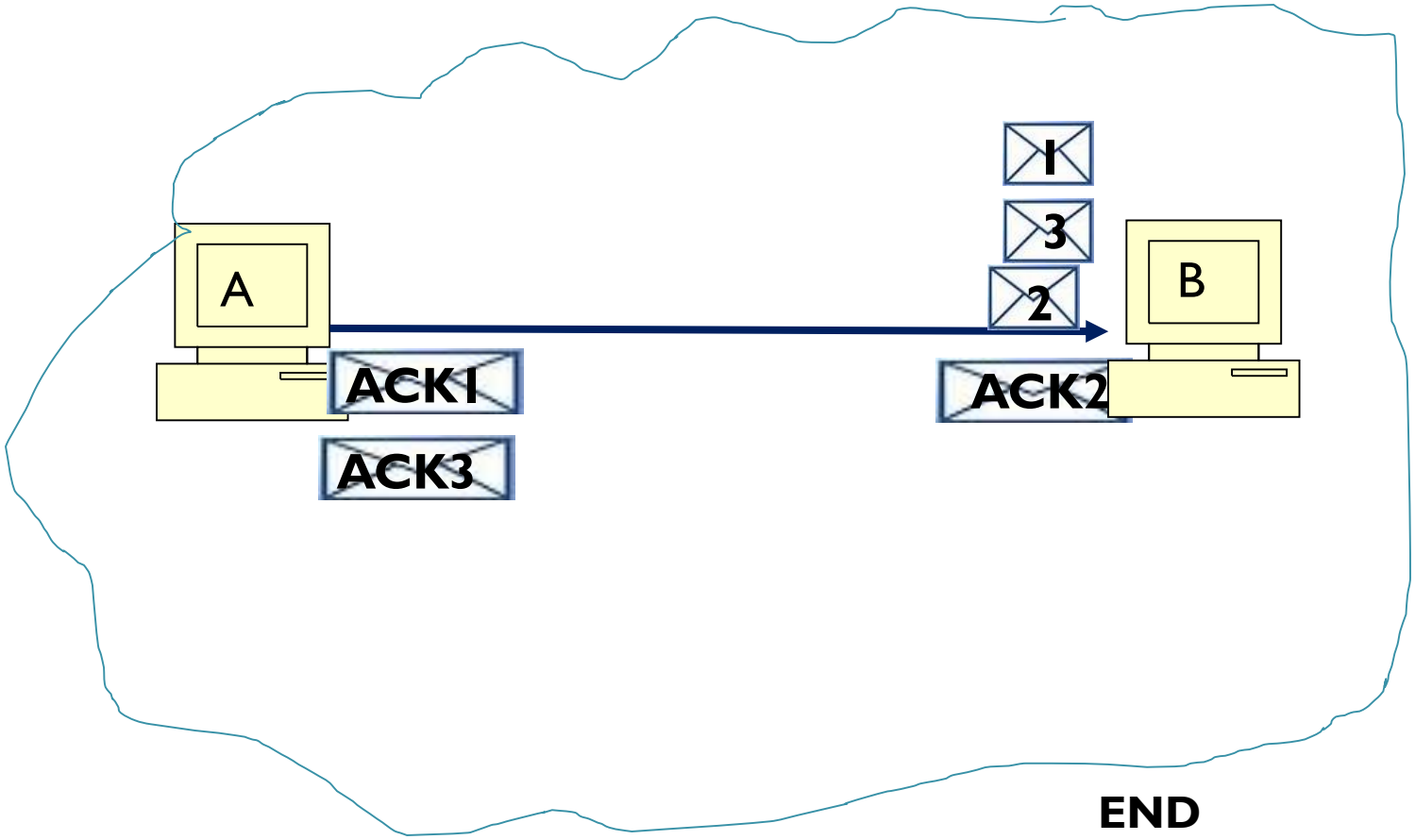
A solution

Computer A receives acknowledgements for **message 1** and **3,** but not for **message 2.**
Thus, **computer A** sends **message 2** again to computer **B.**



