JCAICT 2011

Beyond Cloud Computing
– Computer technologies now and the future –

Toshiaki ARAI
Senior Chief Researcher
Systems Development Laboratory
HITACHI, Ltd., JAPAN

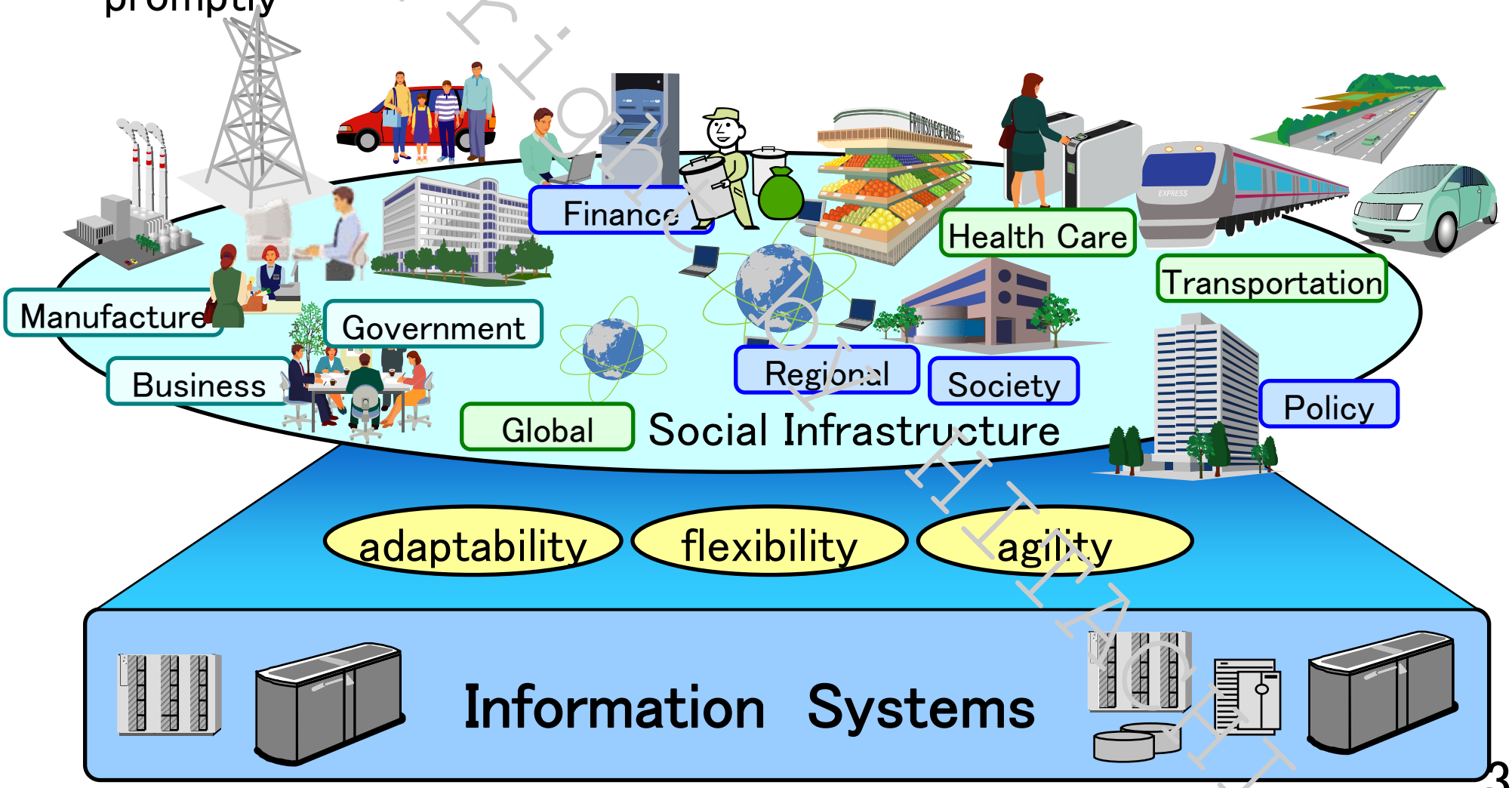
Contents

0

Environment surrounding Information Systems

0-1 Environment Surrounding Information Systems

- Everything in the world is changing rapidly and globally
- Information systems have to be adapted to the changes flexibly and promptly

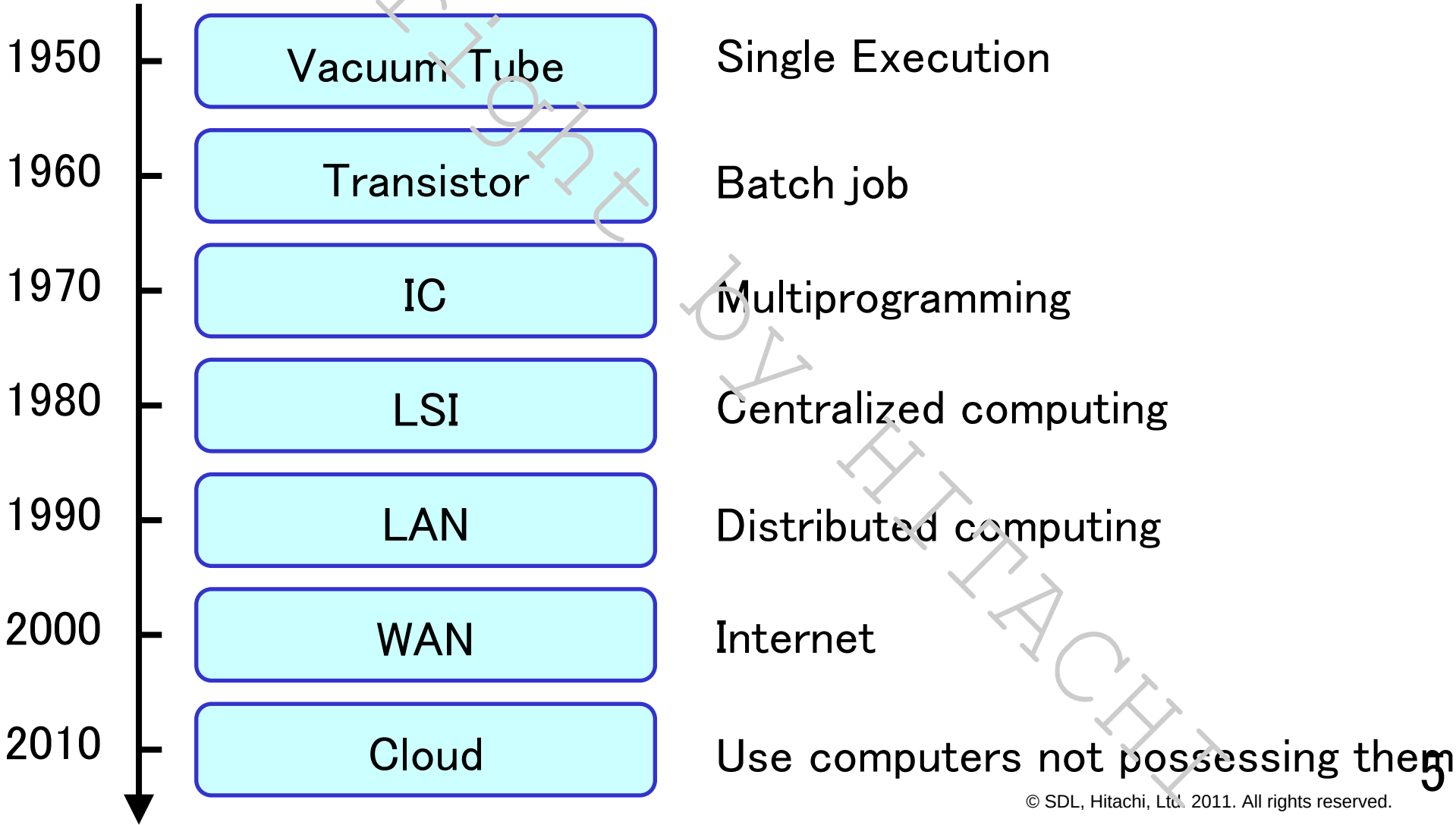


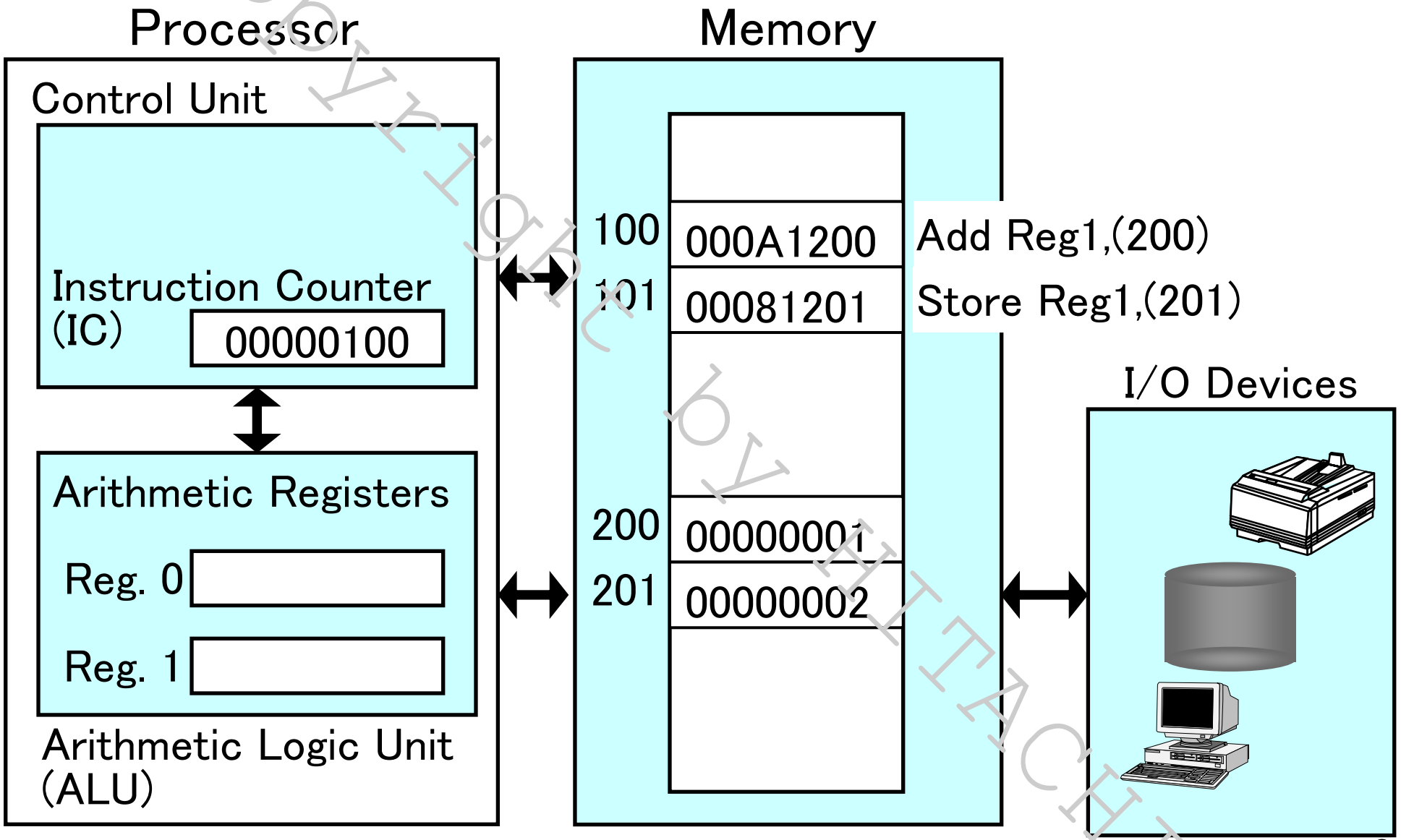
1

History and Evolution of Information Systems

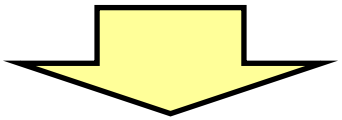
Evolution of Information Systems

- Information system have had great progress about every ten years
- Semiconductor technologies led the progress in the early days
- Network technologies led the progress in the latter



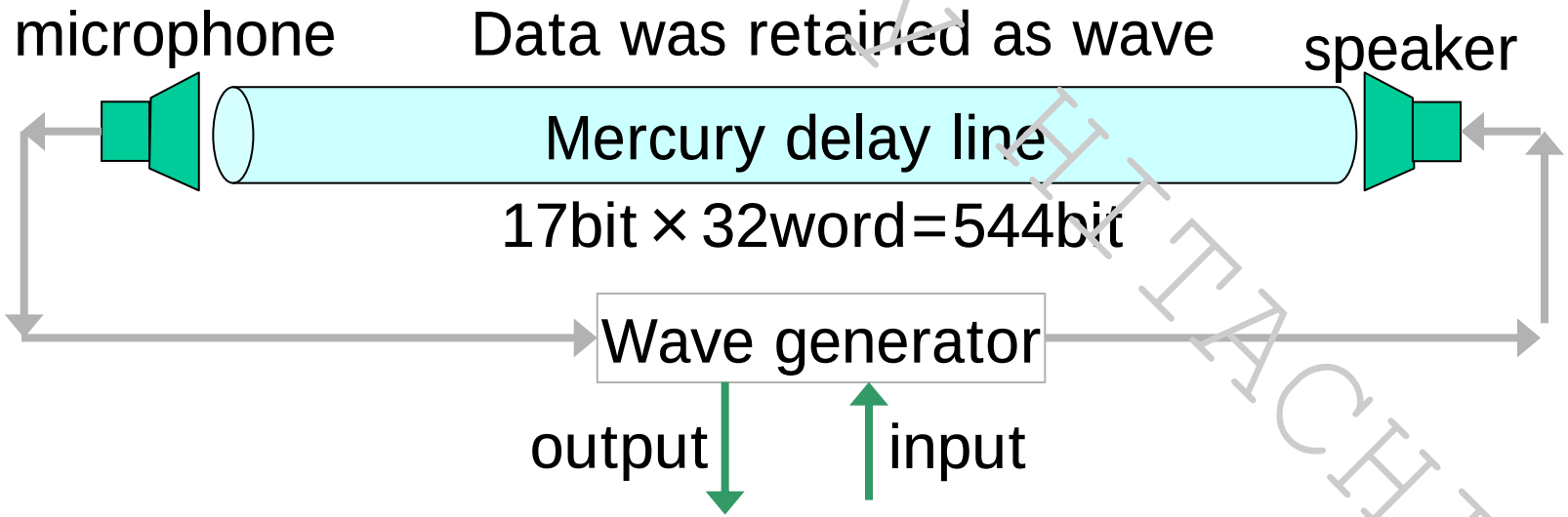


- Keeps programmed instructions in memory as well as data
  - Processor gets an instruction, analyzes it, and executes it
- Earlier computers had hard-wired structure
  - We had to re-wire, re-structure and re-design the computer when executing another program
- Stored program architecture does not need any hardware modification when executing another program
  - Re-load the new program in to the memory

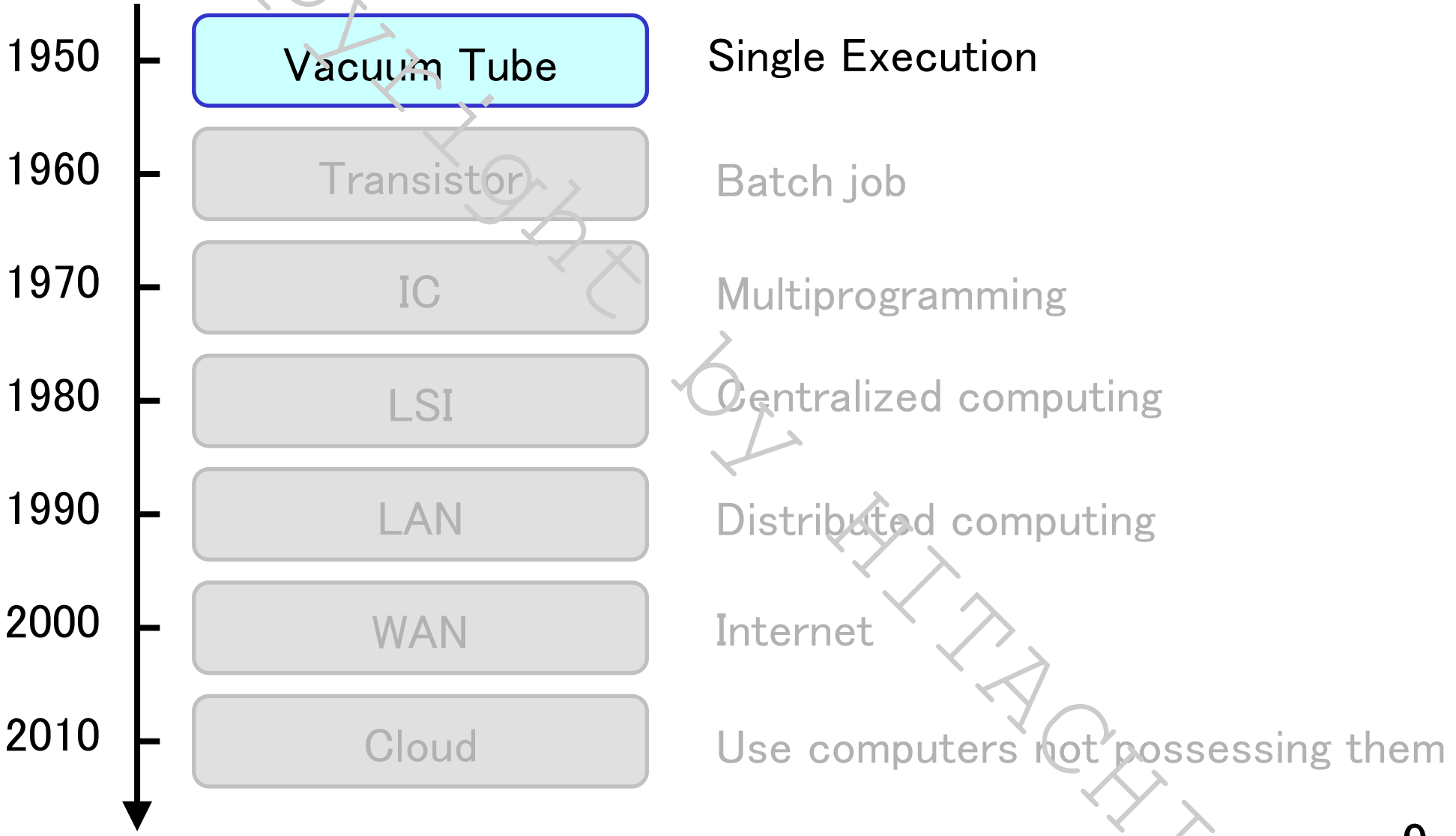


- Concept of software was emerged by stored program architecture
  - Programs can be accumulated and inherited for a long time
  - Information Systems have been able to evolve and progress

- World's First practical Stored Program Computer
  - EDSAC: Electronic Delay Storage Automatic Calculator (1949)
  - Found 73 digit prime number(1951)
  - Computer game：OXO(1952)

