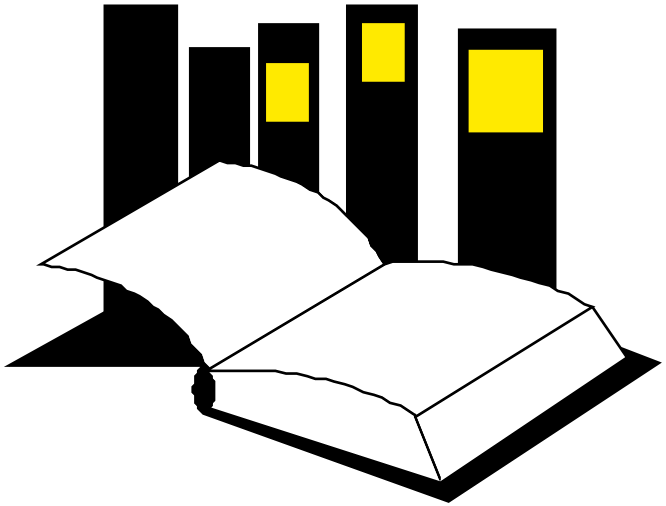
1. Course Overview

- Course Background
- Course Goal
- Course Content
- Course Assessment



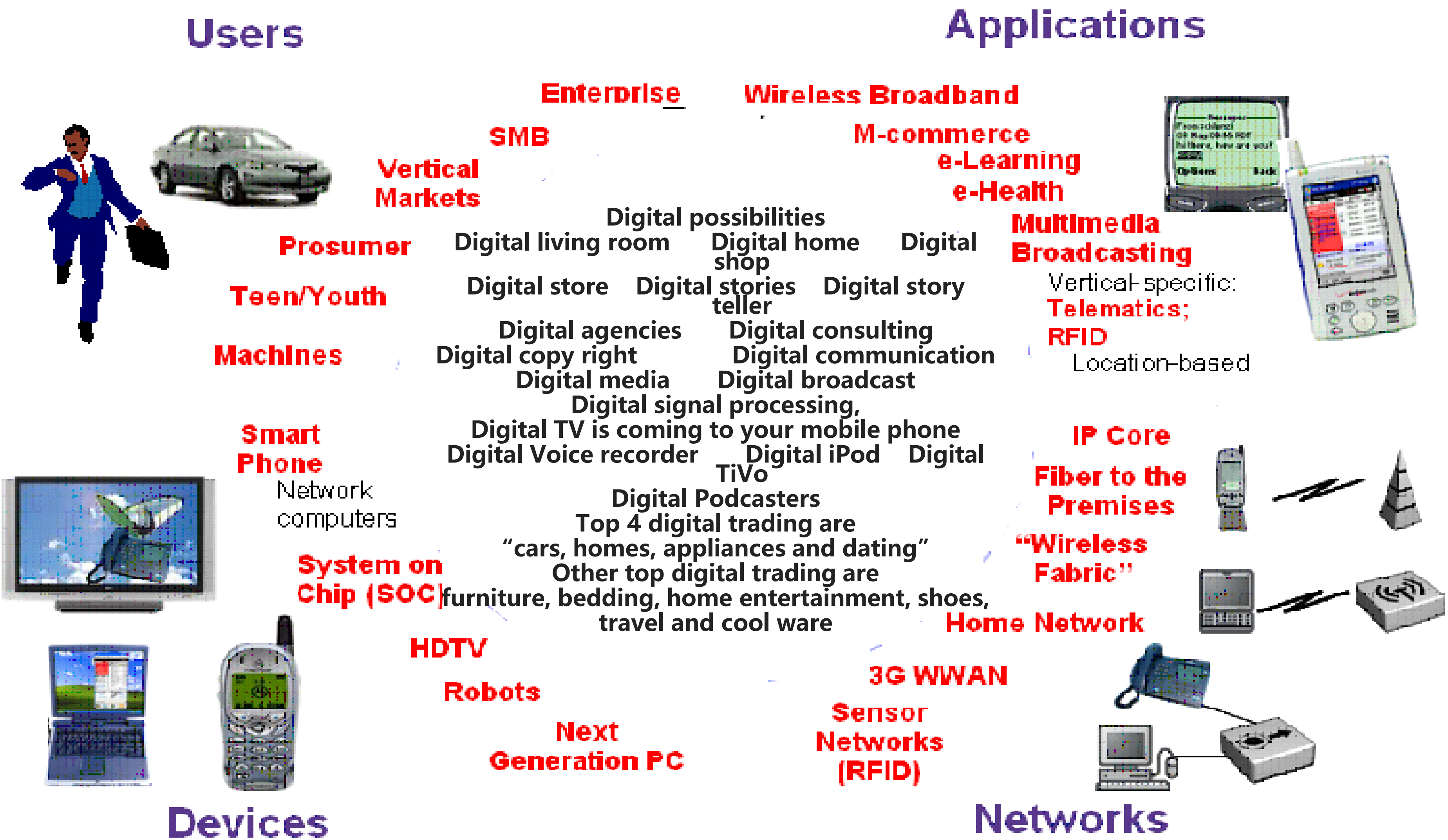
Distributed
Database Systems

Tsinghua University

1. Course Overview

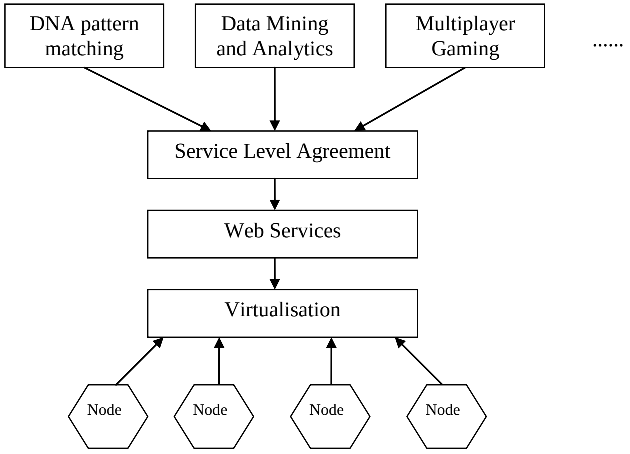
→ Course Background
- Course Goal
- Course Content
- Course Assessment

Digital Living



Cloud Computing

- Internet-based Super Computing Paradigm
  - A large amount of information and processor resources stored on personal computers, mobile phones and other devices are put together to work together.
  - Powerful computing and storage capabilities
  - Service, Platform, Utility