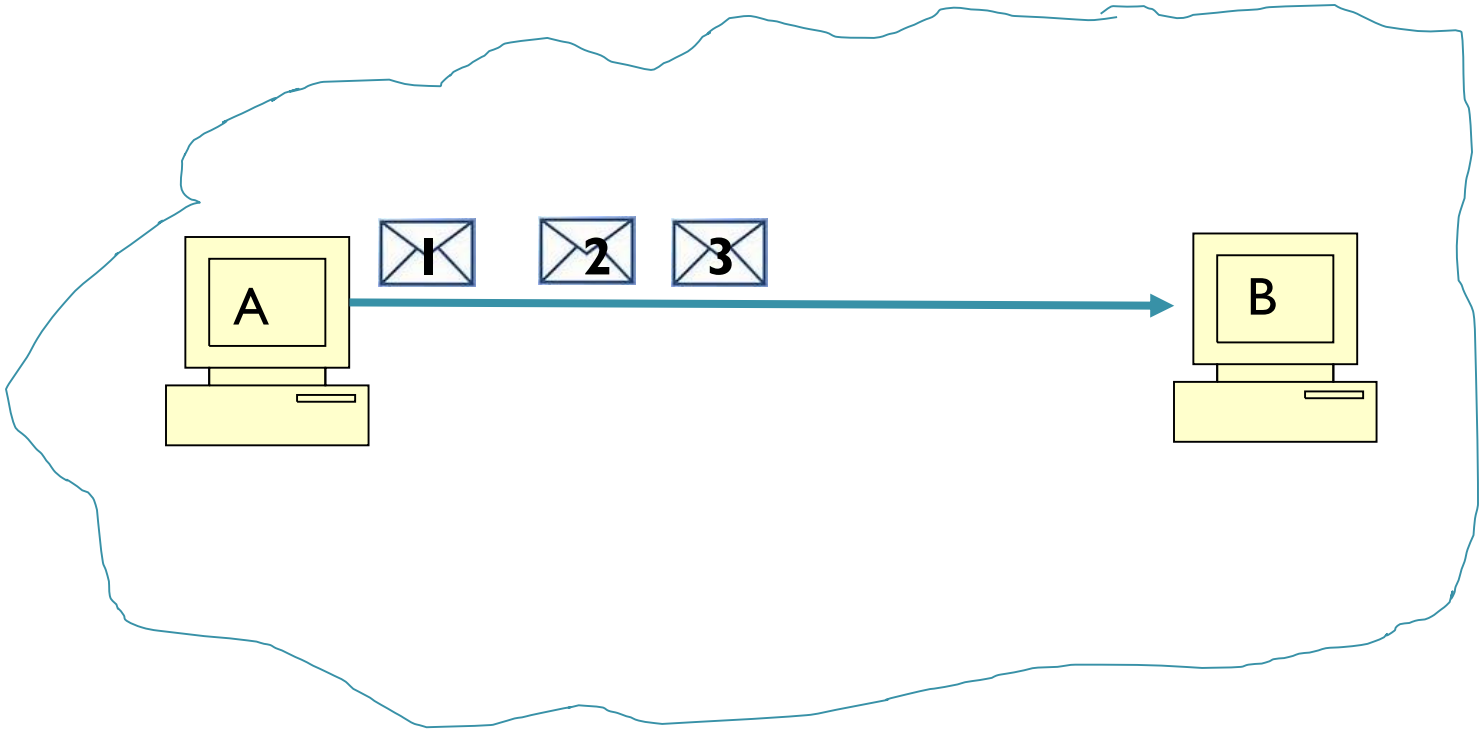
A solution

Computer B sends an acknowledgement to tell **Computer A** that it has received **message 2**.