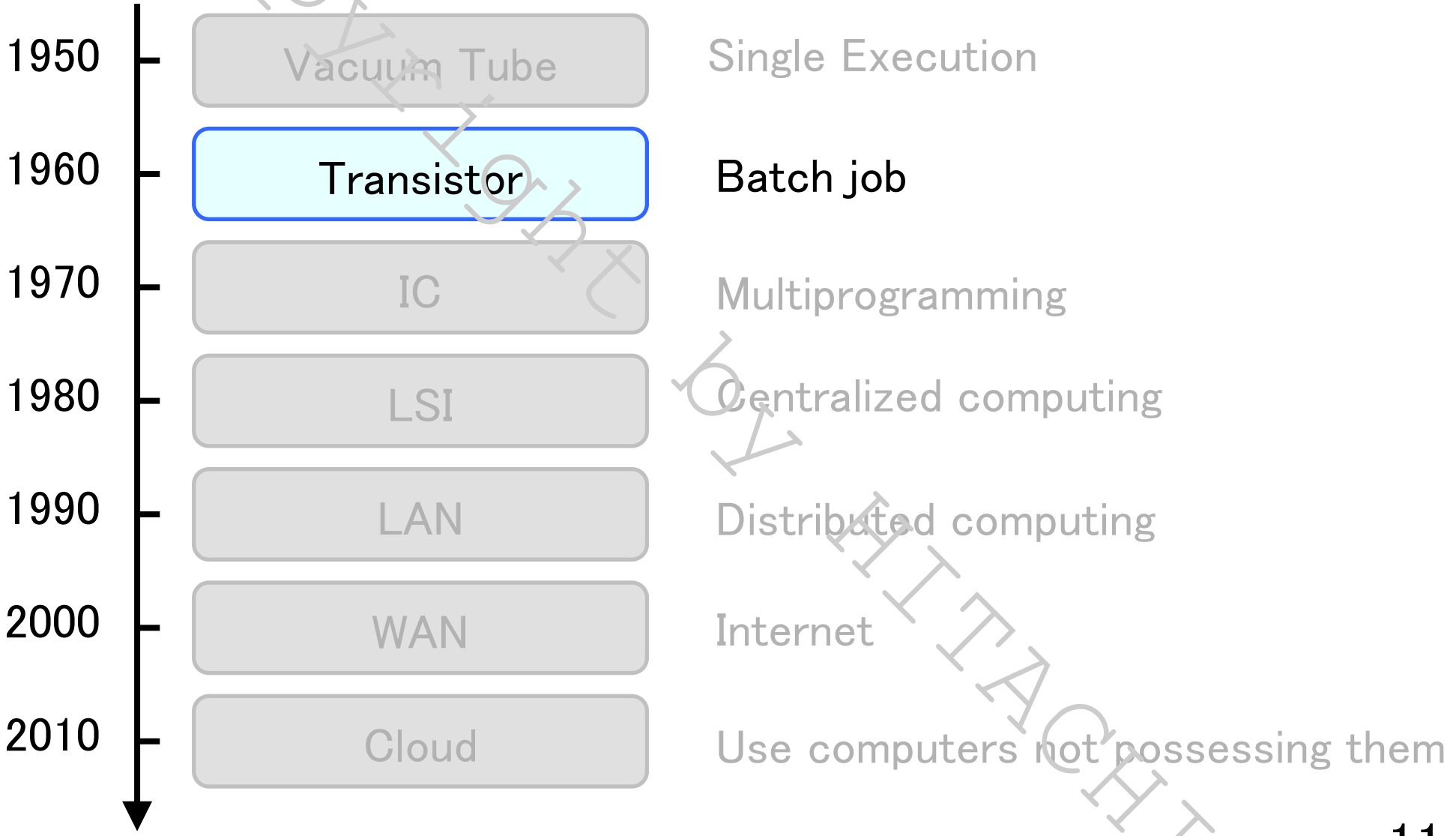


Evolution of Information Systems



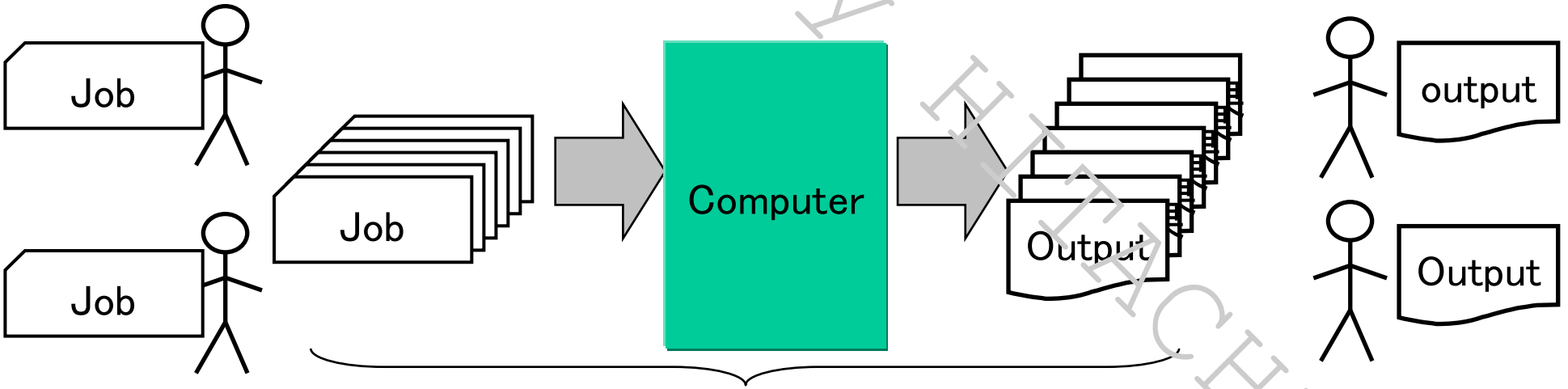
- Thousands of vacuum tubes
- Slow and Low reliable
- A programmer reserved the computer and occupied it
  - The programmer was also the operator, the maintainer, and the designer
- No OS(Operating System)
  - Everything was done manually

Evolution of Information Systems

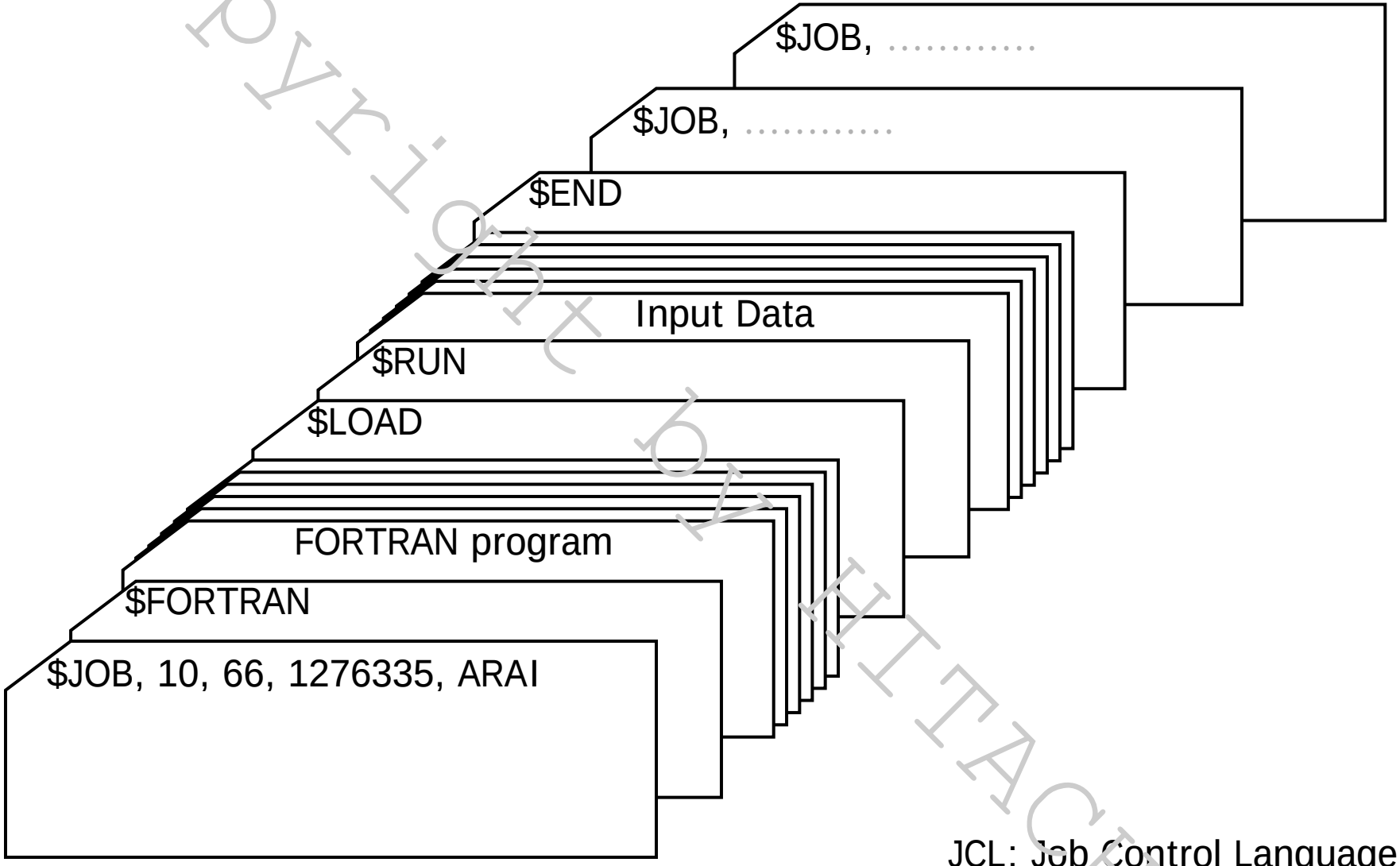


- Transistor was invented in 1947, and started its production in 1954
- Reliability and power consumption were greatly improved
  - MTBF of Transistor: 100,000Hr.
  - MTBF of Vacuum tube: 1,000Hr.
- Computer business started
- High-level Languages and their compilers were introduced
  - FORTRAN(1954): For Scientific Calculation
  - COBOL(1959): For Office Calculation

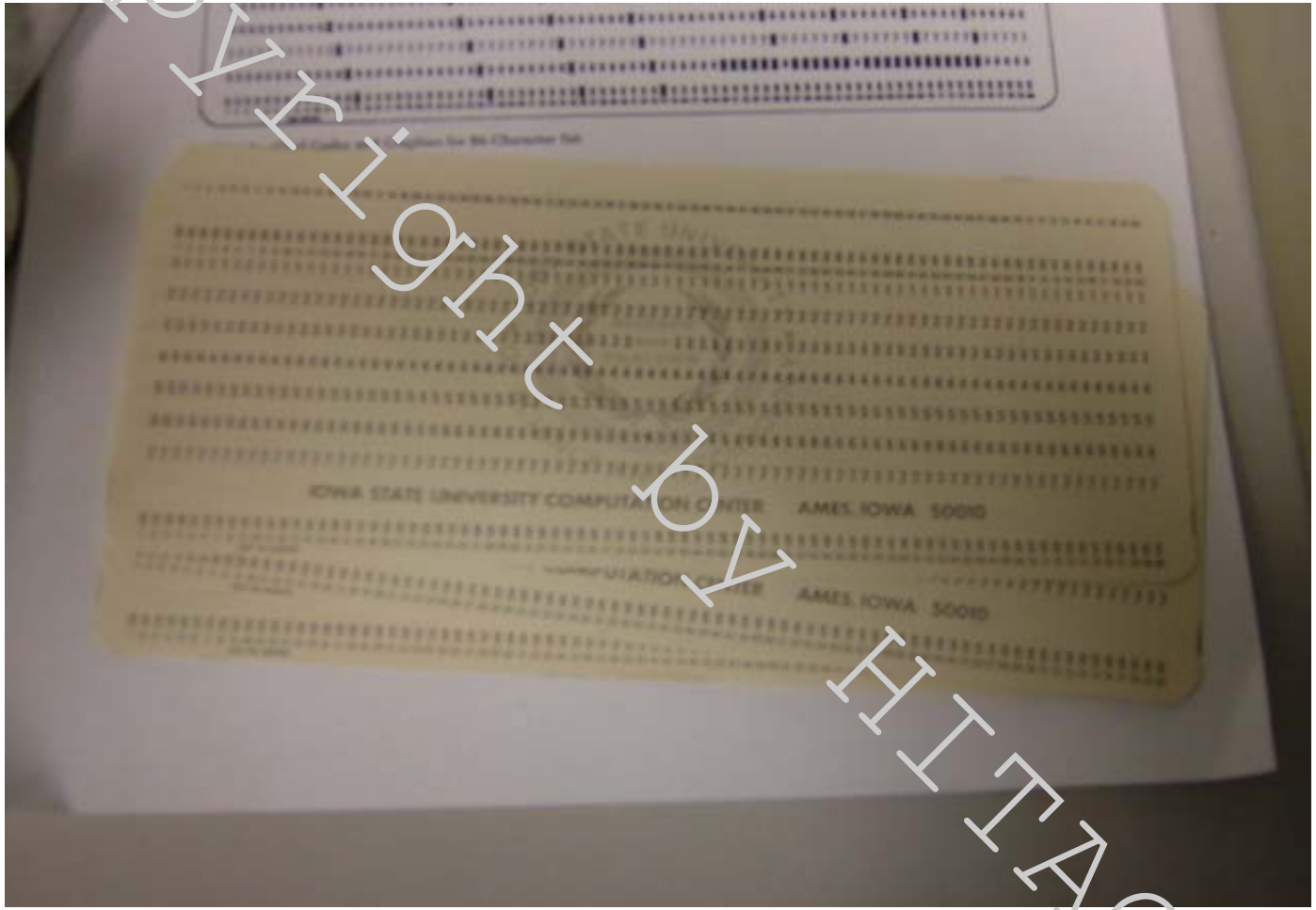
- Computers were still very expensive in spite of being produced by semiconductor devices
- An operator tried to use the computer effectively
  - keep the computer in busy, not in idle state
  - The bottle neck was slow manual operation
- Batch Job execution (Automatic job execution)
  - Jobs are gathered and sent to the computer all together
  - JCL controls jobs' execution instead of the operator