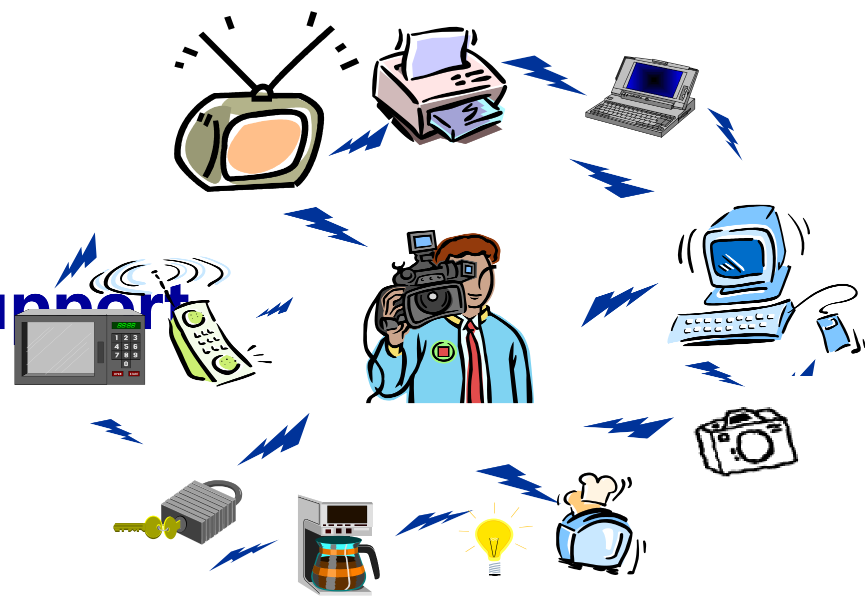


Ambient Intelligence

- Ambient intelligent environments
  - sensitive to people's needs
  - personalized to their requests
  - anticipatory of their behavior
  - responsive to their presence

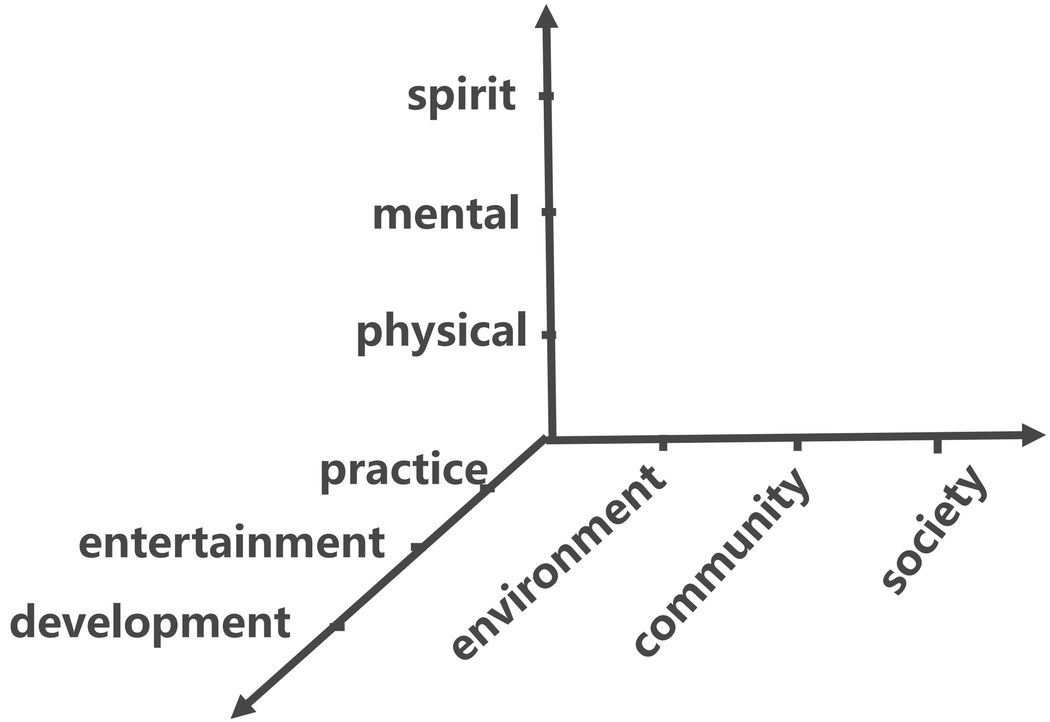
- Emphasis
  - greater user-friendliness
  - more efficient services support
  - user-empowerment

Improve Quality of Life!



Quality of Life

存在 (Being) (who am I?)

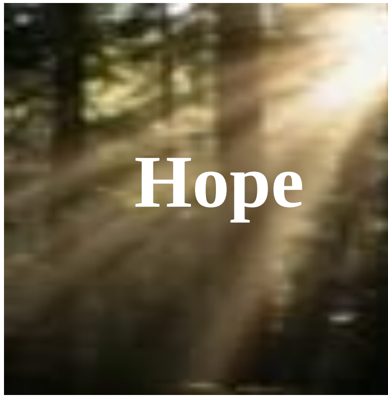


属于(Belonging)
(environment and community)

成为 (Becoming)
(dream and value)

Ambient Environment

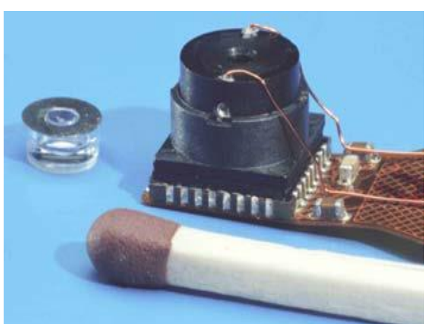
- Satisfy basic needs
- Provide opportunities
- Choice and initiative right