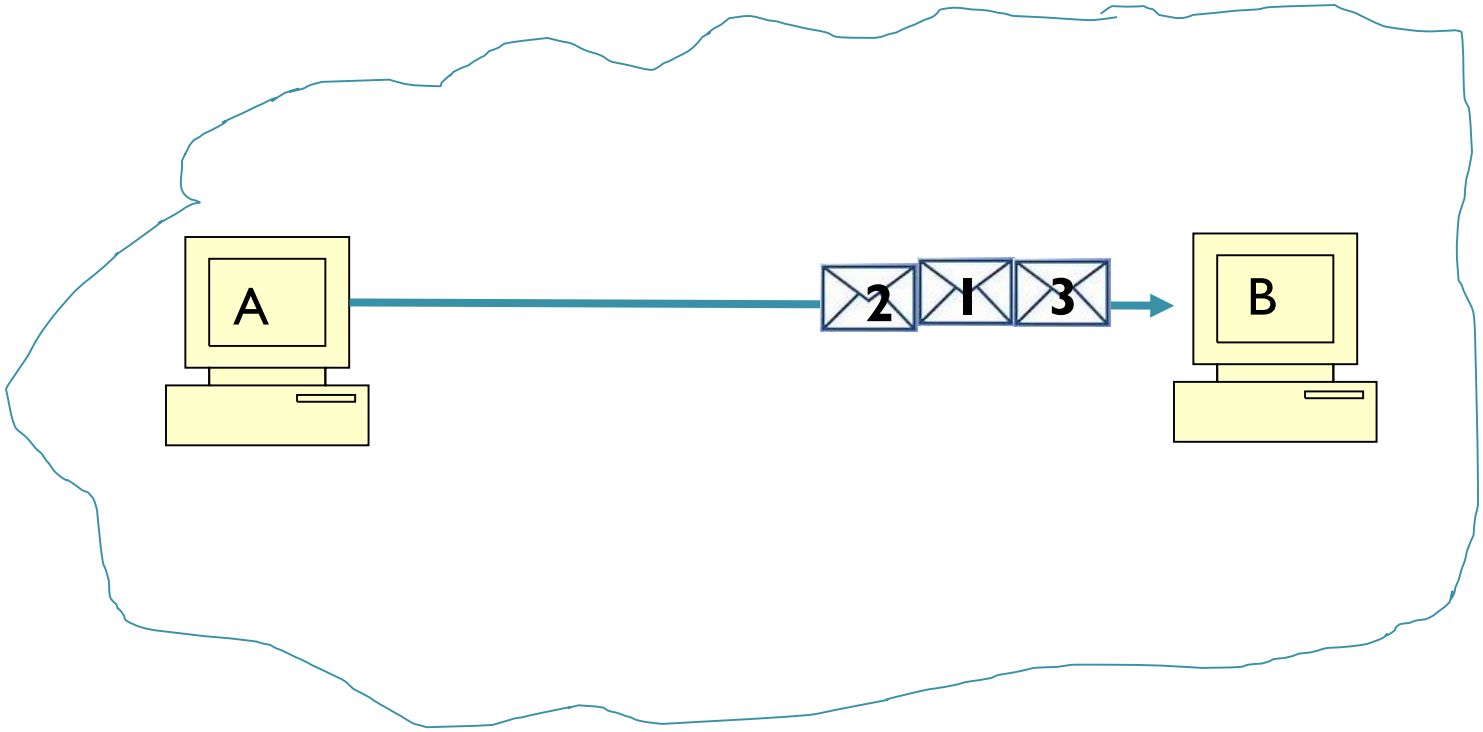
Problem 2: Messages may arrive in an incorrect order

On the internet, it is not guaranteed that messages are received in the order that are sent.