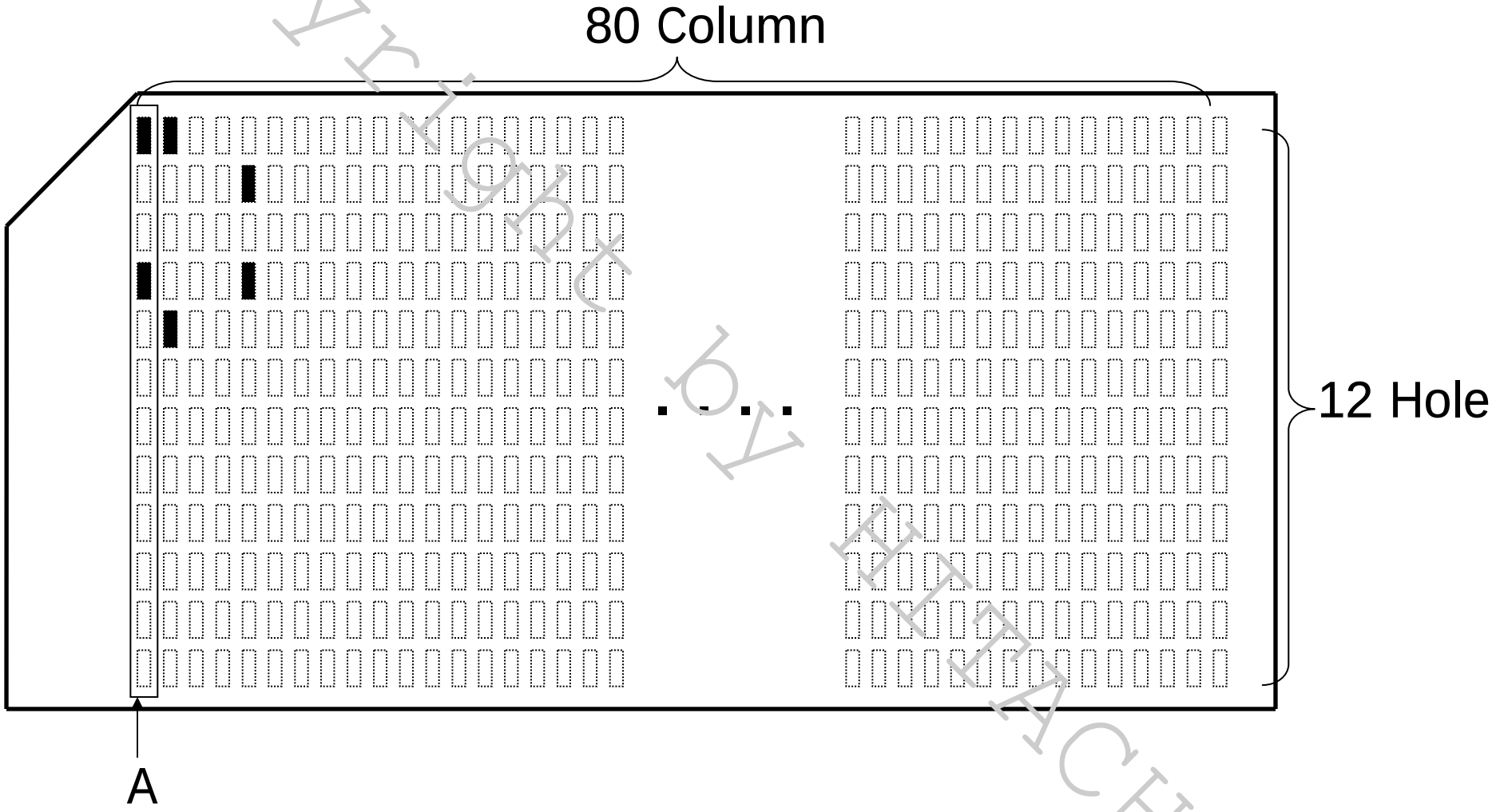


JCL(Job Control Language) controls the Jobs



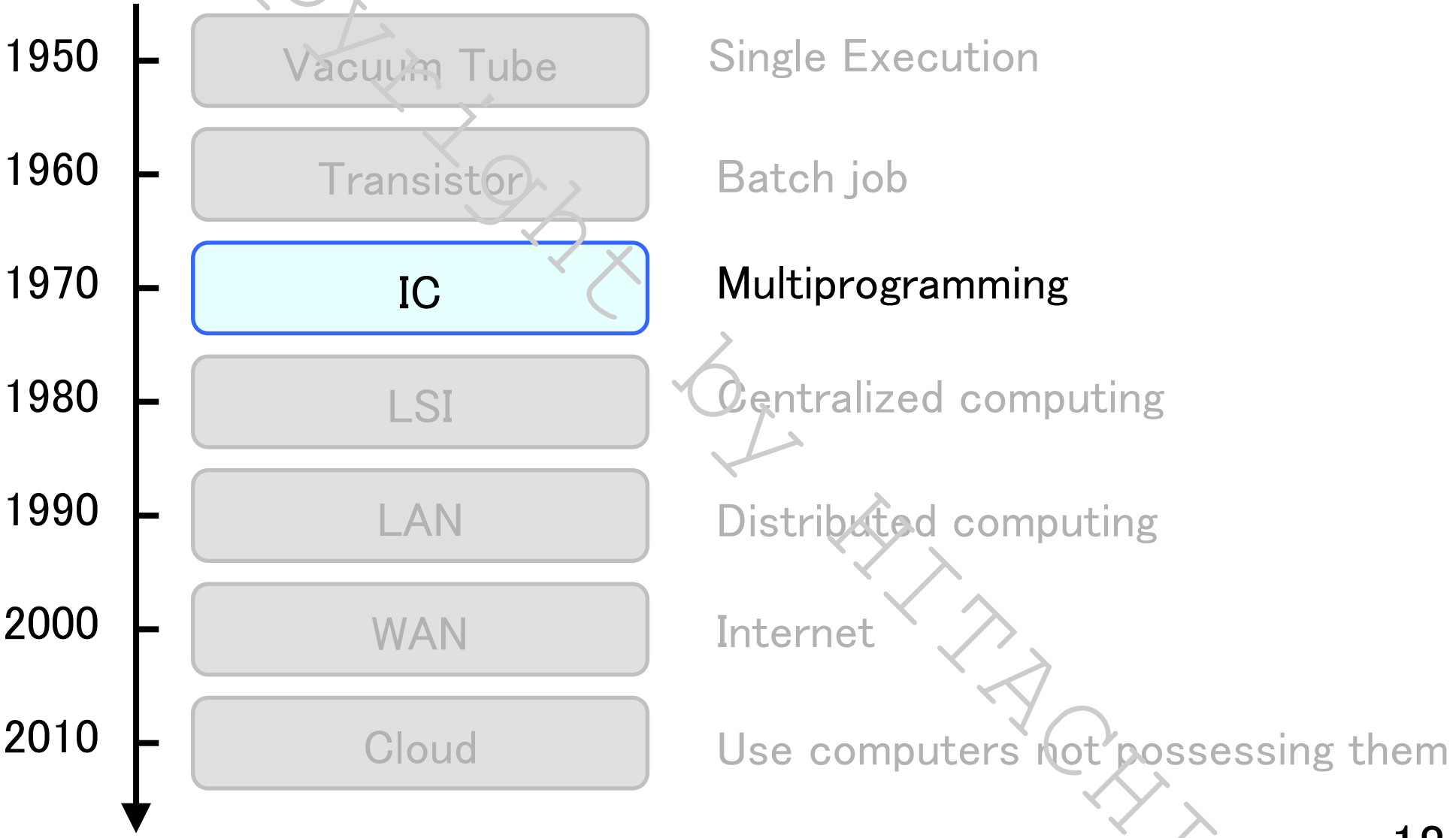
JCL: Job Control Language



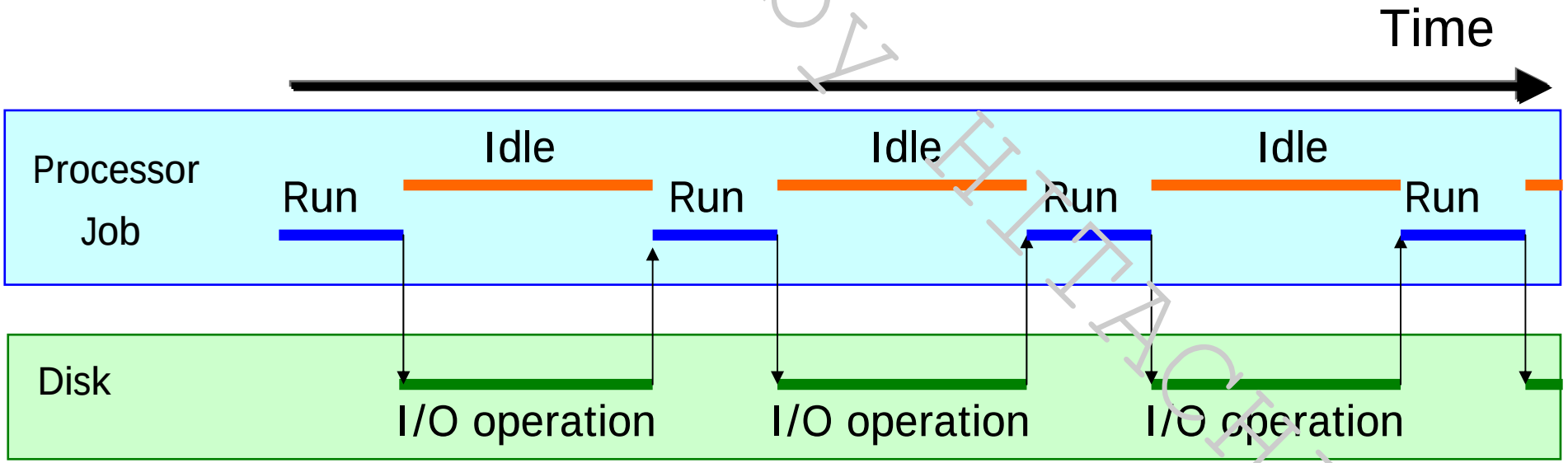


Card Punch Machine

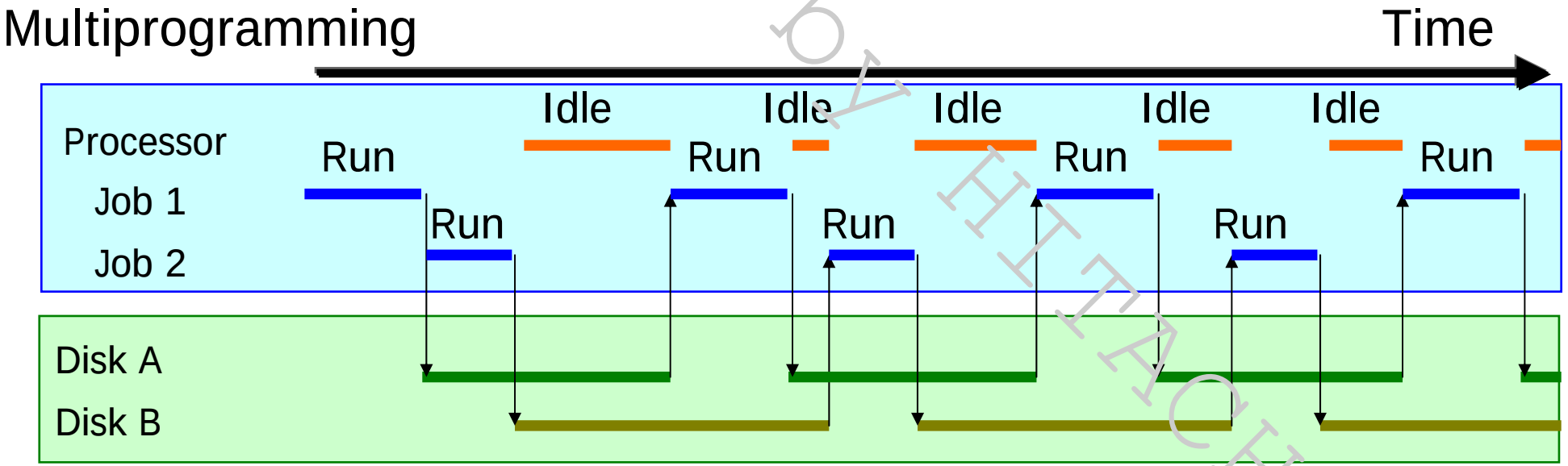
Evolution of Information Systems



- Processor performance was improved by IC technology
- Access Gap Problem: Access speed difference
  - Processor Performance was improved
  - Disk access speed stayed in low because of its mechanical operation
  - Disks became the bottleneck and processors could not be used effectively



- Solution is to overlap processor execution and I/O operation
  - Load multiple jobs in memory
  - When a job enters I/O wait status, another job is executed
  - Processor idle time is reduced

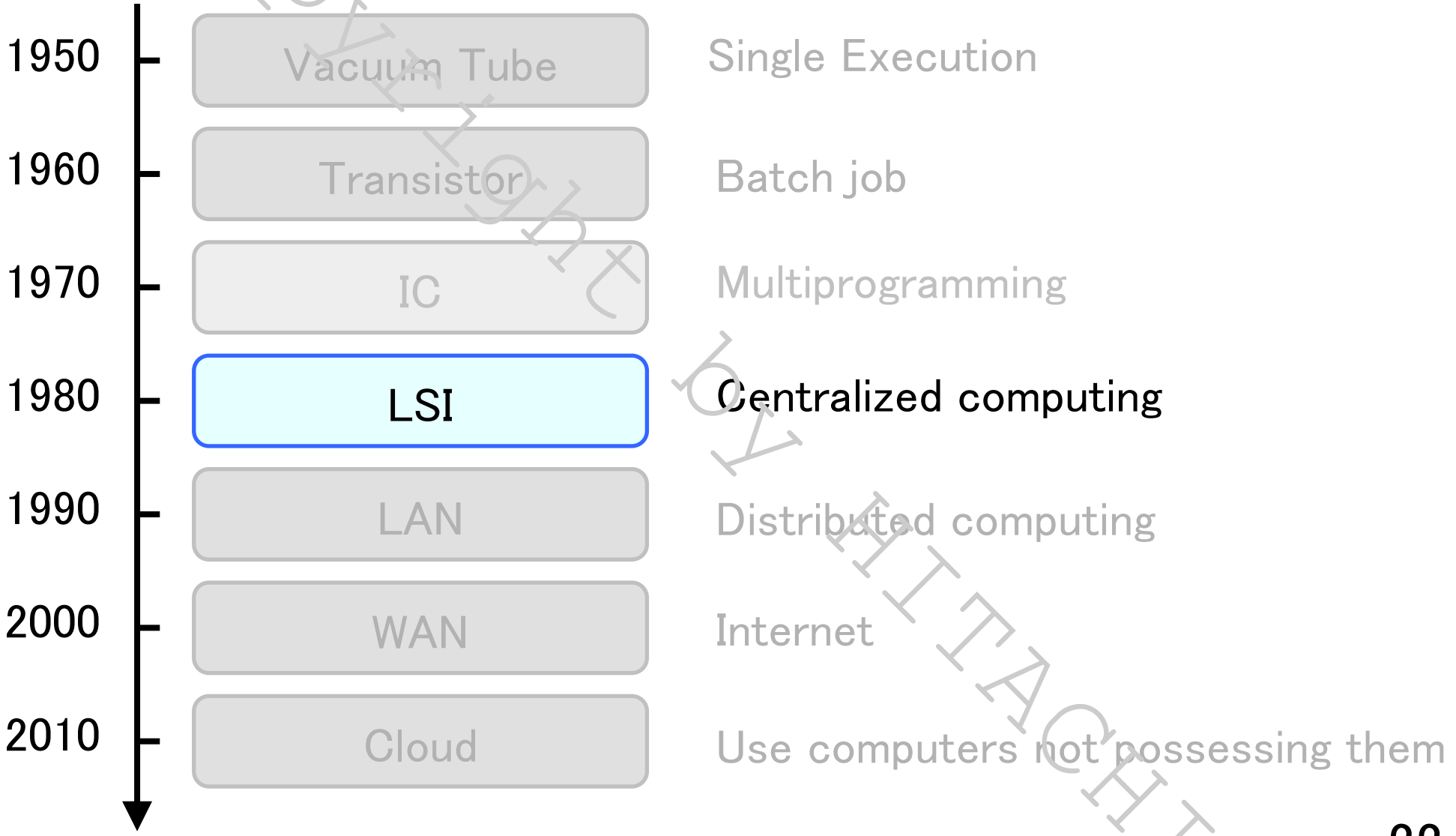


- Job management
  - A job might illegally access another job's memory
    - Memory protection and memory management
  - Which job should be executed first?
    - Scheduling and dispatcher

- Hi-speed processor was bothered by slower peripheral devices
  - Processor was occupied by checking I/O status until the I/O completion, until then
  - An interrupt mechanism was introduced to avoid the wasting processor's valuable time
    - Processor and peripheral devices run independently, and the interrupt mechanism notifies the processor when the device completes the I/O operation
  - Processor can concentrate on its primary tasks

- Major functions of OSs were already developed in this period

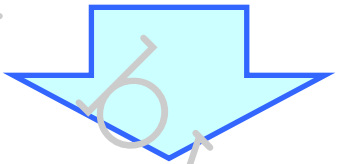
Evolution of Information Systems



- Hardware performance was greatly improved by LSI technology

LSI: Large Scale Integration

- Variety of user demands
  - Use computer anytime, any where, and for any jobs
  - Process and calculate more data and in great detail



- A variety of computing process were executed on one big computer
  - TSS: Time Sharing System
    - Each user uses the computer as if he occupies the computer
  - Online execution, Real-time execution, Big batch processing
- Hundreds of jobs and thousands of TSS users run together

→ Memory problem occurred

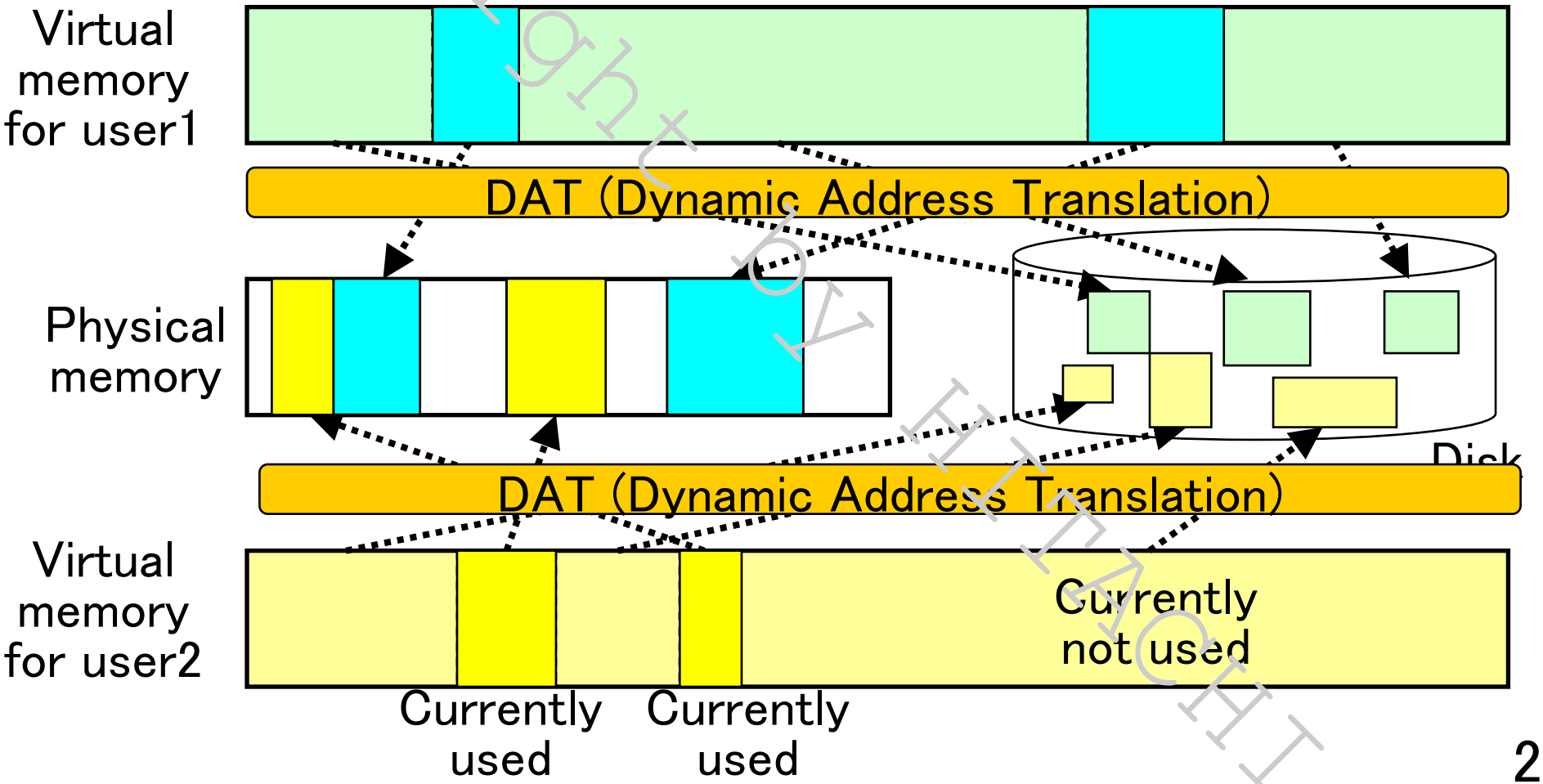
Virtual memory technology

- Demands for memory capacity became severe
  - Programmers wanted much memory than installed
    - Need for storing much data for analysis and calculation
    - For not concerning the size of memory
  - Administrator wanted to raise multi-program level by loading more users into the memory ,and keep processor utilization near 100%