Technology
3D Moore

More Moore

Large Area Electronics

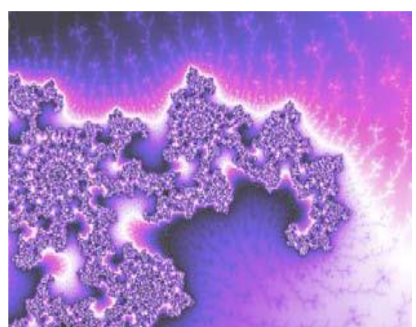


Systems in Package

Economy
3D Business

Globalization

Experience Economy



Creative Industry

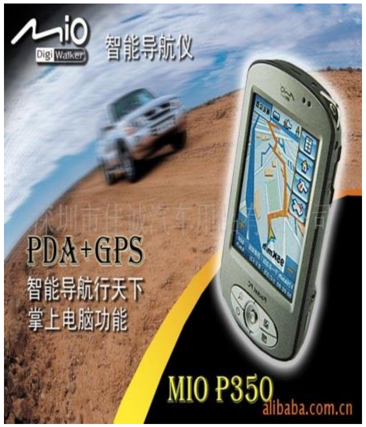
Society
3D People

Freedom

Expression

Care

Smart Living



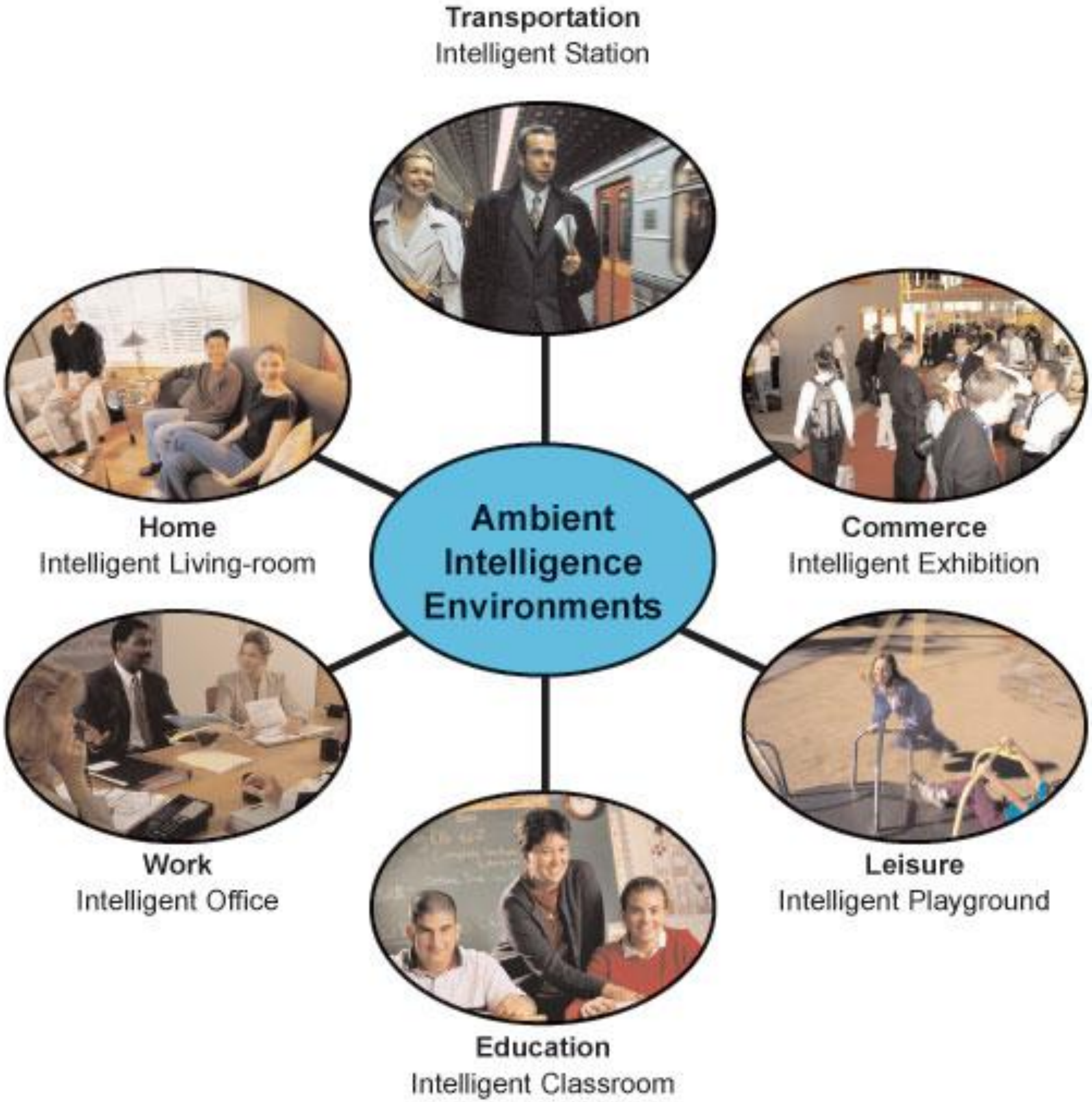
(From Philips)

Smart Kitchen

- The kitchen desktop is a projection display that can recognize food, suggest what to cook, and play tutorial videos
- The table can weigh, time, and cook
- The waste bin can scan and identify the materials. The recyclable garbage is crushed, sorted, and disinfected. After that, it is vacuum-sealed and labeled for future use.
- The sink has two holes, one side keeps the relatively clean water for watering flowers and flushing the toilet; the other side throws away "black water" that cannot be used any more.

Smart Refrigerator

- The refrigerator knows the food inside. When the inventory is reduced to a certain extent, the refrigerator orders food from the online store.
- The refrigerator has an electronic touch screen, which can display the calendar, weather, schedule, and phone number when the phone rings.


Transportation
Intelligent Station
Home
Intelligent Living-room
Ambient Intelligence Environments
Commerce
Intelligent Exhibition
Work
Intelligent Office
Leisure
Intelligent Playground
Education
Intelligent Classroom

From Internet to Internet of Things

- **The Internet of Things (IoT) is the network of physical objects or "things" embedded with electronics, software, sensors, and network connectivity, which enables these objects to collect and exchar**