Problem 2: Messages may arrive in an incorrect order

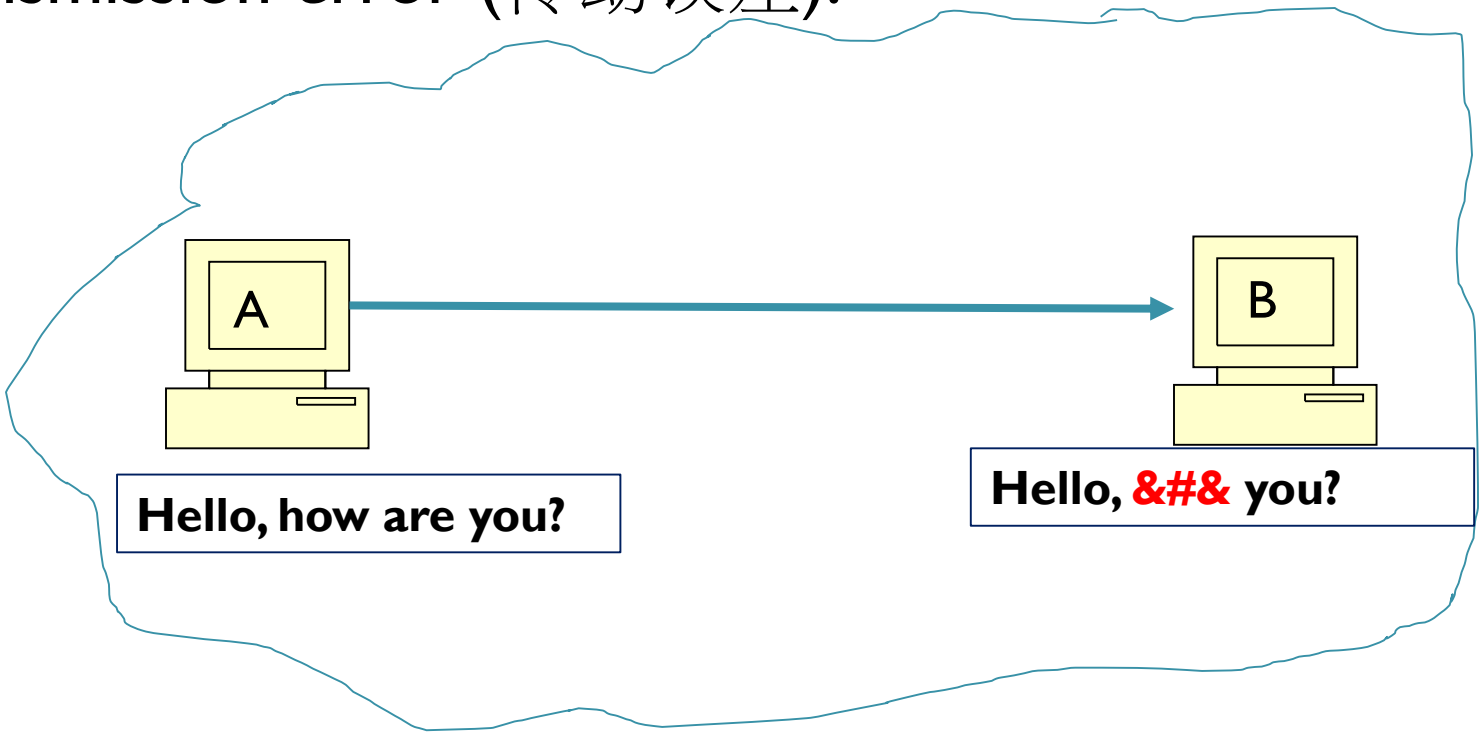
On the Internet, it is not guaranteed that messages are received in the order that are sent.



Solution: The **sequence numbers** (序列号) on messages can be used to **reorder the messages**.

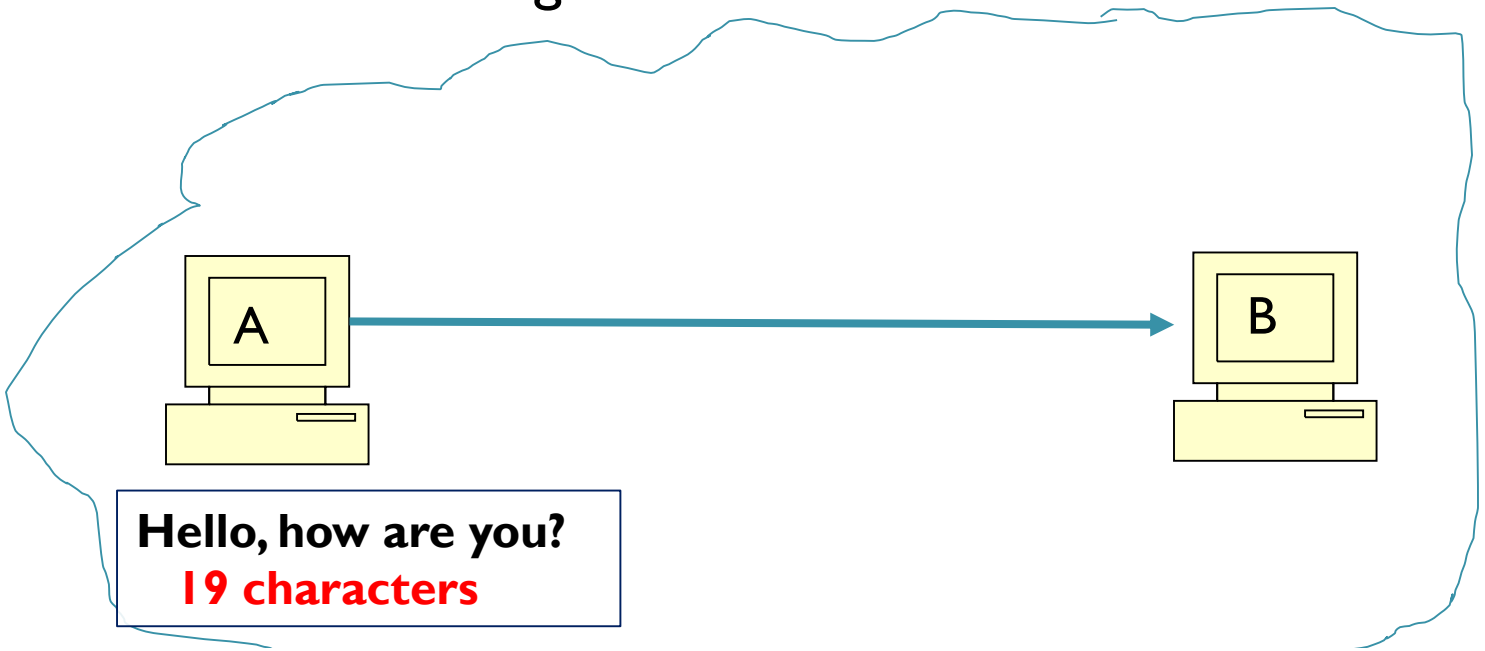
Problem 3: Corrupted messages (损坏的消息)