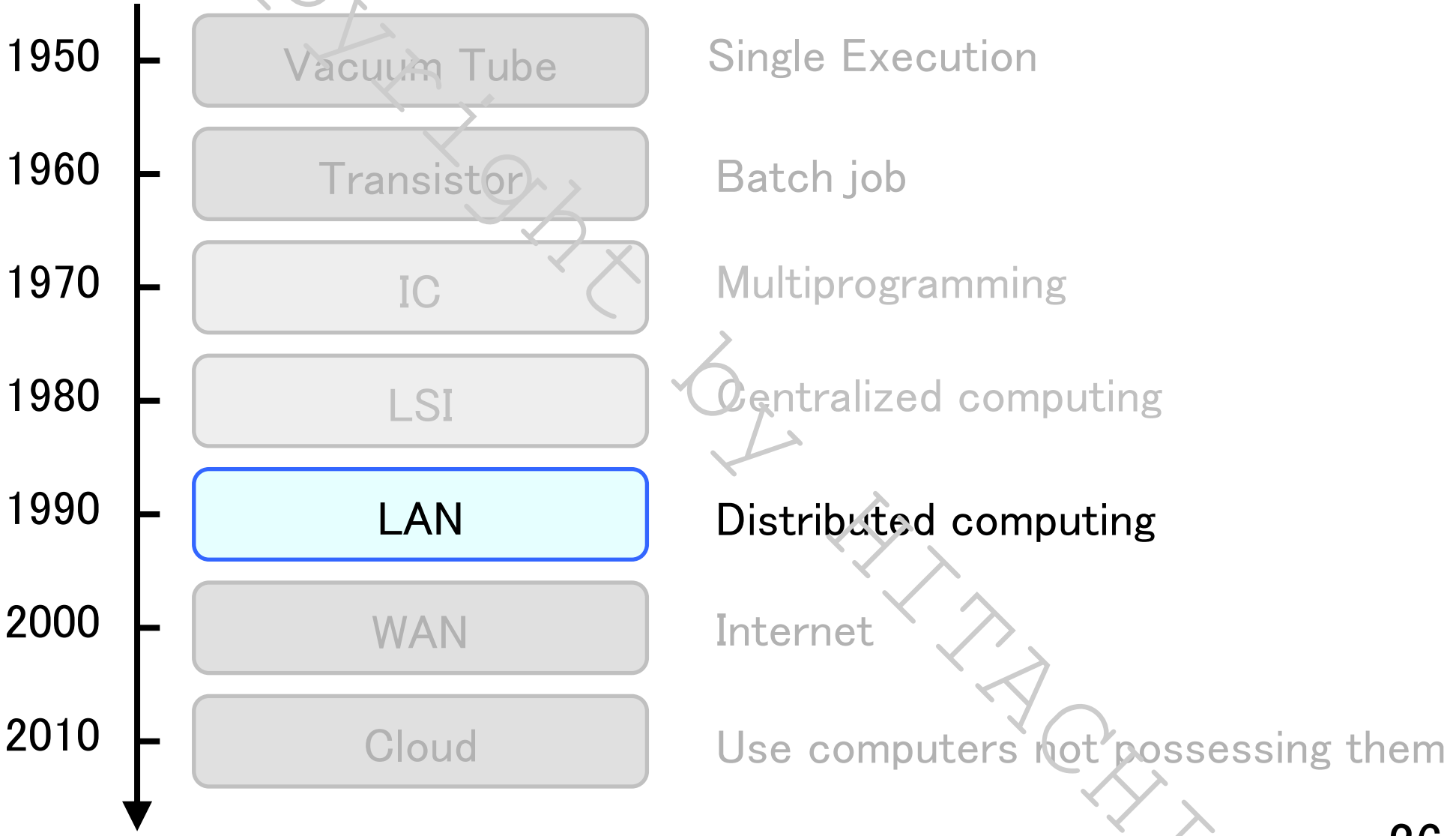
- Virtual memory virtually gives the users more memory capacity than actually installed
- OS supplies the virtual memory by using physical memory and disks

Virtual memory

- Users believe the memory is large enough and contiguous, but in reality the parts it is currently using are scattered around physical memory, and the inactive parts are saved in a disk



Evolution of Information Systems



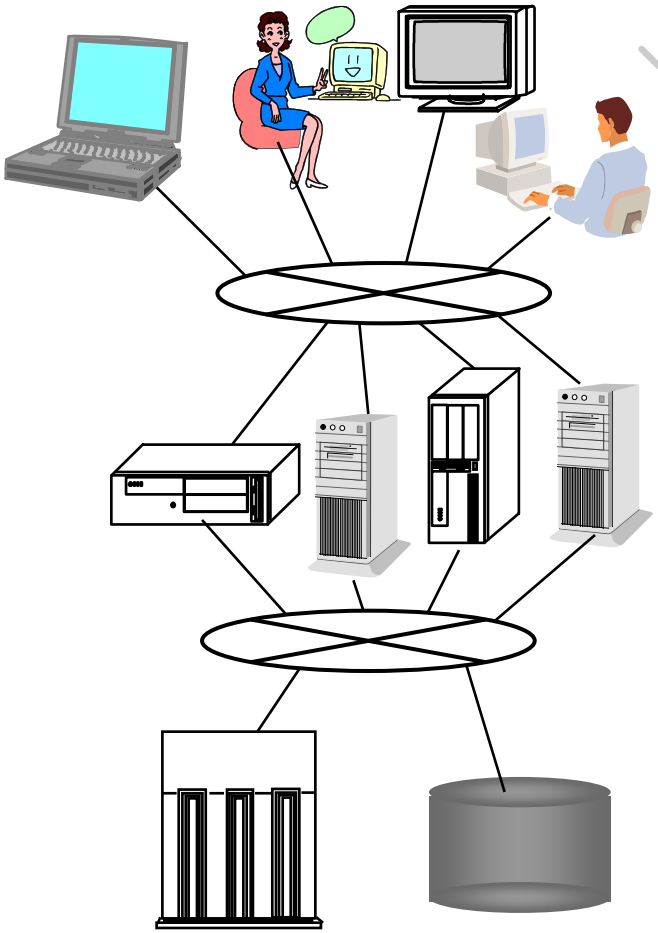
Background of distributed processing

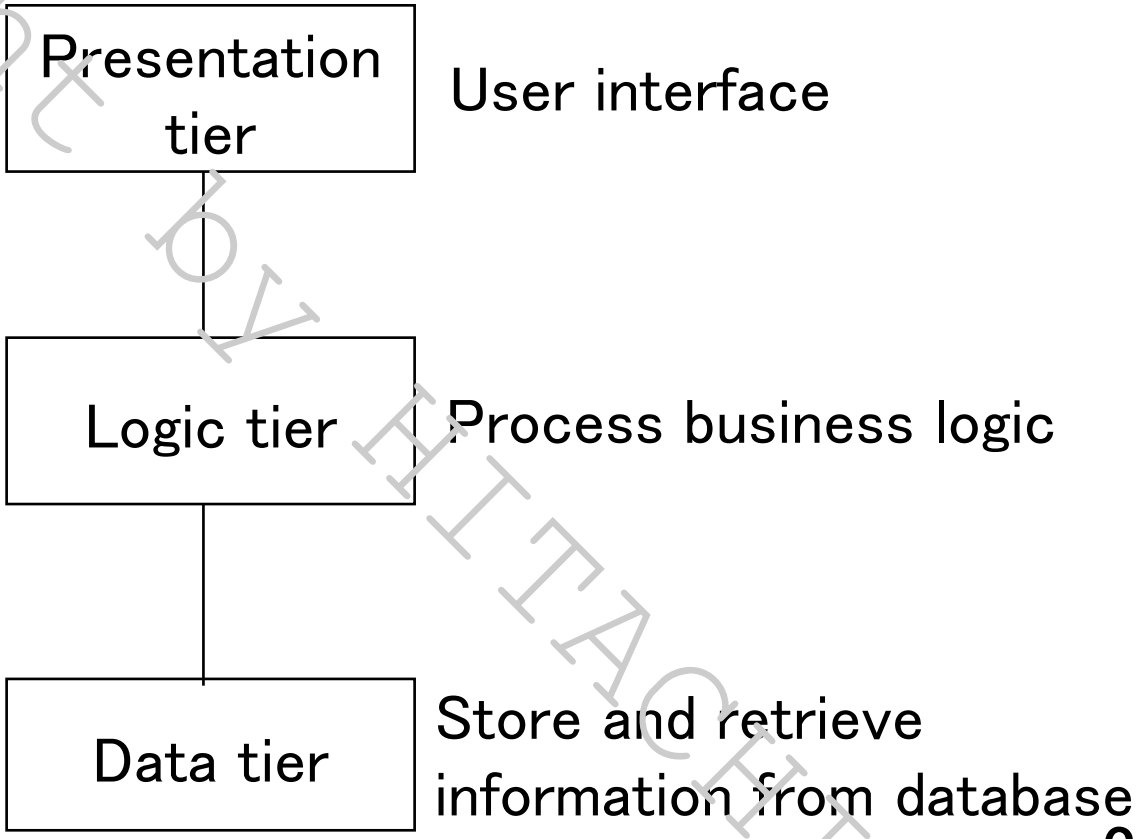
- Microprocessor
  - Intel Corp. began to product microprocessor in 1971
  - A Japanese company asked it to make calculator easily and cheaply
  - Its performance have been improved continually

- ARPANET: The Advanced Research Projects Agency Network
  - DARPA (the Defense Advanced Research Projects Agency) created a new network in 1968
  - The world's first operational packet switching network and the core network of a set that came to compose the global Internet

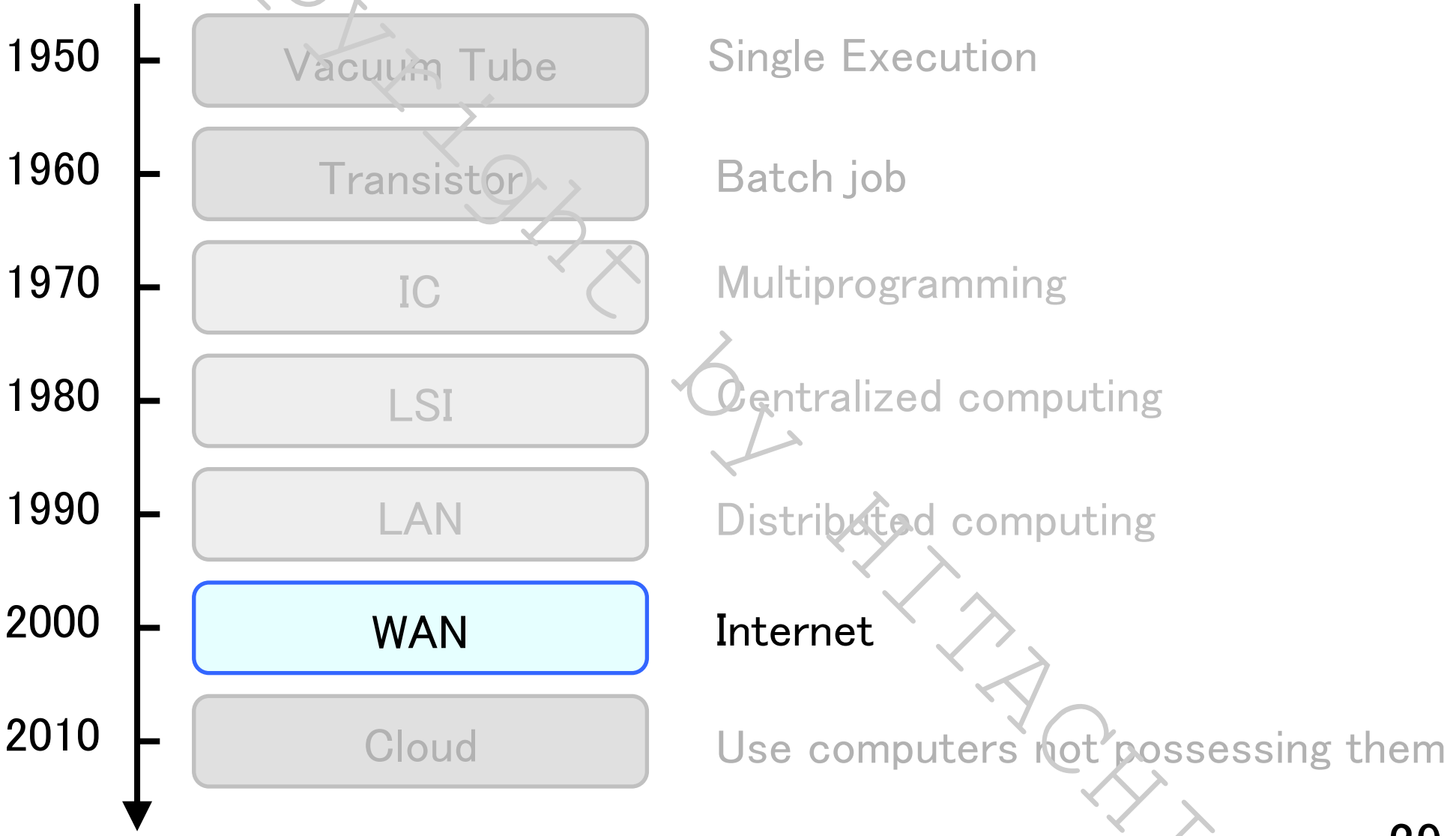
Distributed Computing

- Multiple computers communicate through network in order to achieve common goal
- 3-tier model has advantages in scalability and maintenancebility by sharing the responsibility among the servers





Evolution of Information Systems

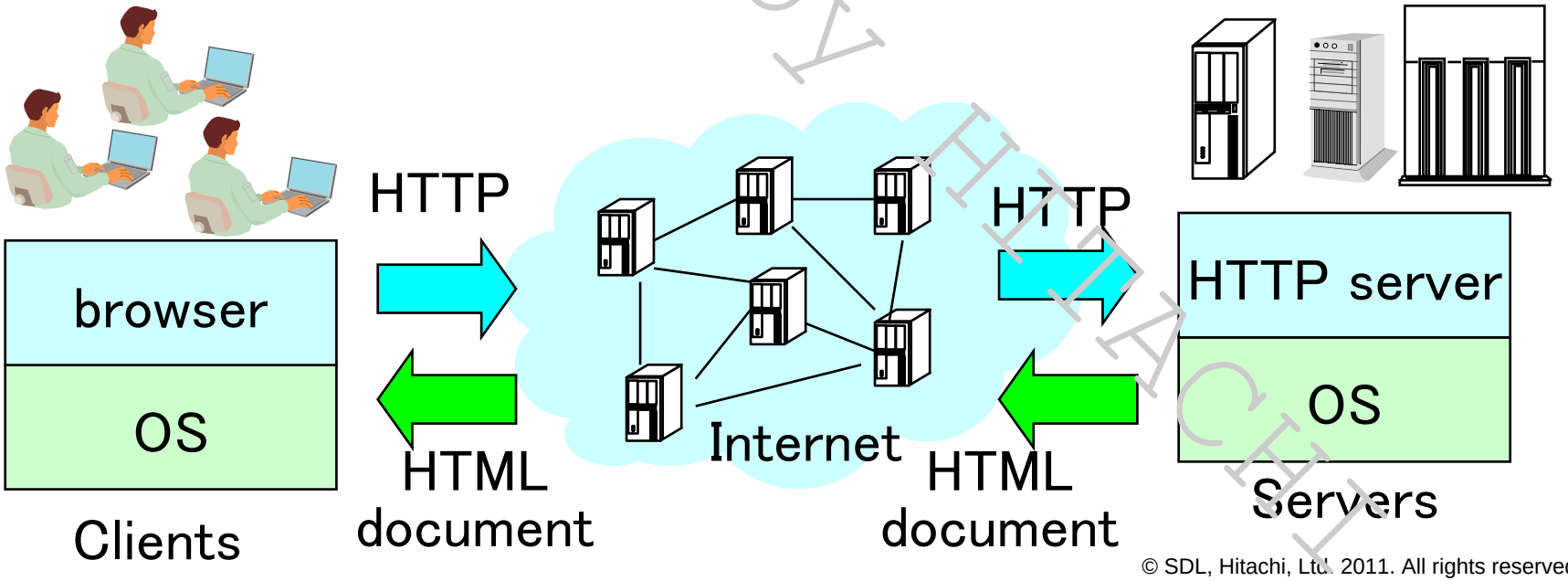


WWW (World Wide Web)

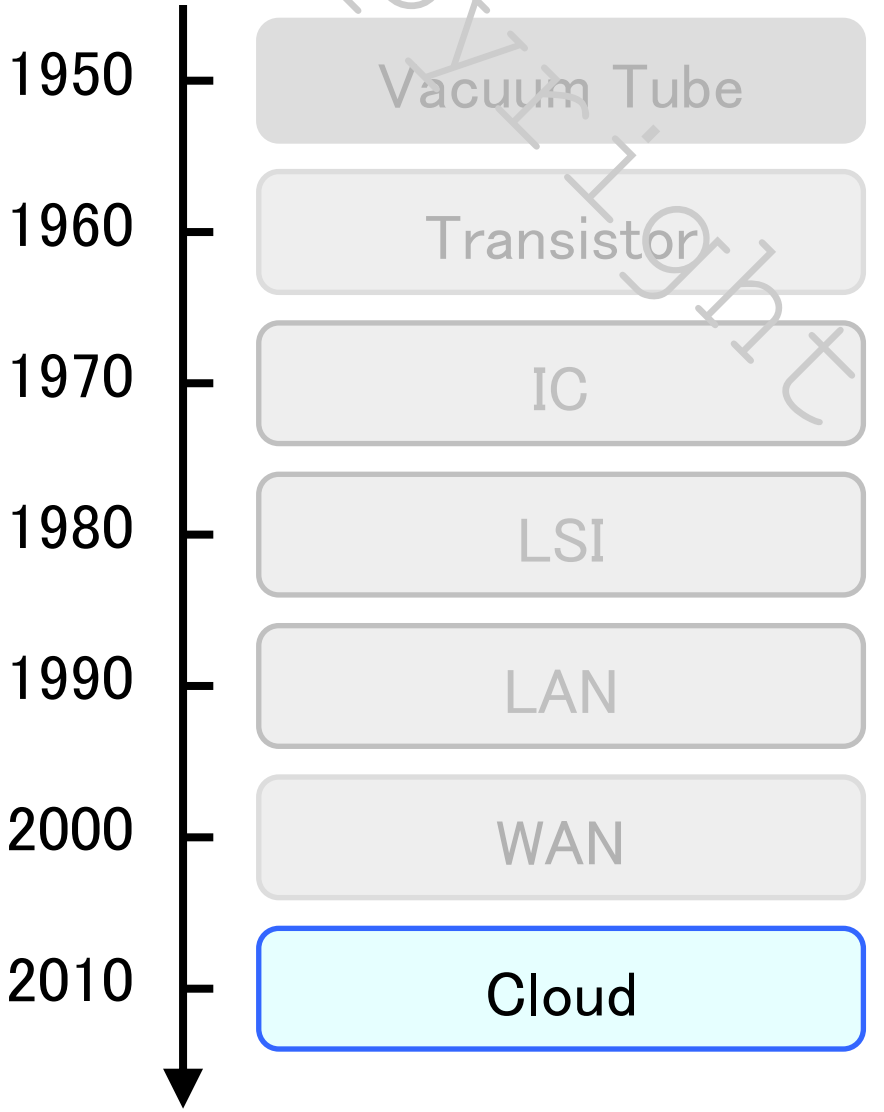
- WWW was developed to be a pool of human knowledge, and human culture, which would allow collaborators in remote sites to share their ideas and all aspects of a common project
- HTTP specifies the protocol among clients and servers
- HTML describes documents and their structures

HTTP: HyperText Transport Protocol
HTML: HyperText Markup Language

- Servers and clients on different hardware or on different OS can communicate each other