Man-Machine → Object-Object Communication

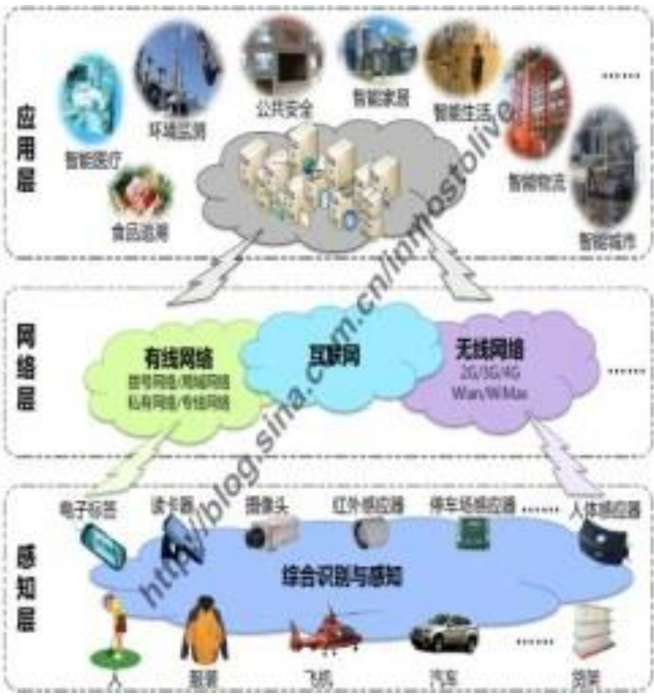
- **All objects, from tires to toothbrushes, from houses to paper towels, can communicate actively through the Internet.**
  - **Clothes "tell" the washing machine's requirements for color and water temperature.**
  - **Vehicle tires communicate with each other about road conditions to avoid traffic jams.**
  - **The fan blades activate the deicing maintenance equipment for intelligent operation and maintenance.**

Vision of IoT

- **Traditional Era**
  - Separate physical infrastructure and IT information infrastructure
  - airports, highways, and buildings vs data centers, personal computers, broadband, etc.
- **IoT Era**
  - Concrete facilities and cables are integrated with chips and broadband into a unified infrastructure
  - Such infrastructure is more like a new earth construction site on which the operation of the world is carried out, including economic management, production operation, social management, and even personal life

IoT → Social Media

Machine / Sensor



Social Media

Passive information receivers → Active creators

Create and Consume Contents

Web Clickstream

DOC / Media

Call Log

Apps

Create value by making connections

Data generated in one day

200万篇博客文章
在网上发布。

每天约有2940亿封
电子邮件发出。

1亿8700万小时的音乐
在Pandora（流媒体音乐网站）播放。

2200万小时
用来在Netflix（在线影片租赁商）
上观看以前的电视节目和电影。

5亿3200万条
状态更新。

1288个新应用
可供下载
超过3500万个应用被下载。

2亿5000万张照片
上传到Facebook。
如果把它们都印出来，
叠起来能有80个
埃菲尔铁塔那么高。

一天内，整个因特网的流量信息可以装满
1亿6800万张DVD光盘。

86万4000小时的视频
上传到YouTube上。

Data generated in one day in China

淘宝网 Taobao.com

Over 50000 GB of data generated per day

Over 40 million GB data stored

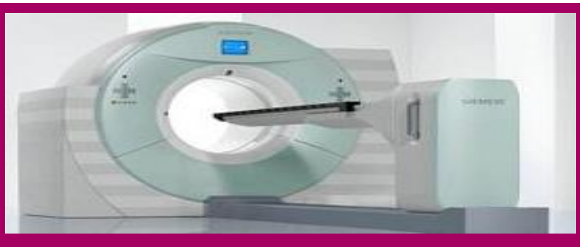
Baidu 百度

Over 1 trillion pages

Processing over 6 billion search requests

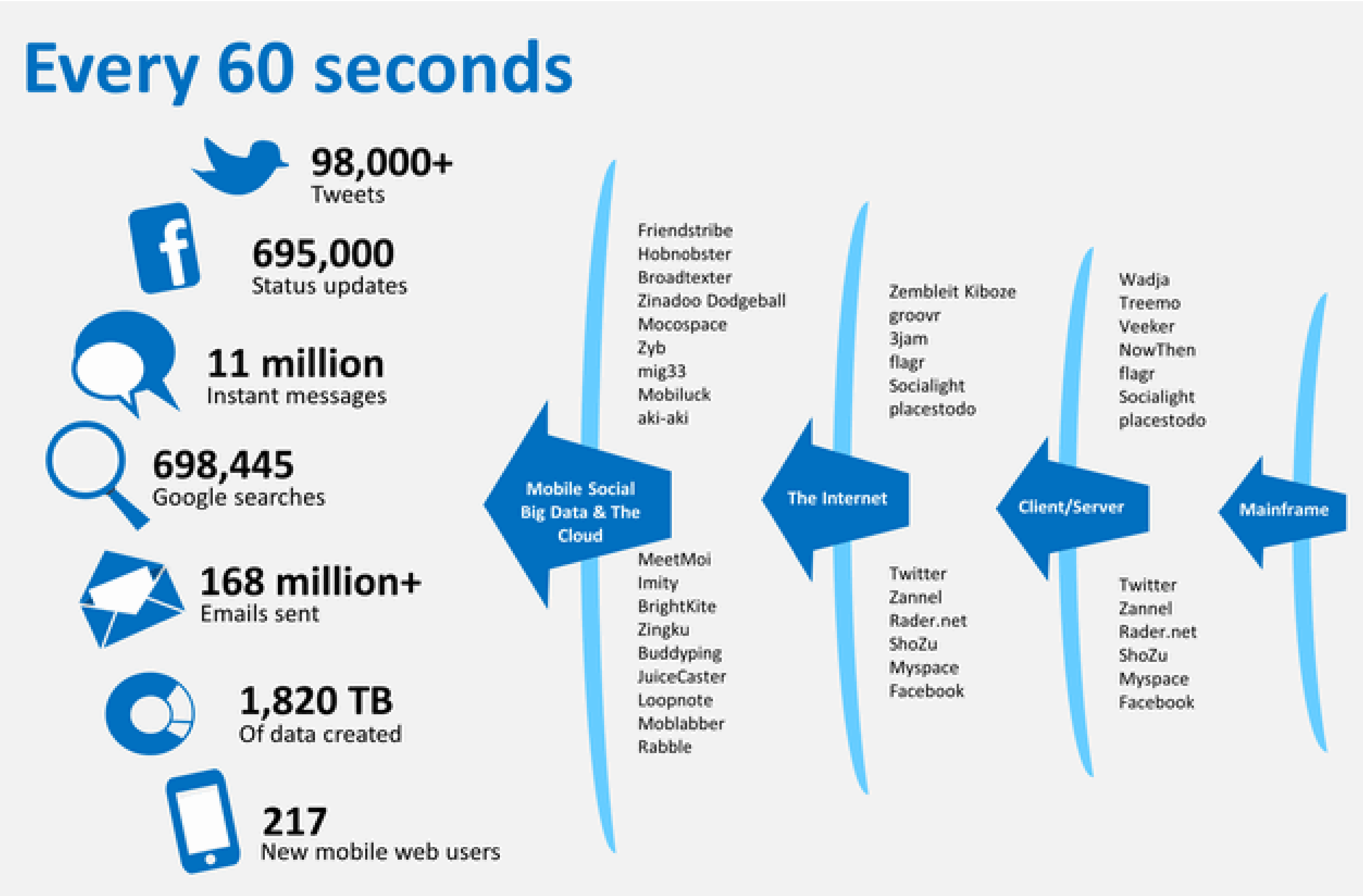
The CT image data of a patient is tens of GB