A message sent on a network may become corrupted due to a transmission error (传动误差).



Solution: use an **error detection code** (错误检测代码)

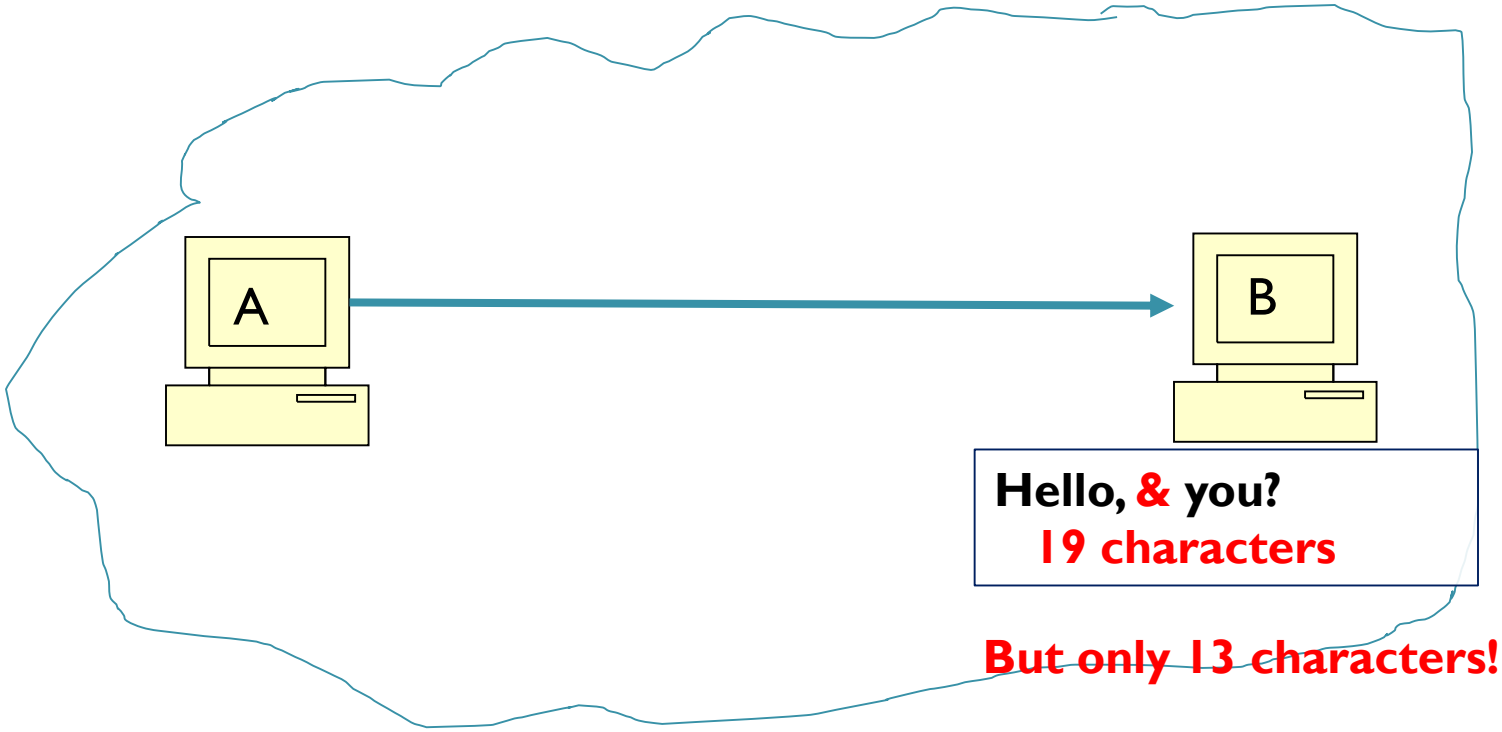
An **error detection code** is calculated for each message. It is sent with the message