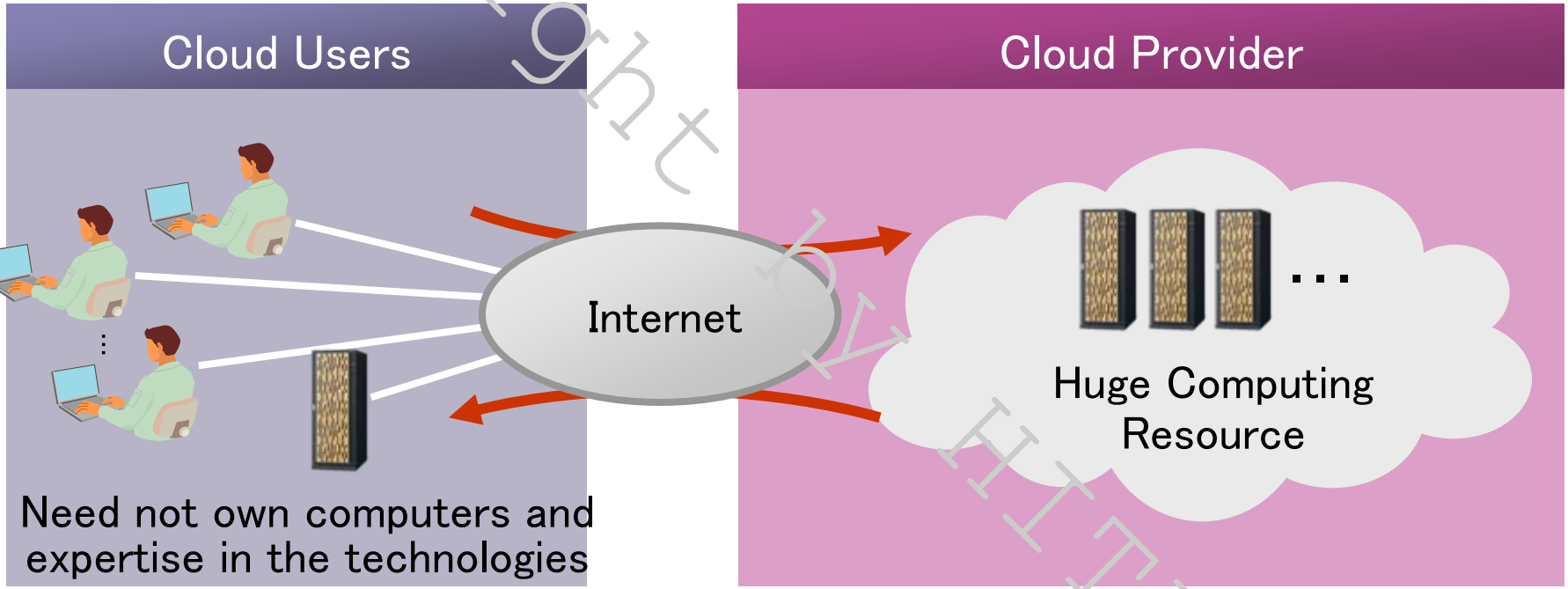
Evolution of Information Systems



2

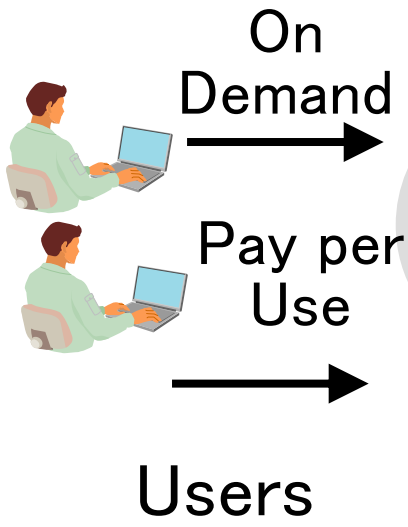
Technologies supporting Cloud Computing

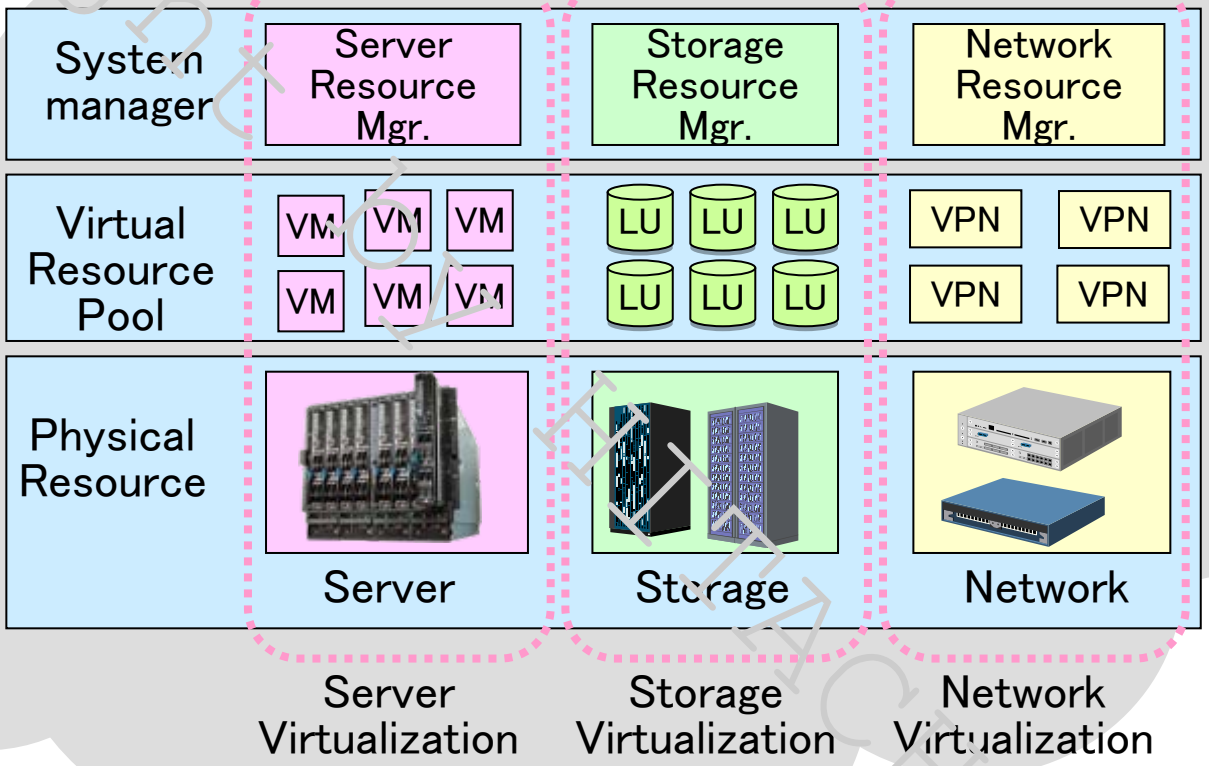
- Cloud is a new computing style that we can use scalable huge computing resources as a service through the Internet
- Cloud computing is spreading rapidly



Utilization | Scalable | On-Demand

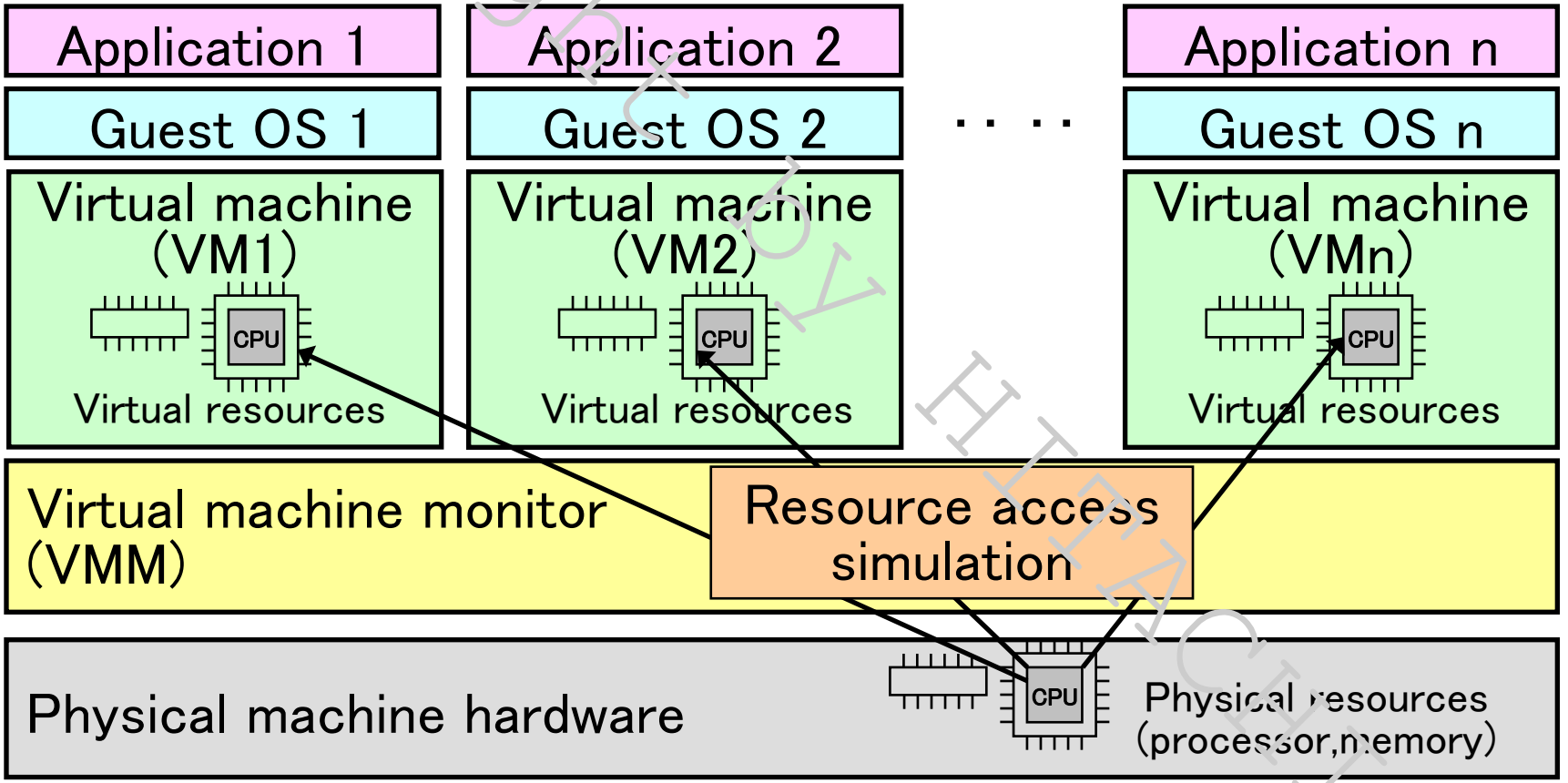
- Cloud Provider supplies the resources rapidly and effectively according to the users' requests
- Cloud consists of a collection of virtualized computer resources





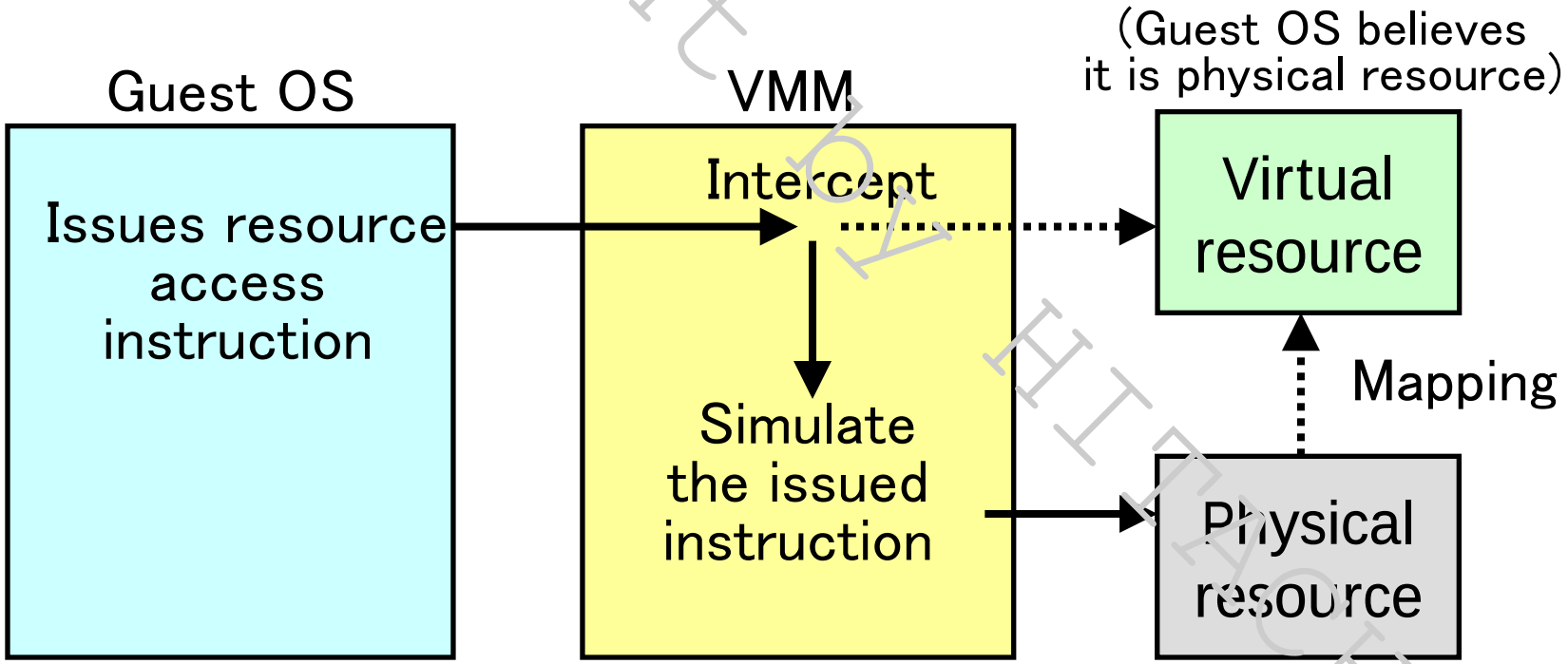
VM: Virtual Machine
LU: Logical Volume
VPN: Virtual Private Network

- Construct multiple virtual machines on a single physical machine, and run an operating system on each virtual machine independently
- VMM virtualizes physical resources such as processor and memory, and assign them to each VM



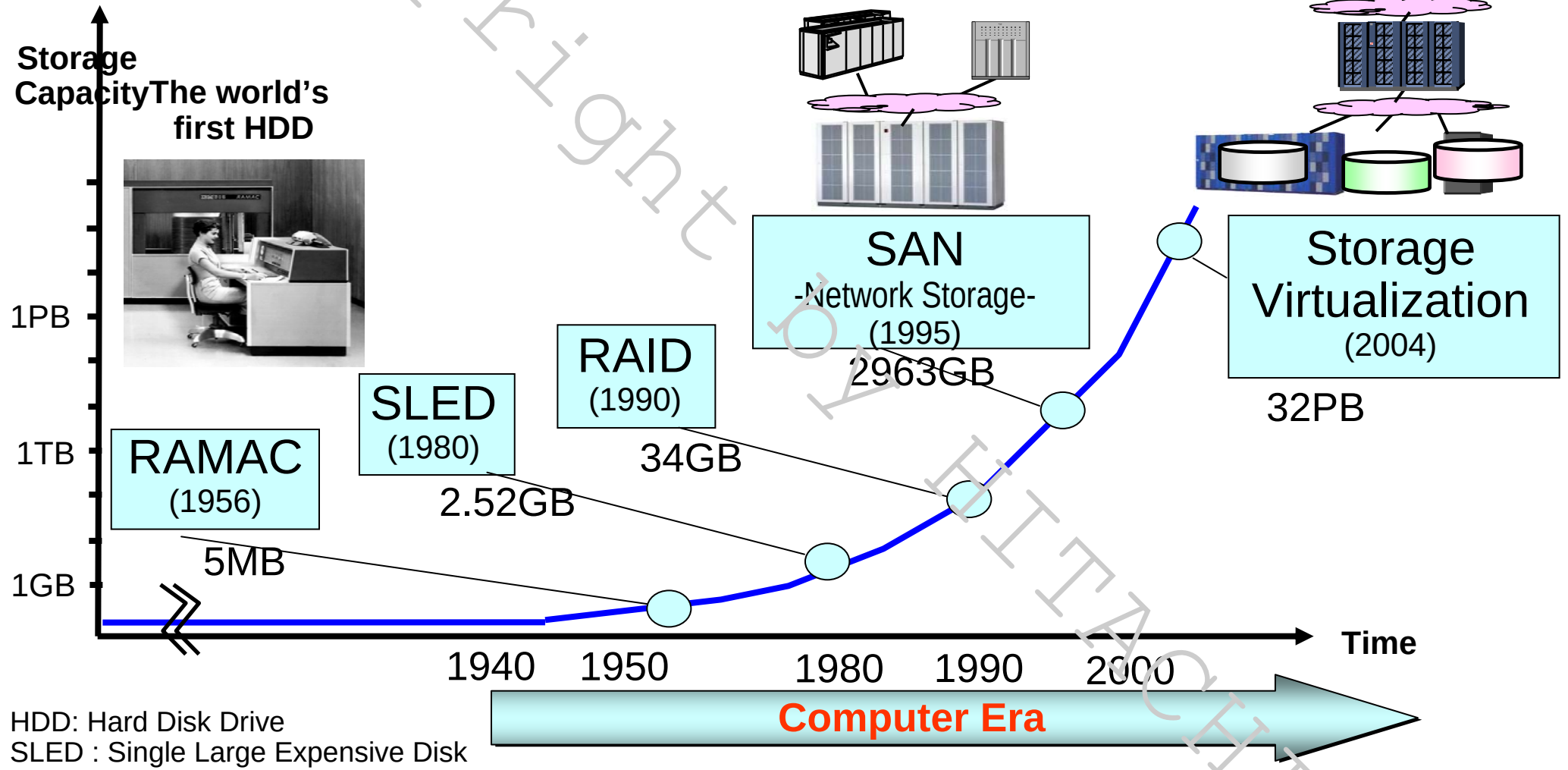
VMM simulates resource access

- Guest OS believes that it accesses physical resources
  - When guest OS issues resource access instruction
  - VMM intercepts the issued instruction
  - And simulates as if issued instruction is executed by accessing the physical resource which is corresponding to the virtual resource



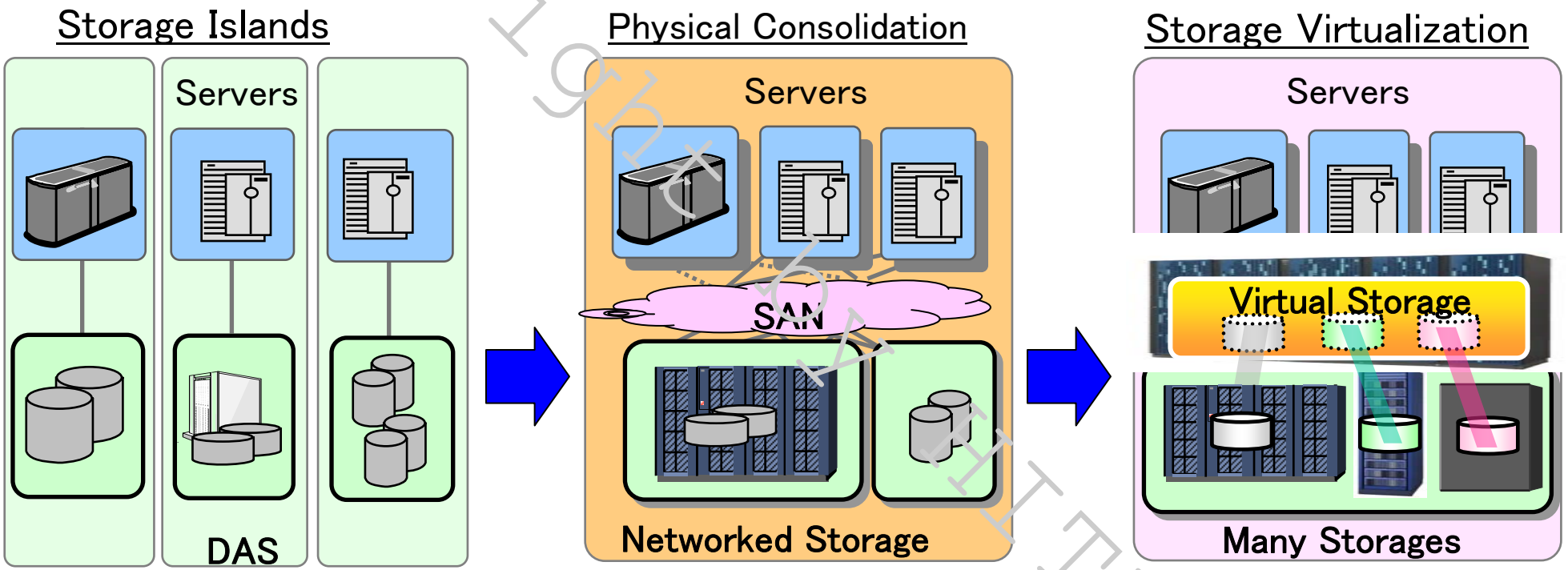
History of storage

- Storage is the most important element in information systems because information is permanently stored only in storage