The data that need to be saved every year is up to 10 billions of gigabytes of data



Generating over 3.6 GB of data per hour

A city produces tens of millions of gigabytes of data every month

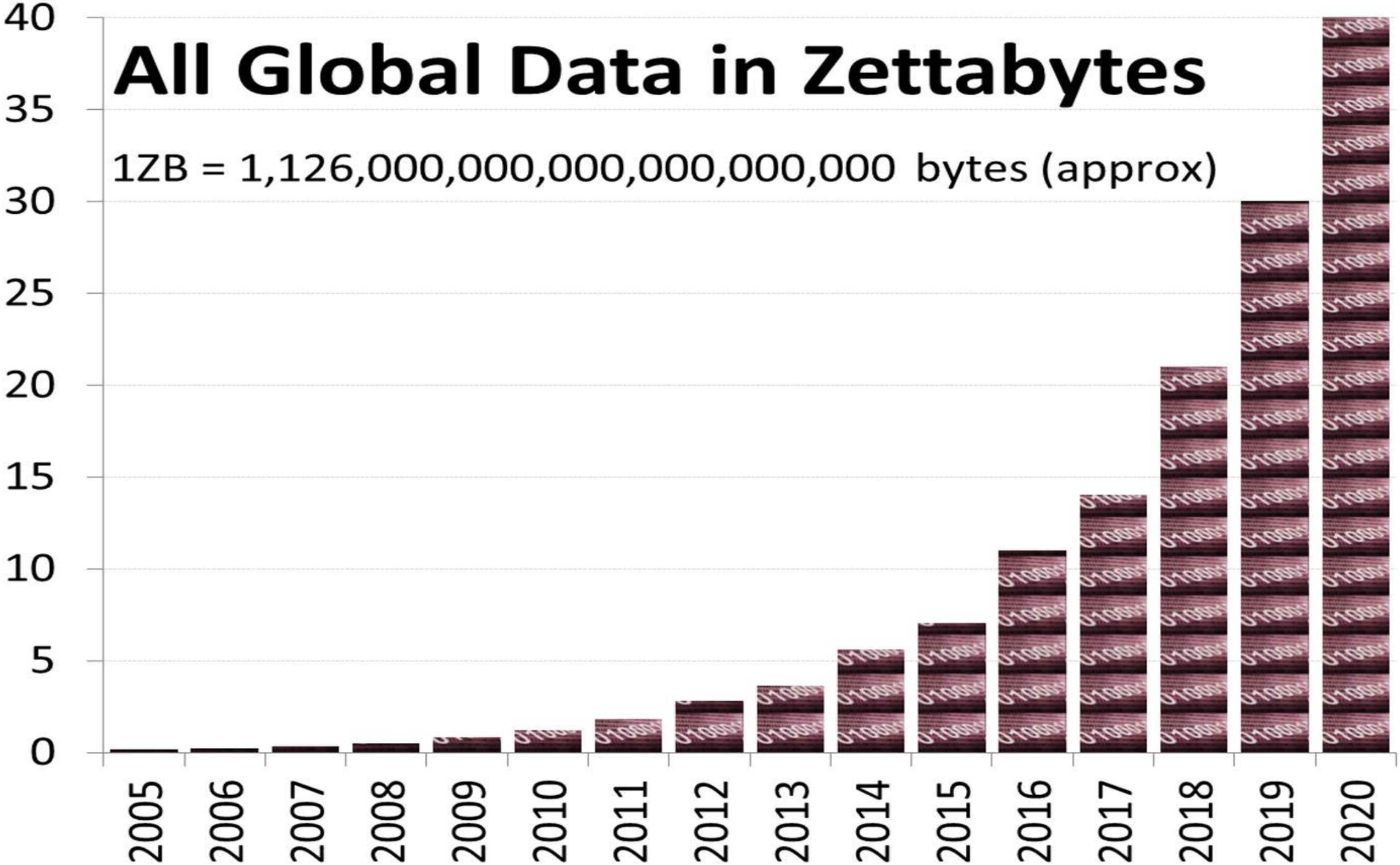

Every 60 seconds
98,000+
Tweets
695,000
Status updates
11 million
Instant messages
698,445
Google searches
168 million+
Emails sent
1,820 TB
Of data created
217
New mobile web users
Friendstribe
Hobnobster
Broadtexter
Zinadoo Dodgeball
Mocospace
Zyb
mig33
Mobiluck
aki-aki
Mobile Social Big Data & The Cloud
MeetMoi
Imity
BrightKite
Zingku
Buddyping
JuiceCaster
Loopnote
Moblabber
Rabble
Zembleit Kiboze
groovr
3jam
flagr
Socialight
placestodo
The Internet
Twitter
Zannel
Rader.net
ShoZu
Myspace
Facebook
Wadja
Treemo
Veeker
NowThen
flagr
Socialight
placestodo
Client/Server
Twitter
Zannel
Rader.net
ShoZu
Myspace
Facebook
Mainframe