- **Computer A** calculates a **code** before sending his message.
- The **code** is the number of characters in the message (including spaces).

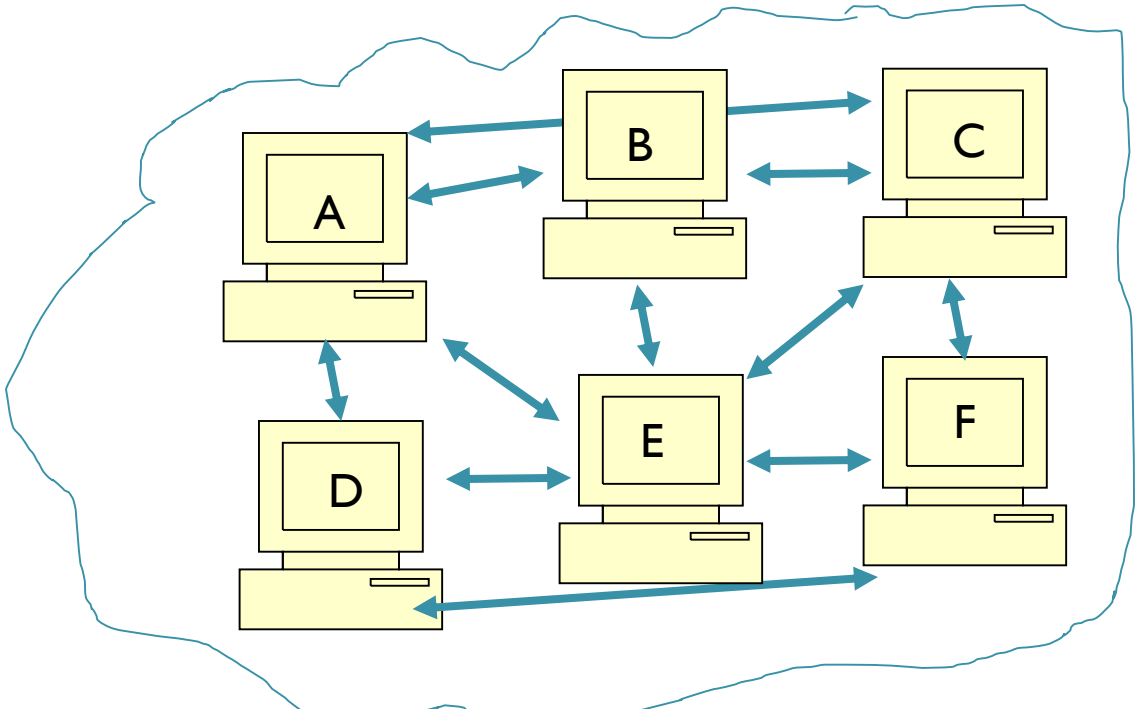
Solution: use an **error detection code** (错误检测代码)

- **Computer B** recalculates the code.



Other challenge: how to know the state of a distributed system?