HDD: Hard Disk Drive
SLED : Single Large Expensive Disk
RAID : Redundant Arrays of Inexpensive Disks
SAN : Storage Area Network

- Storage configuration has been evolved for easy usage
  - SAN consolidated storages and solved capacity unbalance
  - Storage virtualization reduces management cost



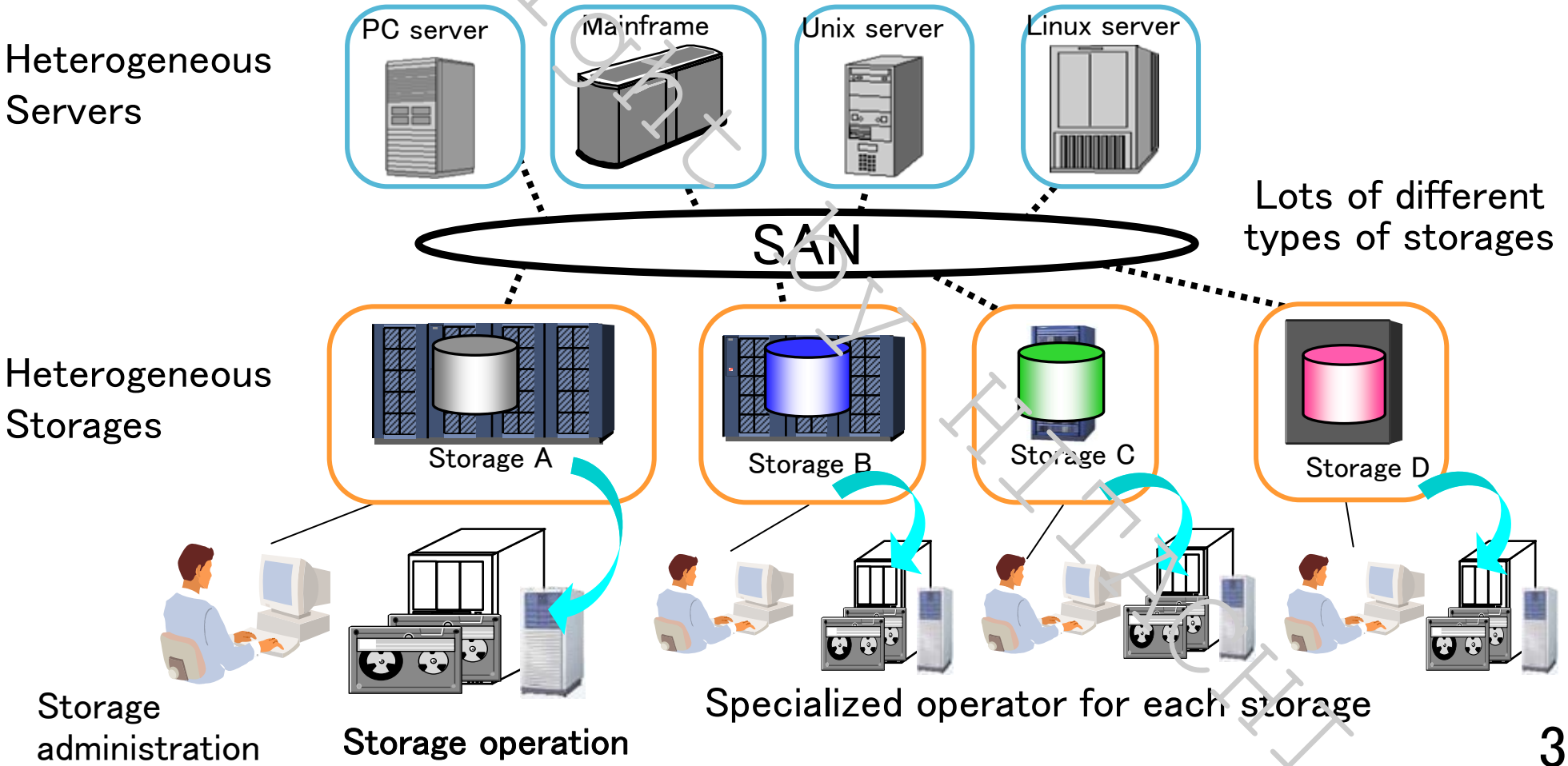
A storage is directly attached to a server.

Servers and storages are connected with one another.

Servers are connected to a virtual storage.

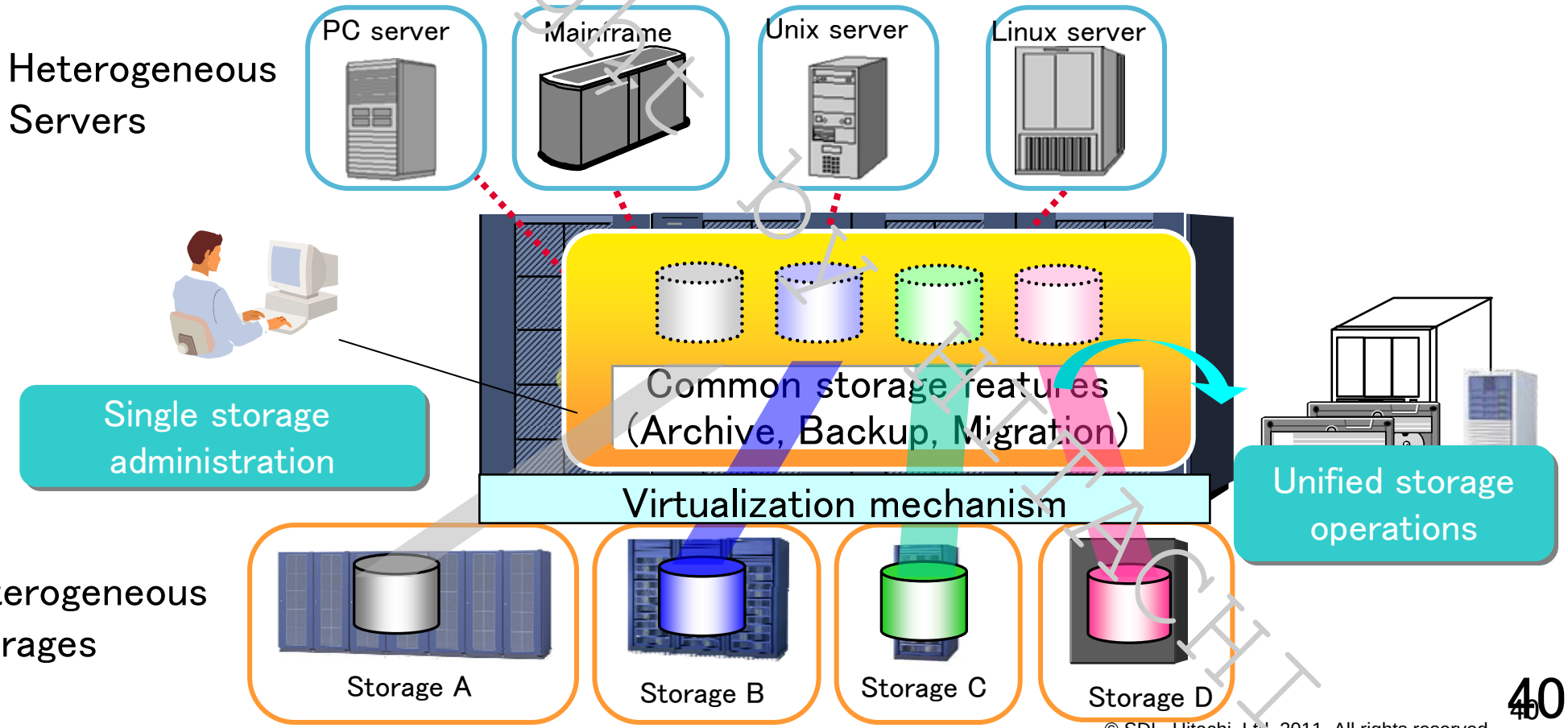
DAS: Direct Attached Storage, SAN: Storage Area Network

- Multiple heterogeneous storages in a SAN due to information explosion
  - Each storage is operated by the specialized operator and administrator
  - Storage operation and administration cost increased



2-8 Solution (Storage Virtualization)

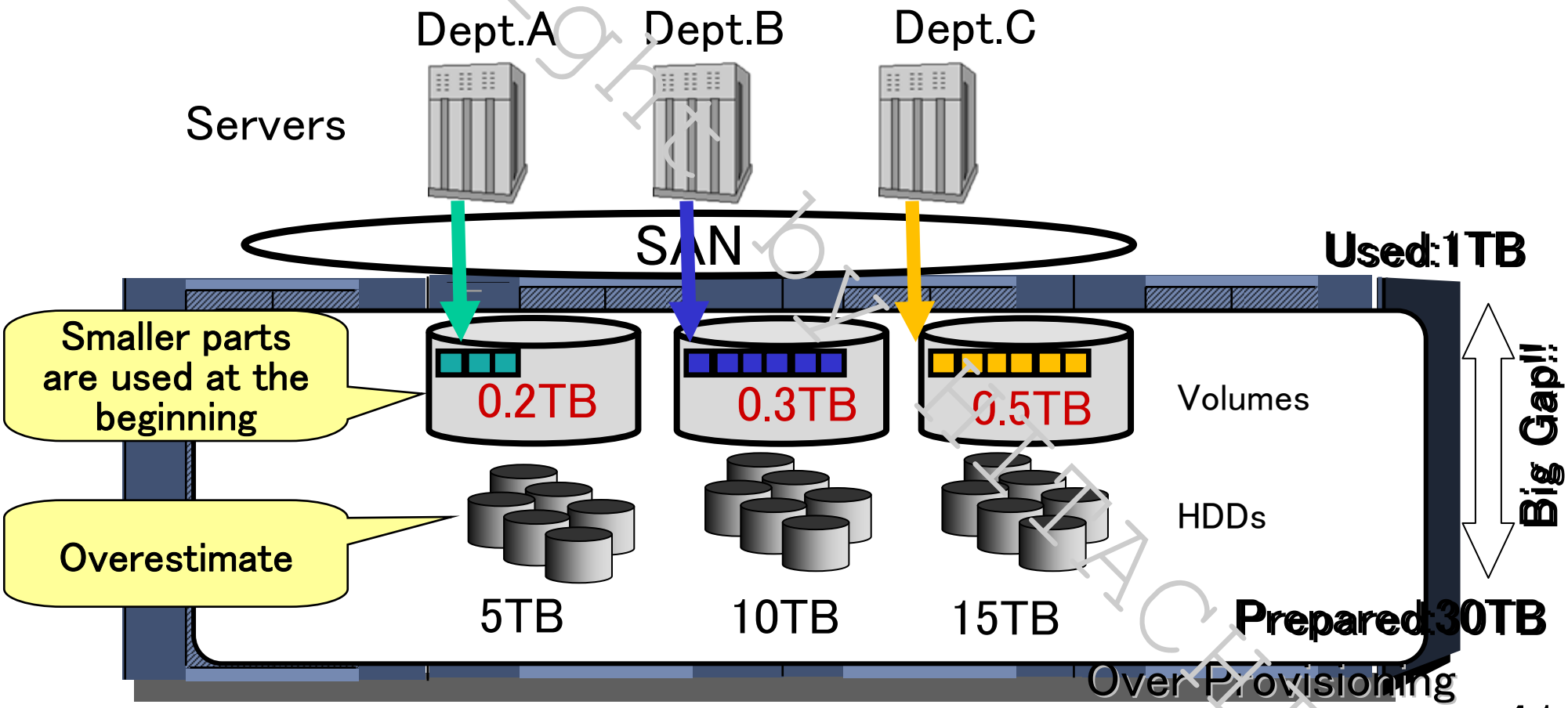
- Storage virtualization provides virtual storage with common feature
  - Virtualization mechanism hides the characteristics of heterogeneous storages
  - Storage operations are unified
  - Administration and operation cost decreased



2-9 Another Problem in SAN

- Provisioning difficulty due to information explosion.
  - Increasing cost of HDDs for over-provisioned volumes
  - Need service outage for expanding capacity

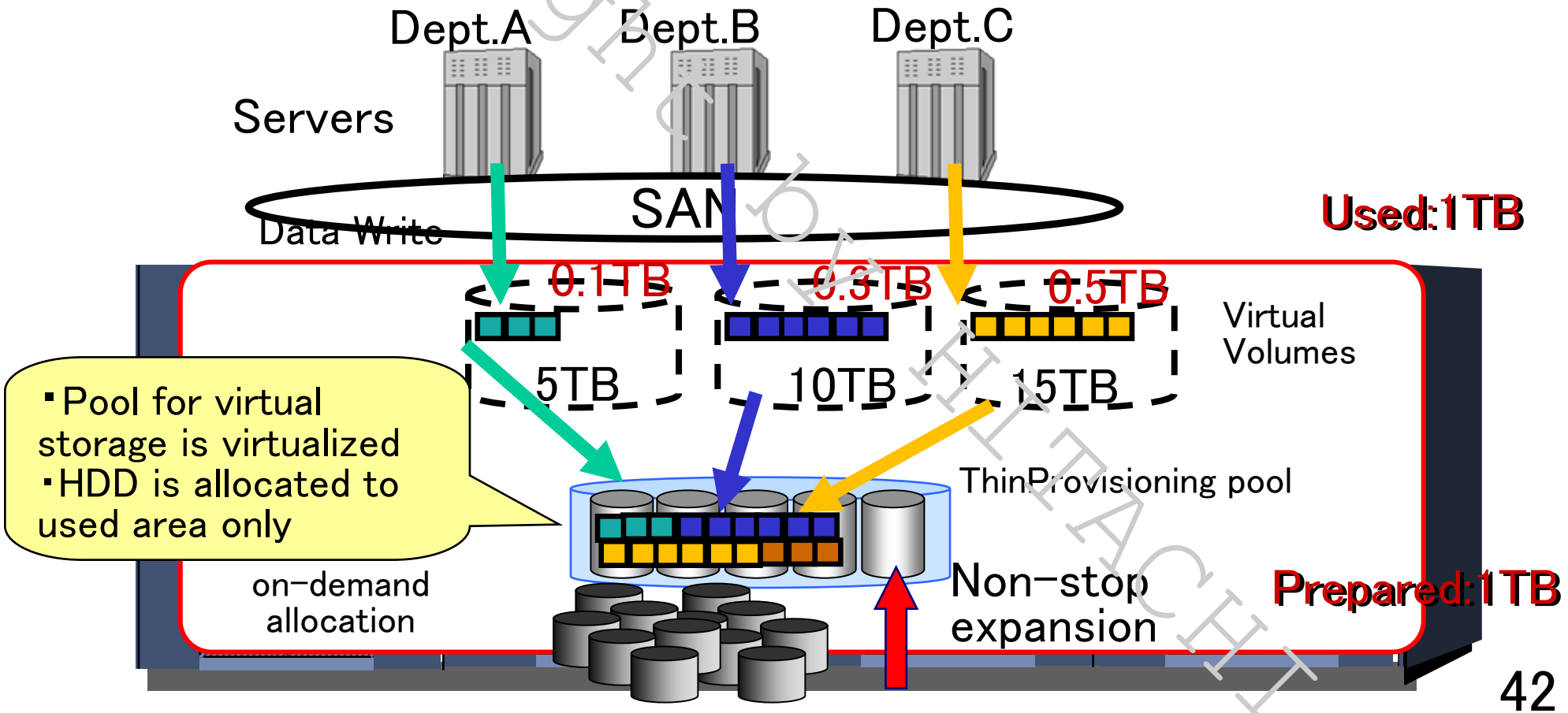
(Provisioning: Capacity planning for volumes)



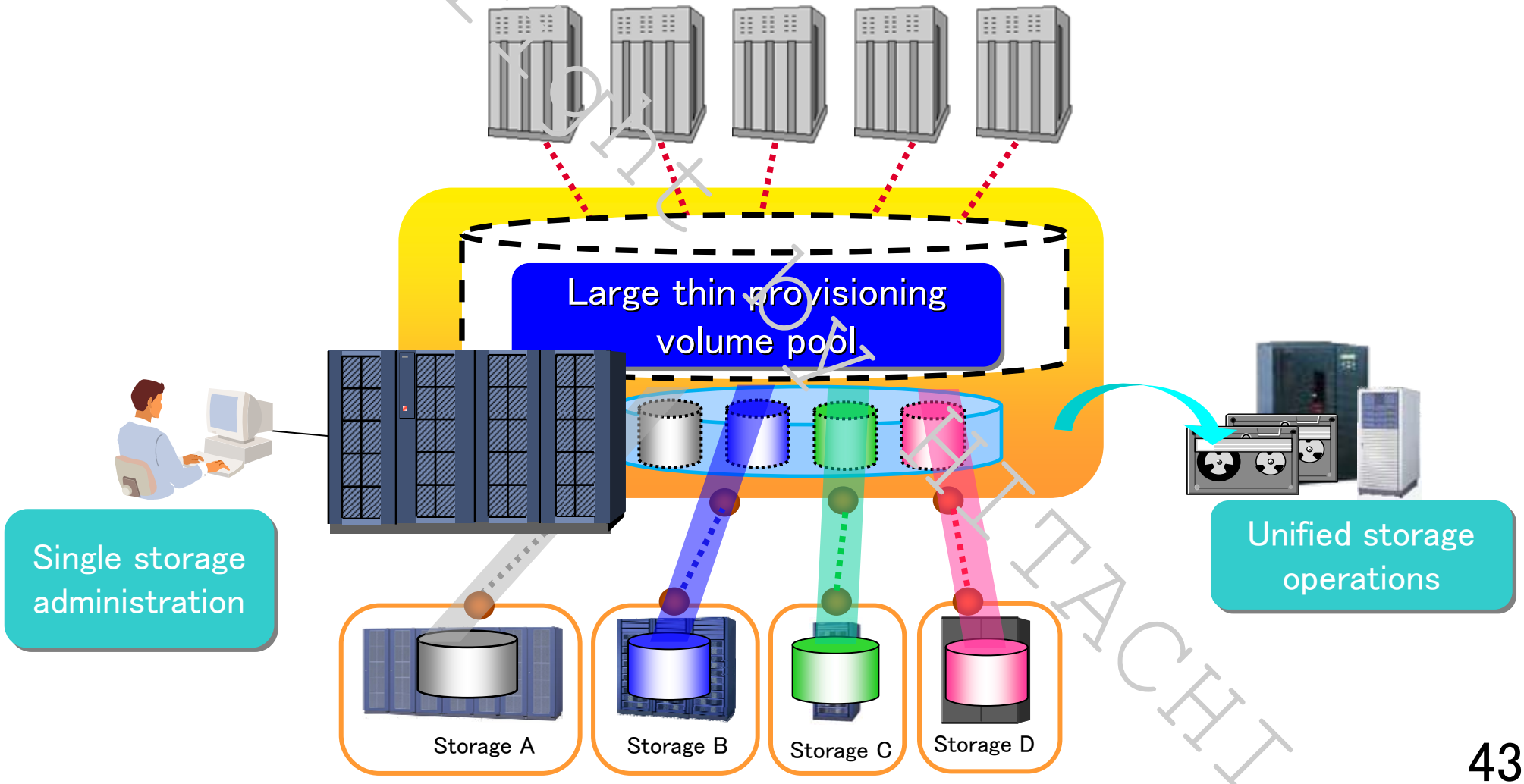
Conventional Storage

2-10 Solution (Thin Provisioning)