IoT– Fan Sensing Data

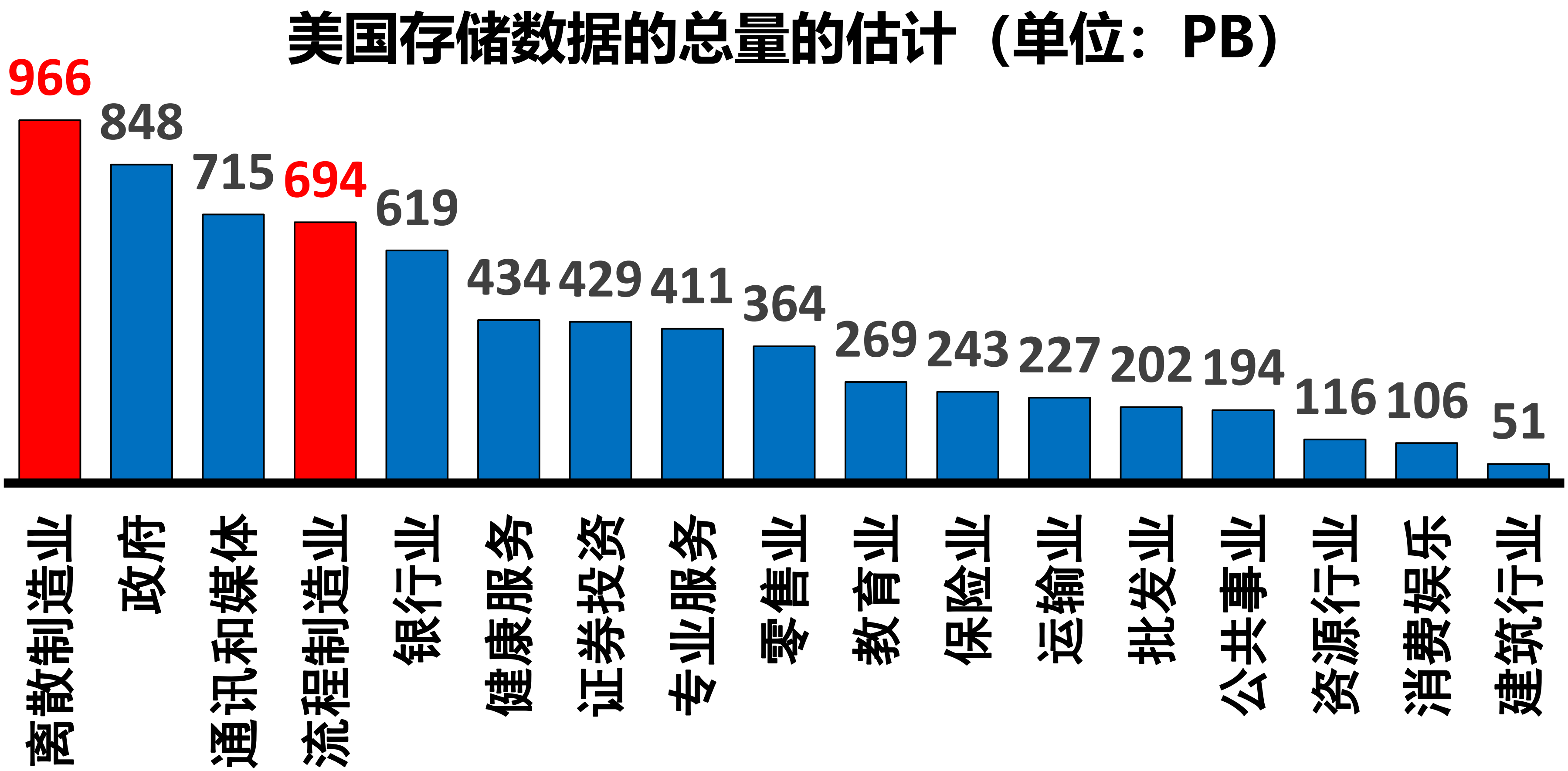
- 500 sensing points in a fan
- Fault detection requires millisecond acquisition frequency
- One fan generates thousands to tens of thousands of data points per second

20000+ fans, acquisition frequency 20ms~7s

300 Billons data records per day
7*24*365

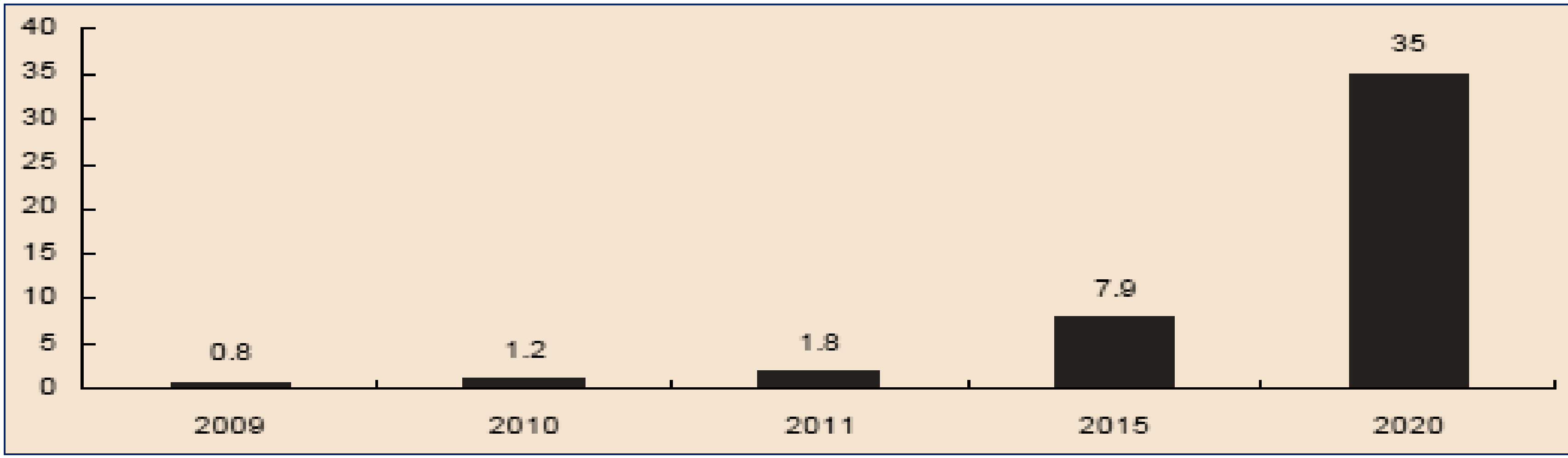

All Global Data in Zettabytes
1ZB = 1,126,000,000,000,000,000,000 bytes (approx)
0
5
10
15
20
25
30
35
40
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020

Most of the data generated by the manufacturing industry



资料来源：Big data: The next frontier for innovation, competition, and productivity (麦肯锡)

Global Data Amount



- Entering the TB era in 2006, approximately 180 EB of data generated worldwide.
- Reached 1.8 ZB in 2011.
- In 2020, the total amount of data in the world will increase nearly 44 times to 35.2 ZB.