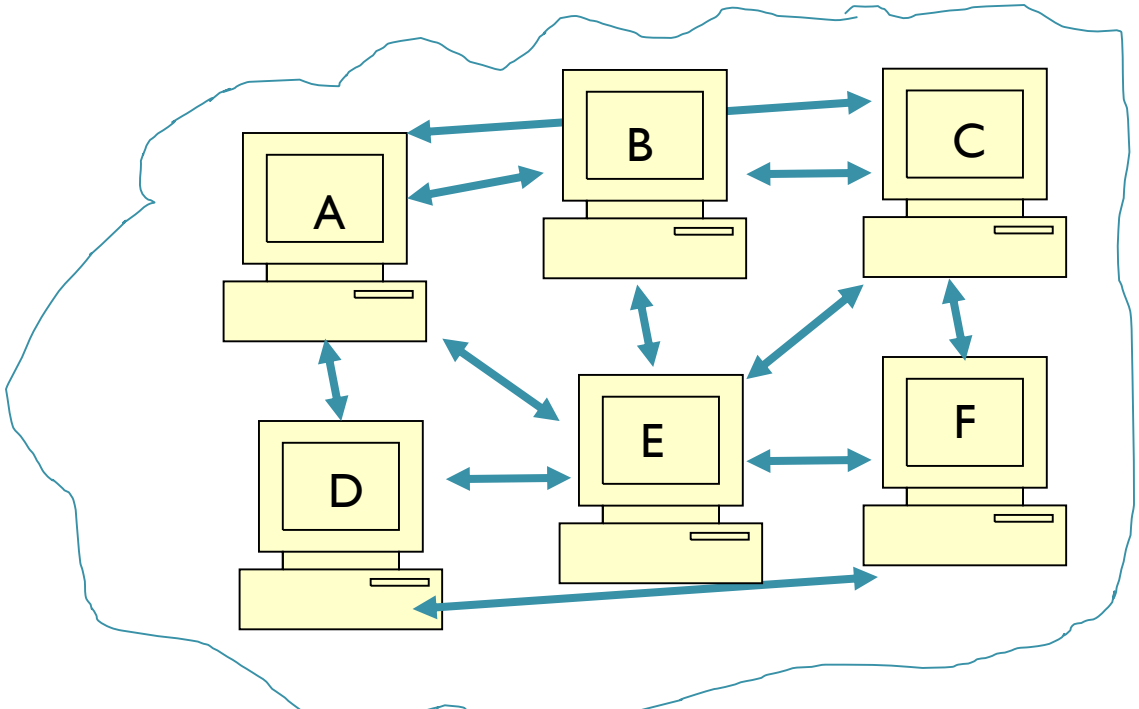
- It is difficult to know the state of the cloud because it is **always changing**!
- If computer A asks what computer B is doing, before A receives the answers, B is already doing something else!



A solution

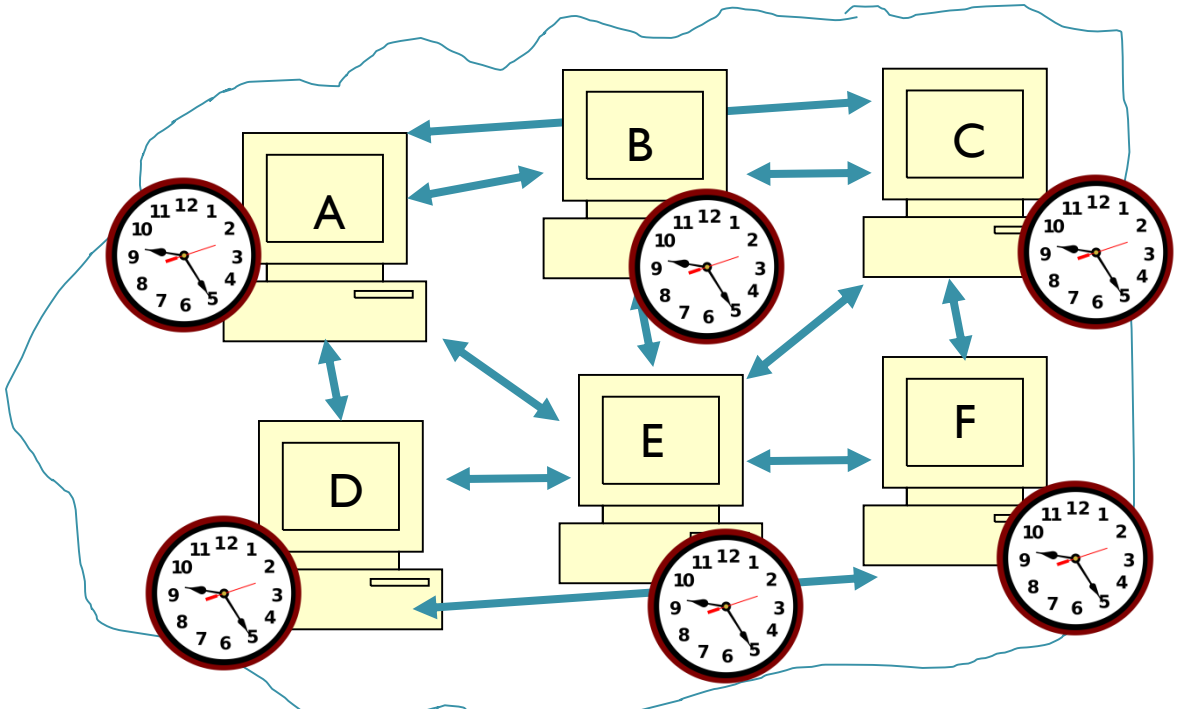
There exists some advanced algorithms (算法) for capturing the state of a distributed system.