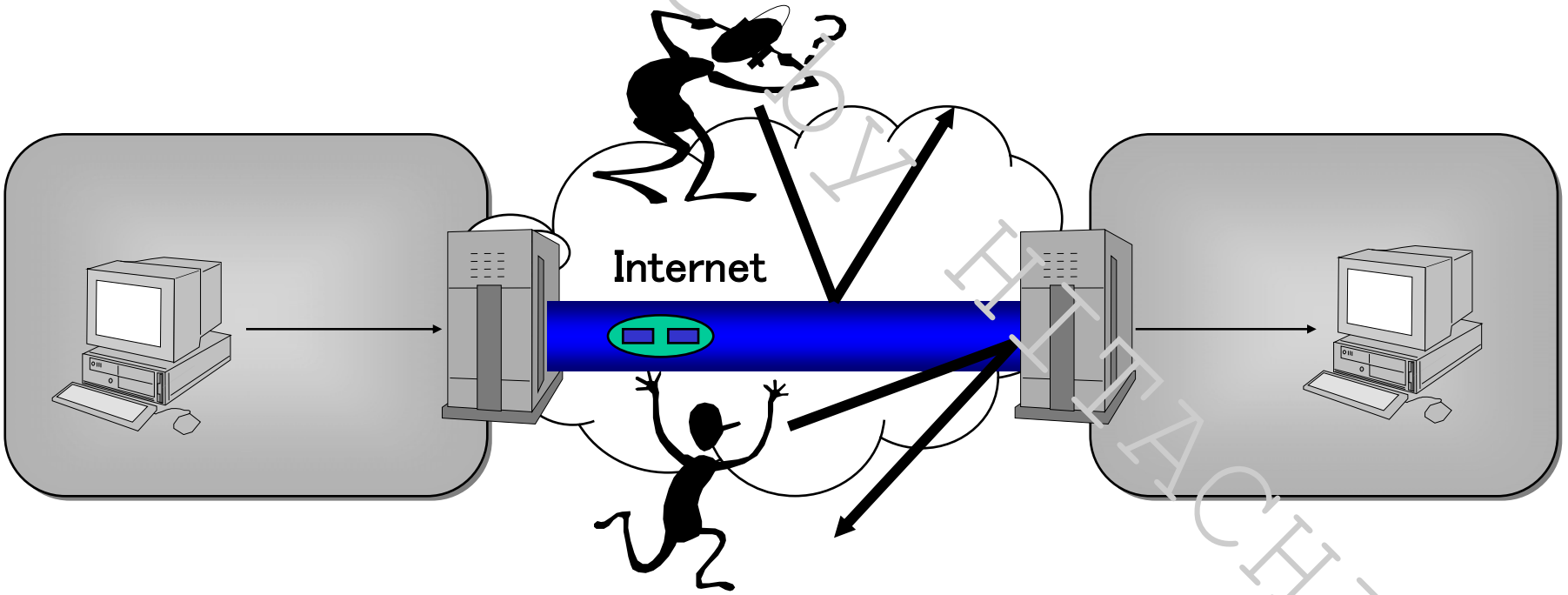
- Thin provisioning technology provides on-demand allocation for HDDs during data writes.
  - Pool for virtual storage is virtualized
  - Decreasing cost of HDDs due to using less HDDs
  - Non-stop expansion for Thin provisioning pool



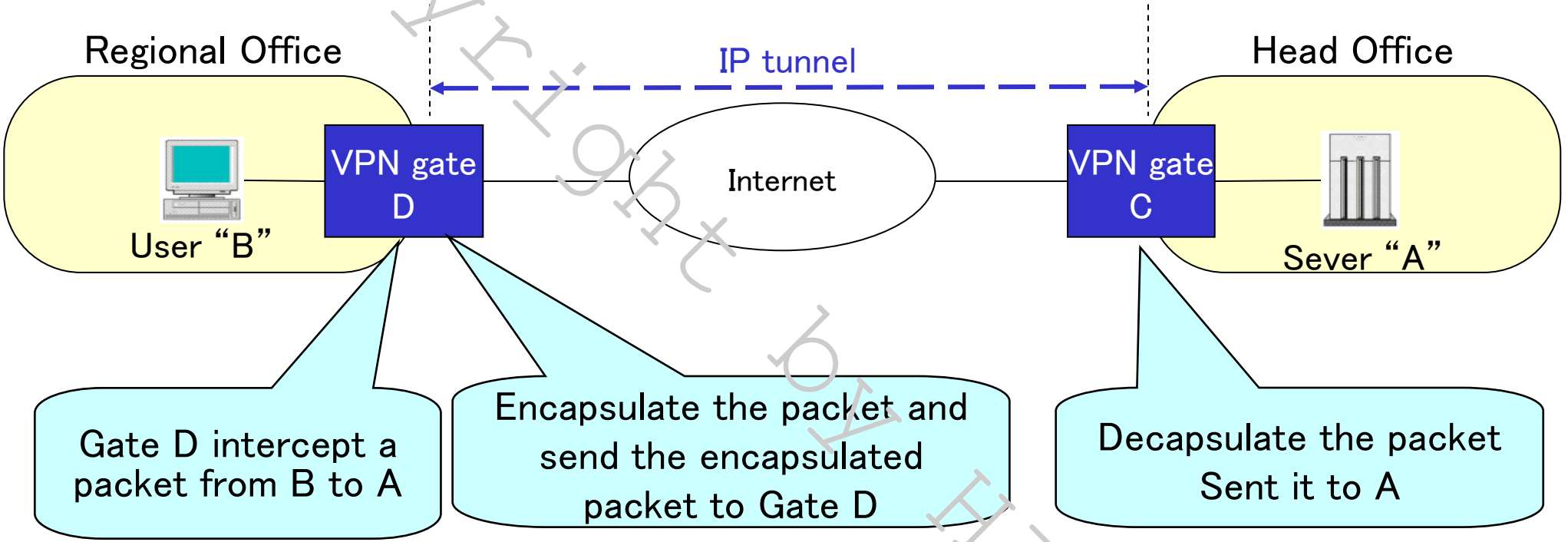
- Combination of storage virtualization and thin provisioning
  - Decreasing both operational cost and HDD cost

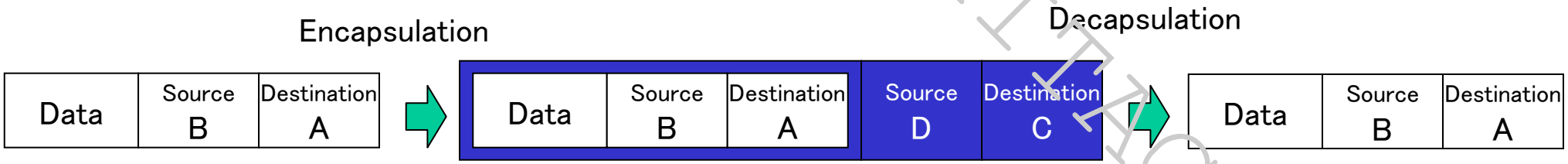


- VPN creates secure network as if it is a private network on a public network
  - Cryptograph data on a public network
  - Users unaware of using public network



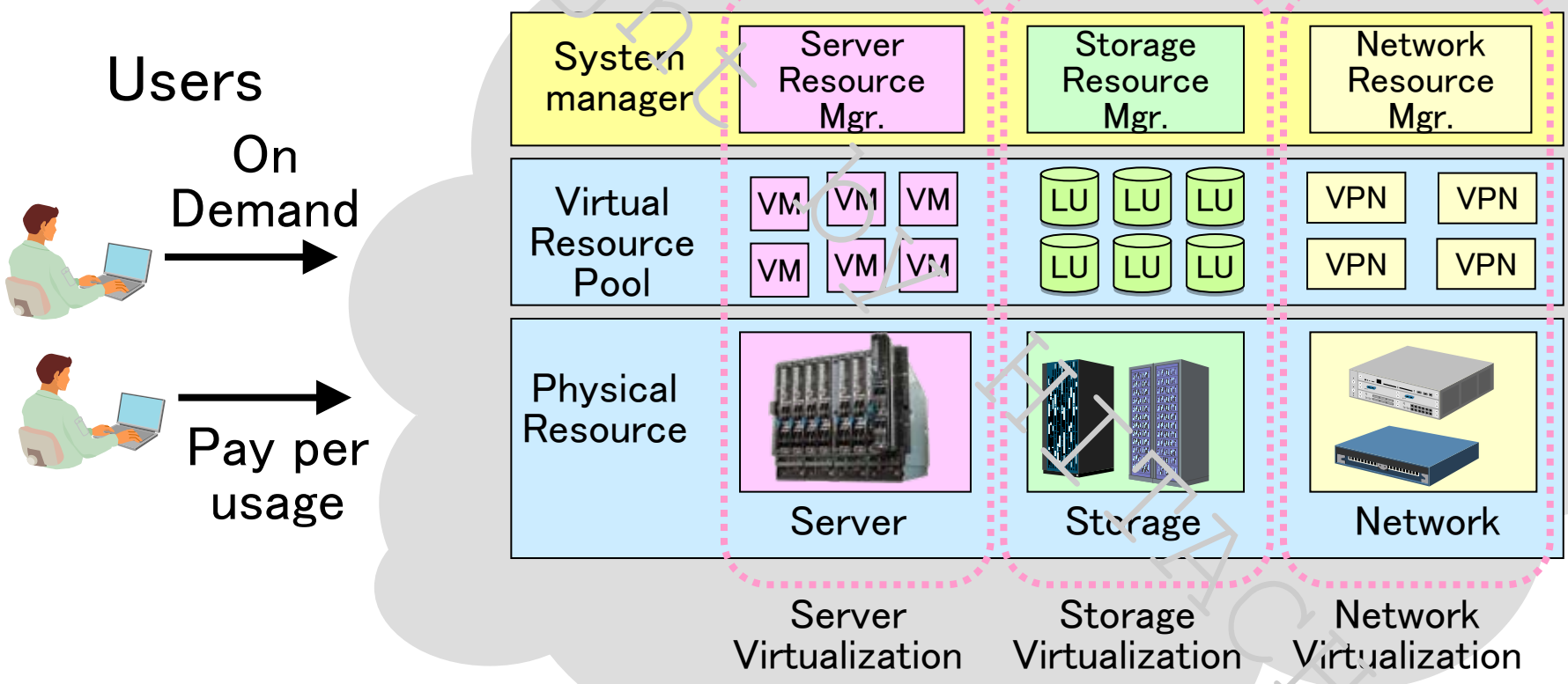
Tunneling





System Manager

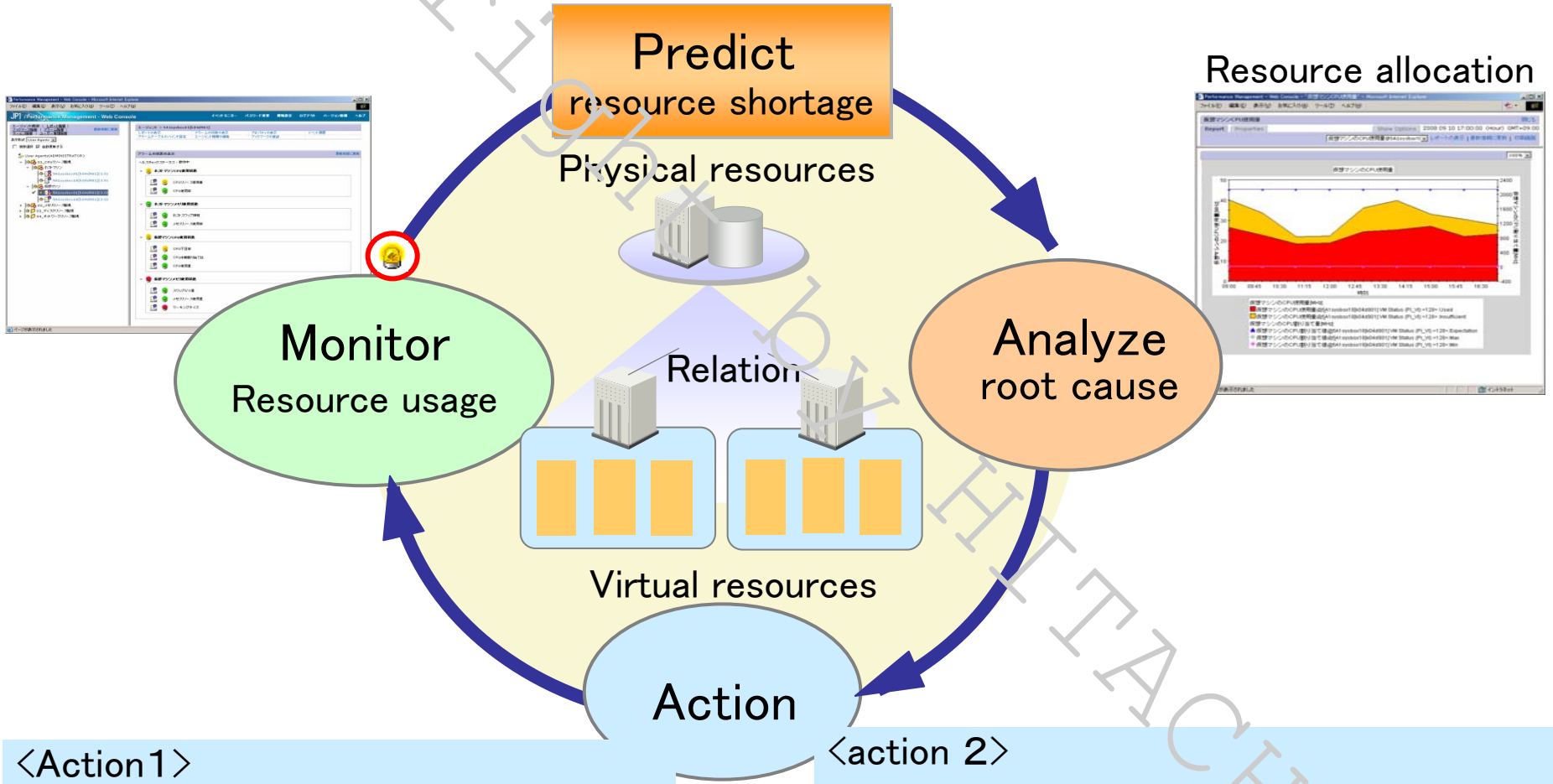
- System manager allocates virtualized resources in the pool to the users on the demand
- Voiding the resource shortage and effective allocation are required



VM: Virtual Machine
LU: Logical Volume
VPN: Virtual Private Network

System Manager Inside Cloud

- Repeat the management cycle (monitor, predict, analyze, action)
- Advantages of cloud computing are achieved by system manager



<Action1>
Modify resource allocation rate

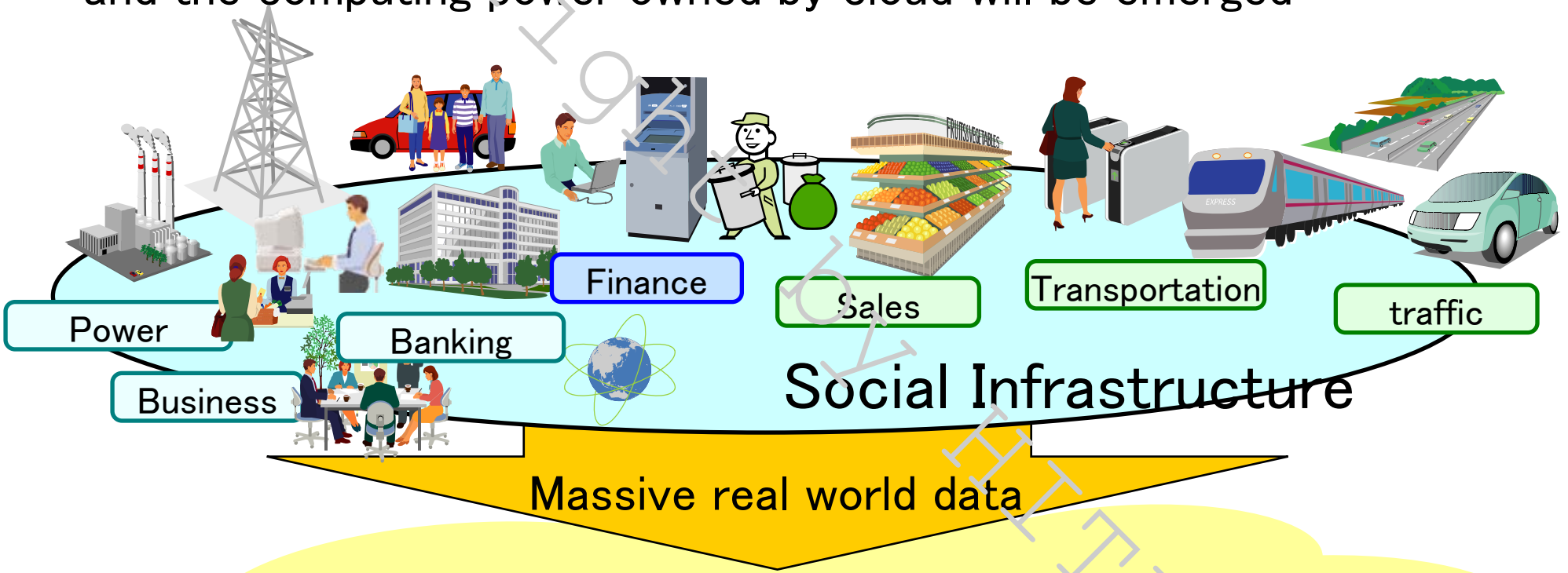
<action 2>
Migrate a virtual server to another physical server