Chinese Masterpiece《Dream of Red Mansion 红楼梦》

has about 870,000 Chinese characters

1 Chinese character = 2*8 bits = 2 bytes

1GB ≈ 671 masterpieces

1TB ≈ 631,903 masterpieces

1PB ≈ 647,068,911 masterpieces

1EB = 4000 times of the amount of information stored in the library of Congress in USA

1Byte = 8 Bit

1KB = 1,024 Bytes

1MB = 1,024 KB = 1,048,576 Bytes

1GB = 1,024 MB = 1,048,576 KB = 1,073,741,824 Bytes

1TB = 1,024 GB = 1,048,576 MB = 1,099,511,627,776 Bytes

1PB = 1,024 TB = 1,048,576 GB =1,125,899,906,842,624 Bytes

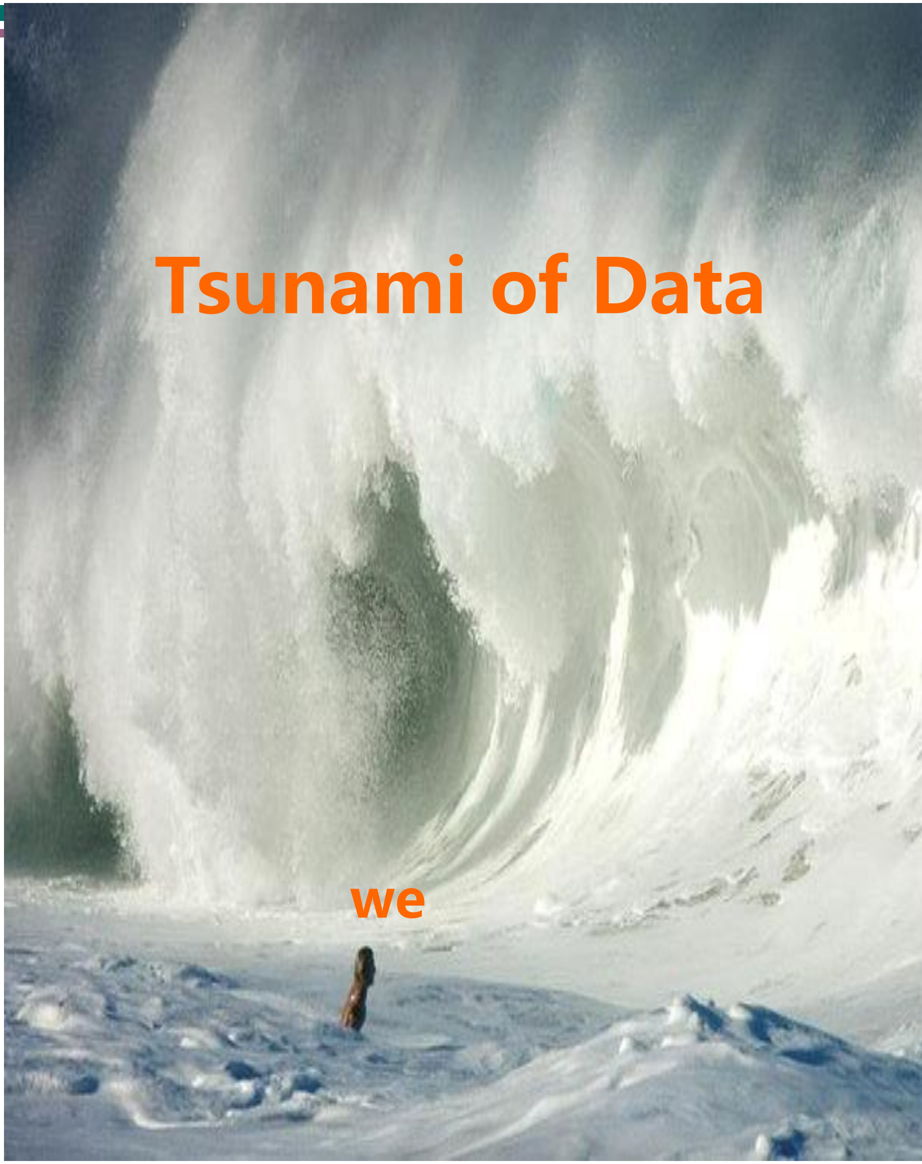
1EB = 1,024 PB = 1,048,576 TB = 1,152,921,504,606,846,976 Bytes