We will not talk about this.



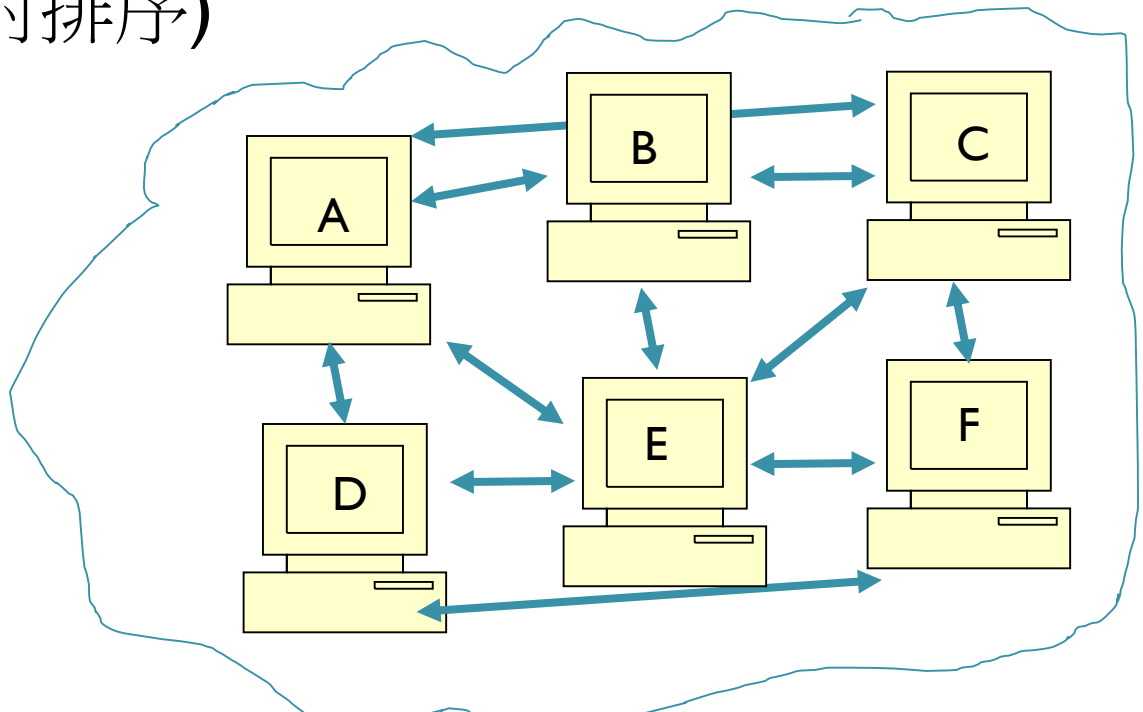
Challenge: how can we measure the time ?

- It is desirable to know **when** something has happened.
- However, clocks of computers are not synchronized.
- How to know if some event happened before some other event in a distributed system?



A solution : logical clocks (逻辑时钟)

- There exists advanced algorithms (算法) for creating some **logical clocks**.
- A logical clock does not measure the exact time of events, but it measures their relative ordering (相对排序)



Conclusion

- Today, we have:
  - continued the introduction,
  - discussed challenges of parallel and distributed computing

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

References

- Chapter 1 and 2. D. C. Marinescu. Cloud Computing Theory and Practice, Morgan Kaufmann, 2013.