3

Beyond Cloud Computing

3-1 Cloud era. has just started

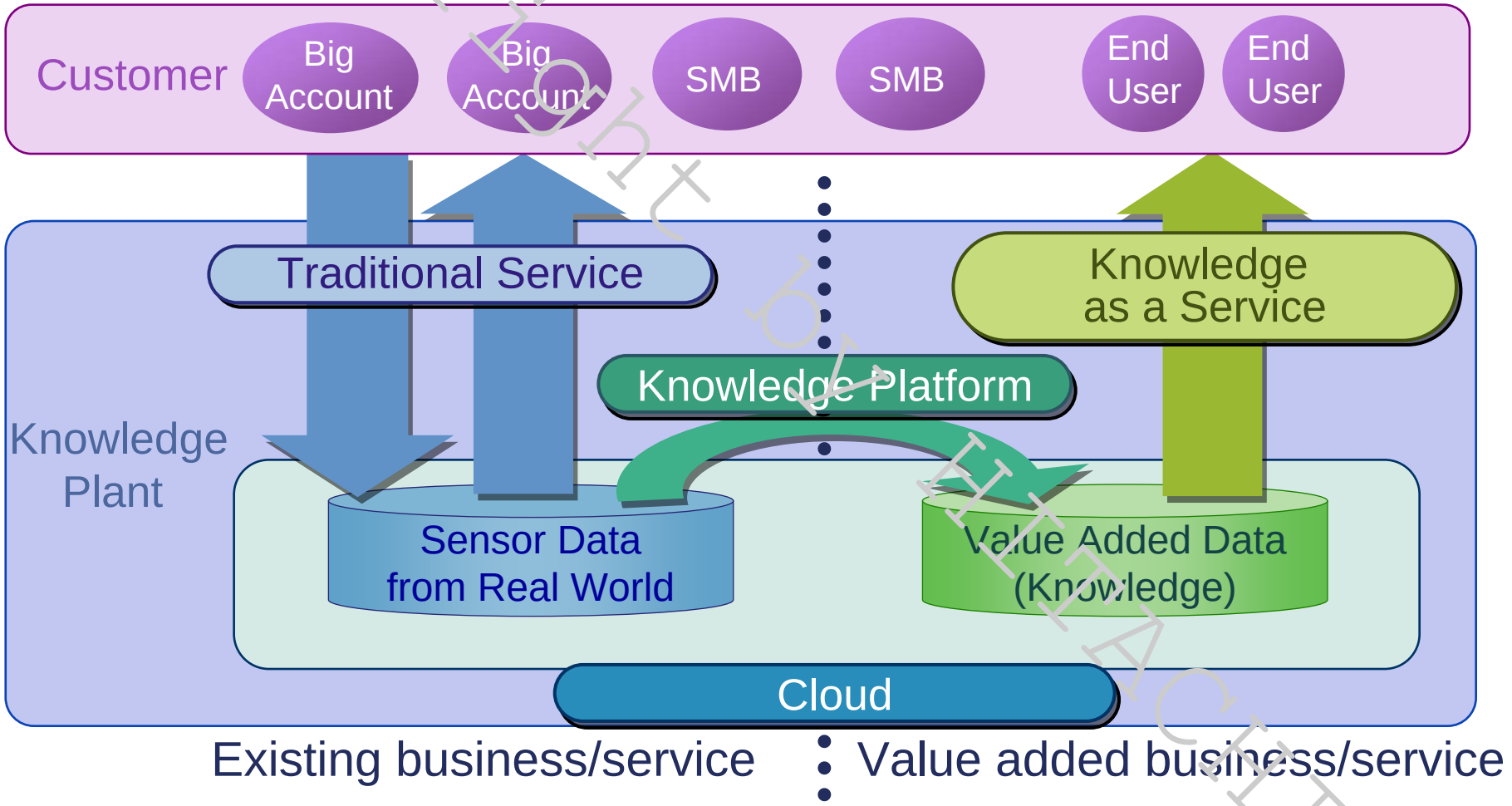
- As cloud computing becomes naturally used as social infrastructure, various types of data will be stored to the cloud
- "Knowledge"-empowered service using the massive data collected and the computing power owned by cloud will be emerged



KaaS: Knowledge as a Service
New business model for IT business to grow to in the cloud era.

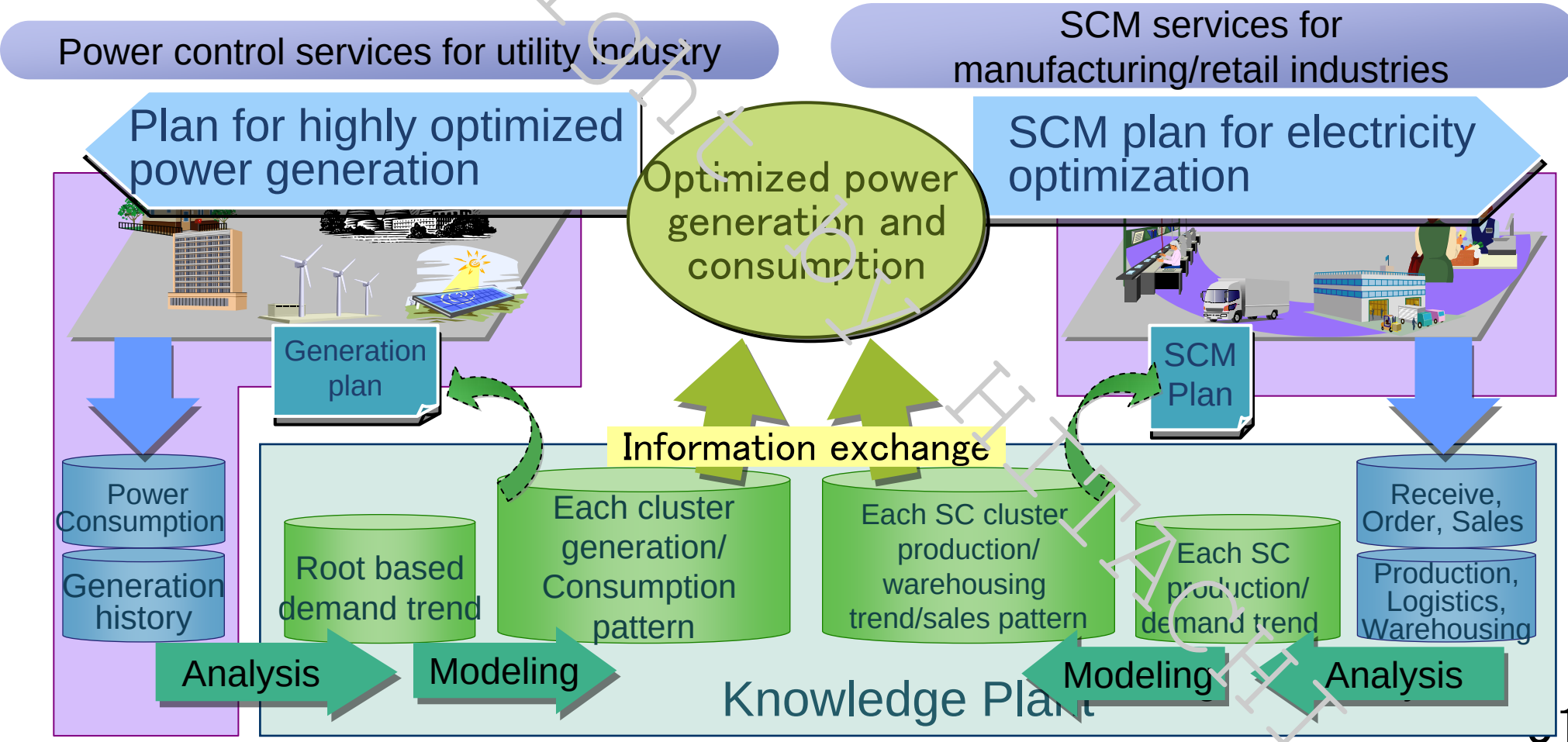
Cloud

- KaaS services extract and present valuable information(knowledge) from the large volumes of data generated by social infrastructure



3-3 KaaS Example (Knowledge Convergence)

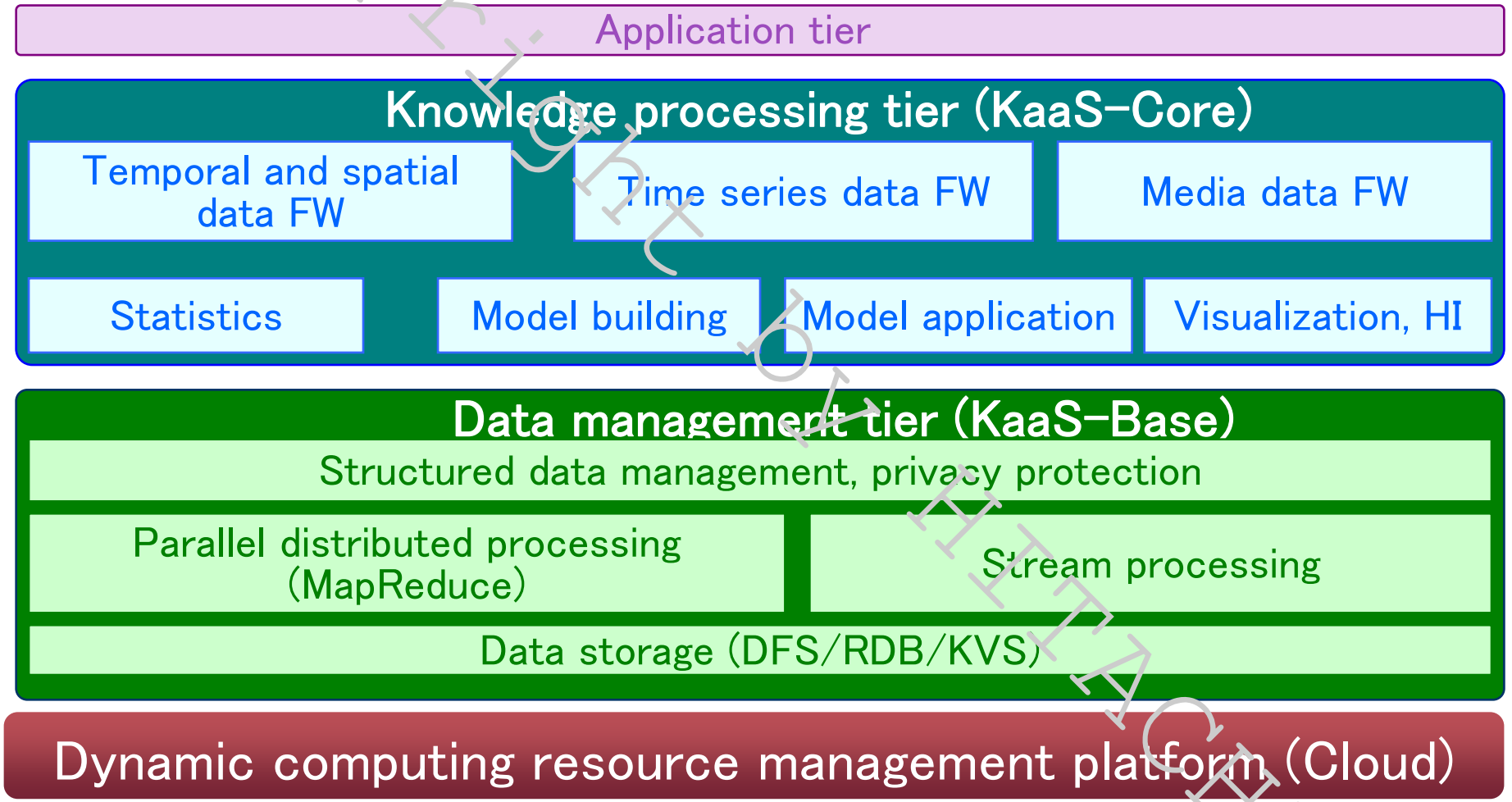
- The electric power optimization of whole society by the collaboration of power companies with manufacturing companies
- Kaas would help to optimize our social system totally by leveraging information beyond industries



SCM: Supply Chain Management

KaaS System Architecture

- Handle large volumes of data and perform complex computation
- Analyze the data from a range of different perspectives

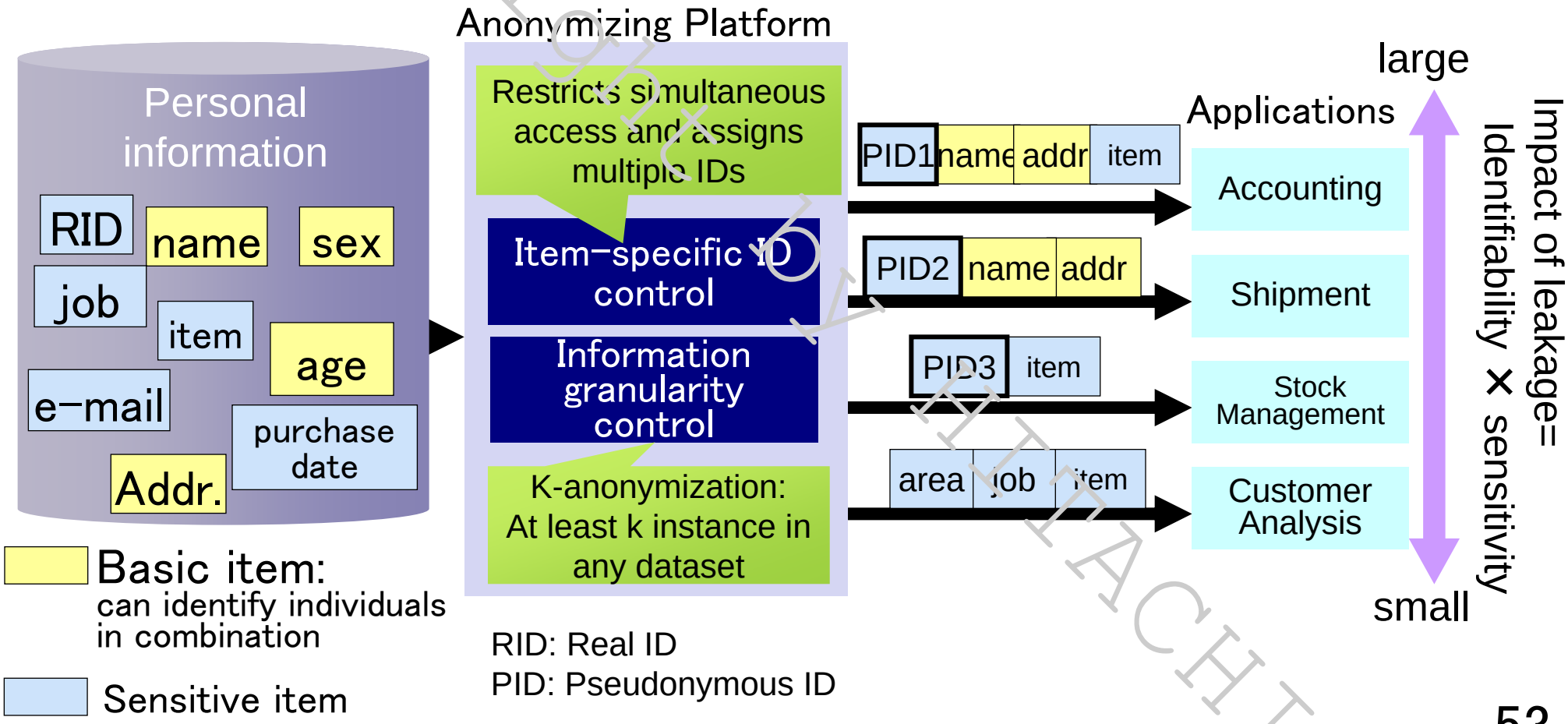


the integrated and efficient handling of various types of data

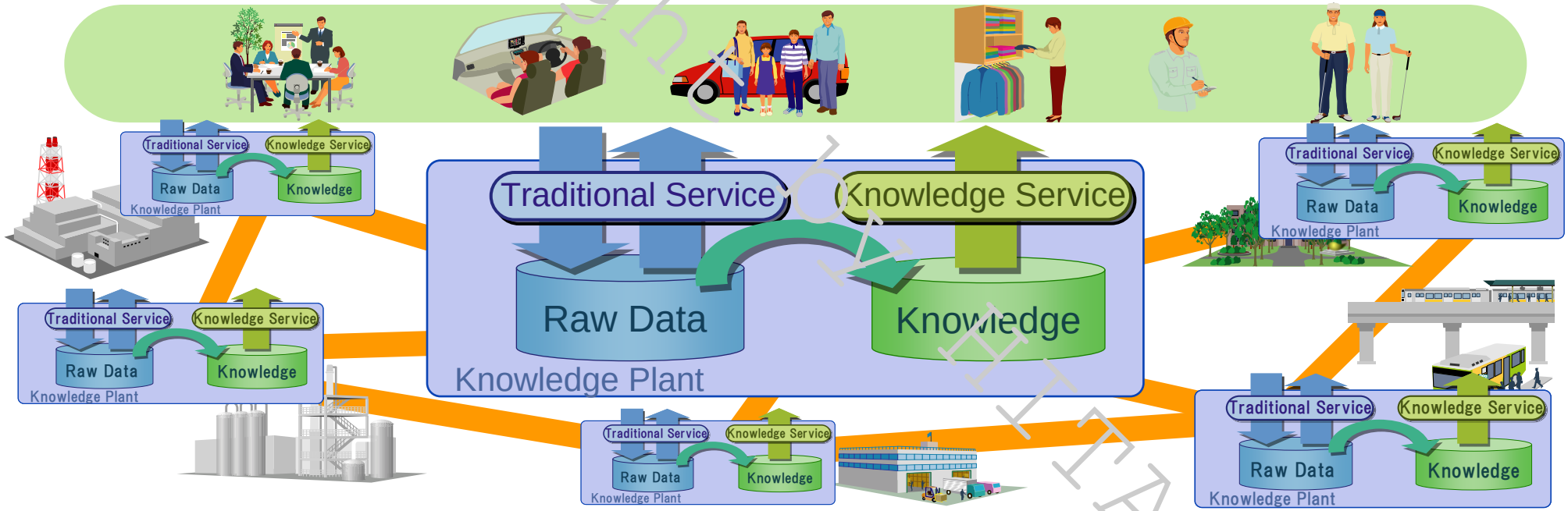
FW: framework, HI: human interface, DFS: distributed file system, RDB: Relational Database, KVS: Key Value Store

Anonymization for Personal Data

- Privacy protection is the one of most important issues in cloud
- Anonymizing makes it impossible to identify individuals' identities
- Control personal IDs according to identification risk of personal data and property of applications



- The knowledge-based society using Cloud Computing is coming
  - Based on the real world information, it supports resolution of today's global challenges (eg. Energy,environment,transportation).
  - Anyone can unconsciously obtain the benefits from the knowledge.



Kaas concept will help to optimize our social system totally

Conclusion

- Information Systems have been evolving for more than 60 years
- Cloud computing is a one of compilation of the researches and development on Information systems
- We believe that a new service “KaaS” contributes to innovation in social infrastructure