1ZB = 1,024 EB = 1,180,591,620,717,411,303,424 Bytes

1YB = 1,024 ZB = 1,208,925,819,614,629,174,706,176 Bytes

Entering Big Data Era

Nature（2008年9月3日）



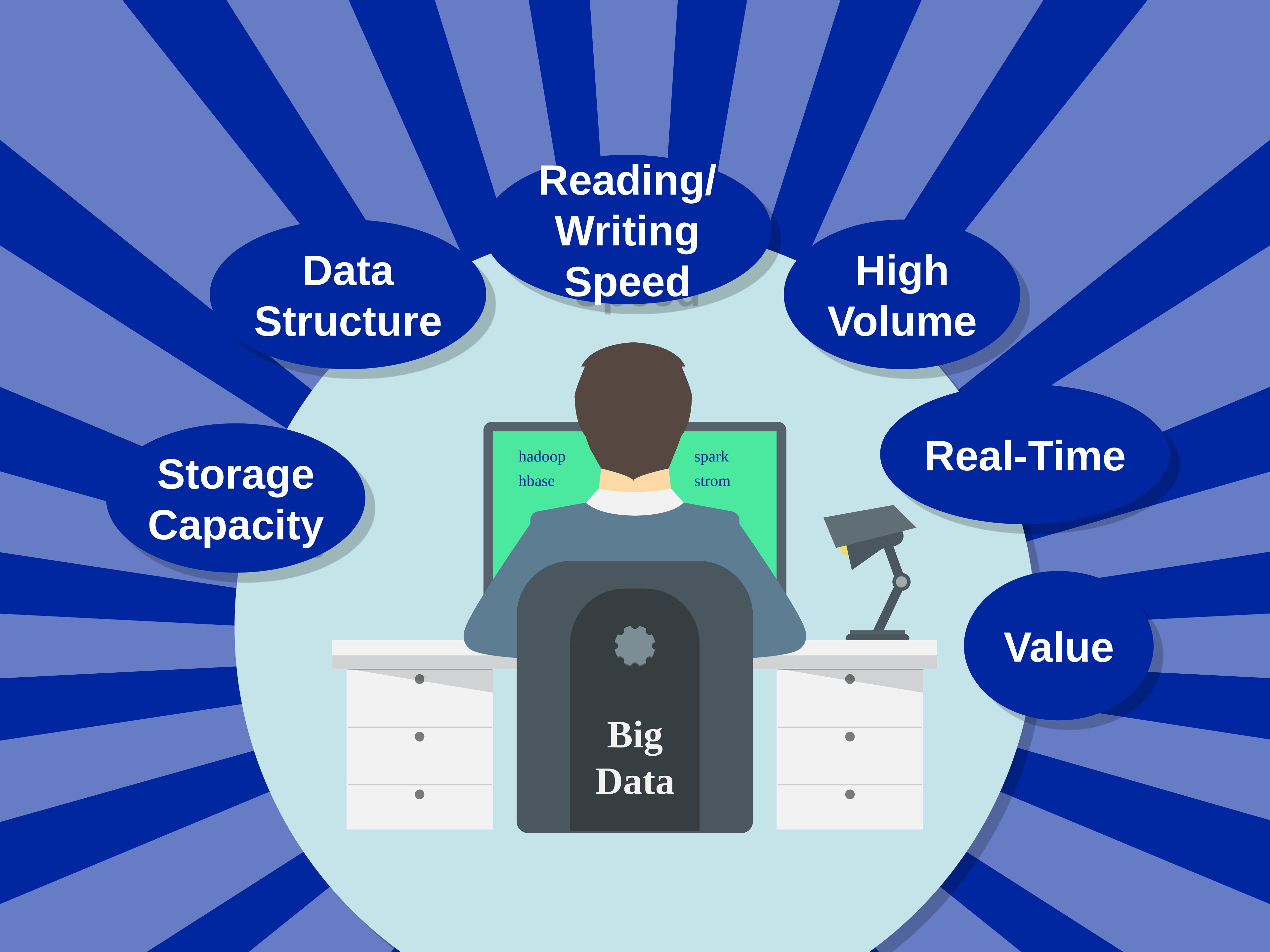




Definition of Big Data

Big data **is an all-encompassing term for any collection of *data sets* so *large and complex* that it becomes *difficult* to process using *traditional* data processing applications.**

-- http://en.wikipedia.org/wiki/Big_data


Reading/
Writing
Speed
Data
Structure
High
Volume
Storage
Capacity
Real-Time
Value
hadoop
hbase
spark
strom
Big
Data

High Storage Capability

The amount of data continues to increase, and storage capabilities continue to increase accordingly.

High Reading/Writing Speed

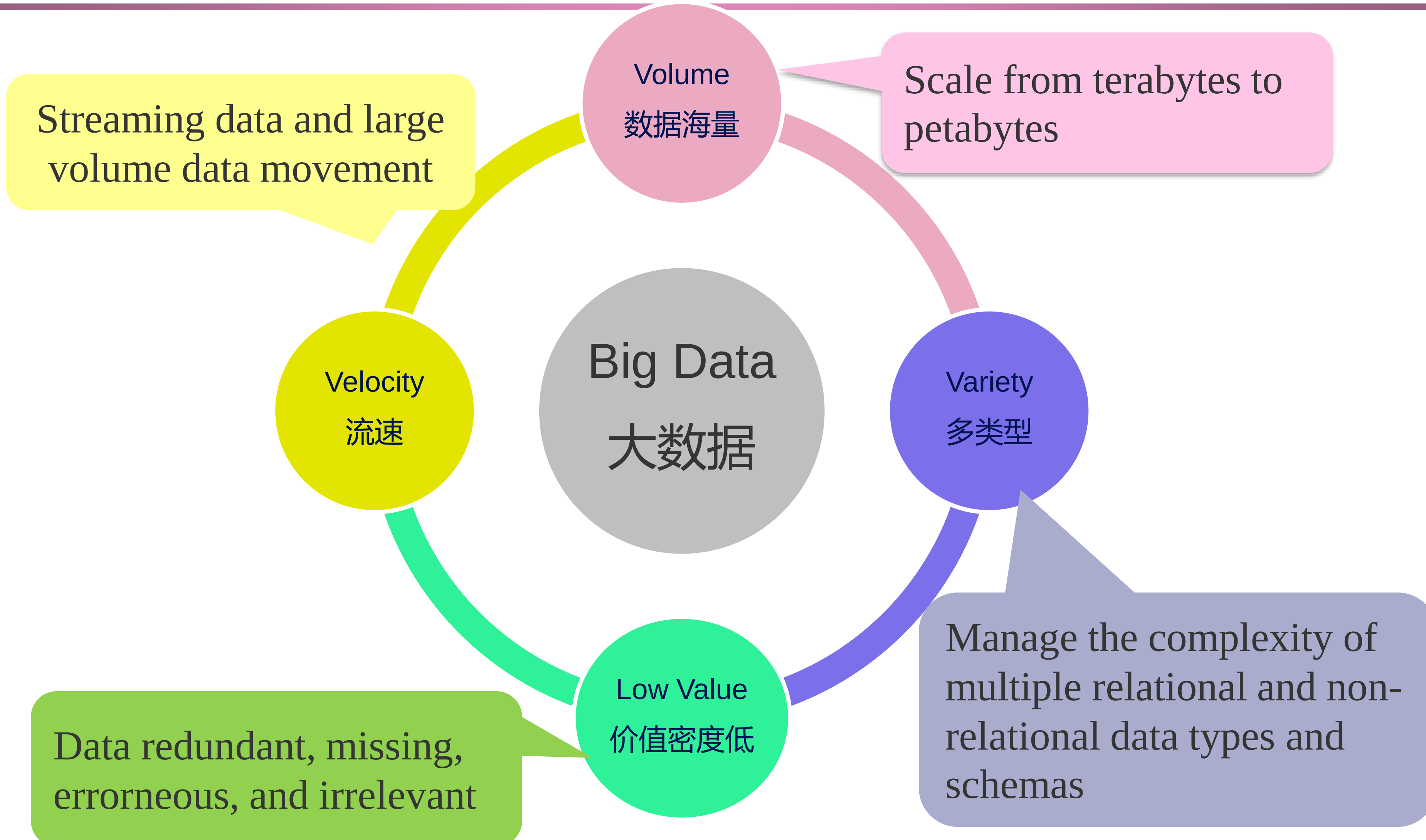
From the huge amount of data, the speed of reading and writing the required data is significantly improved.

Complex and Variety

video, text, image, search history, shopping records, etc.

The data structure is more complex, and develops in diversity.

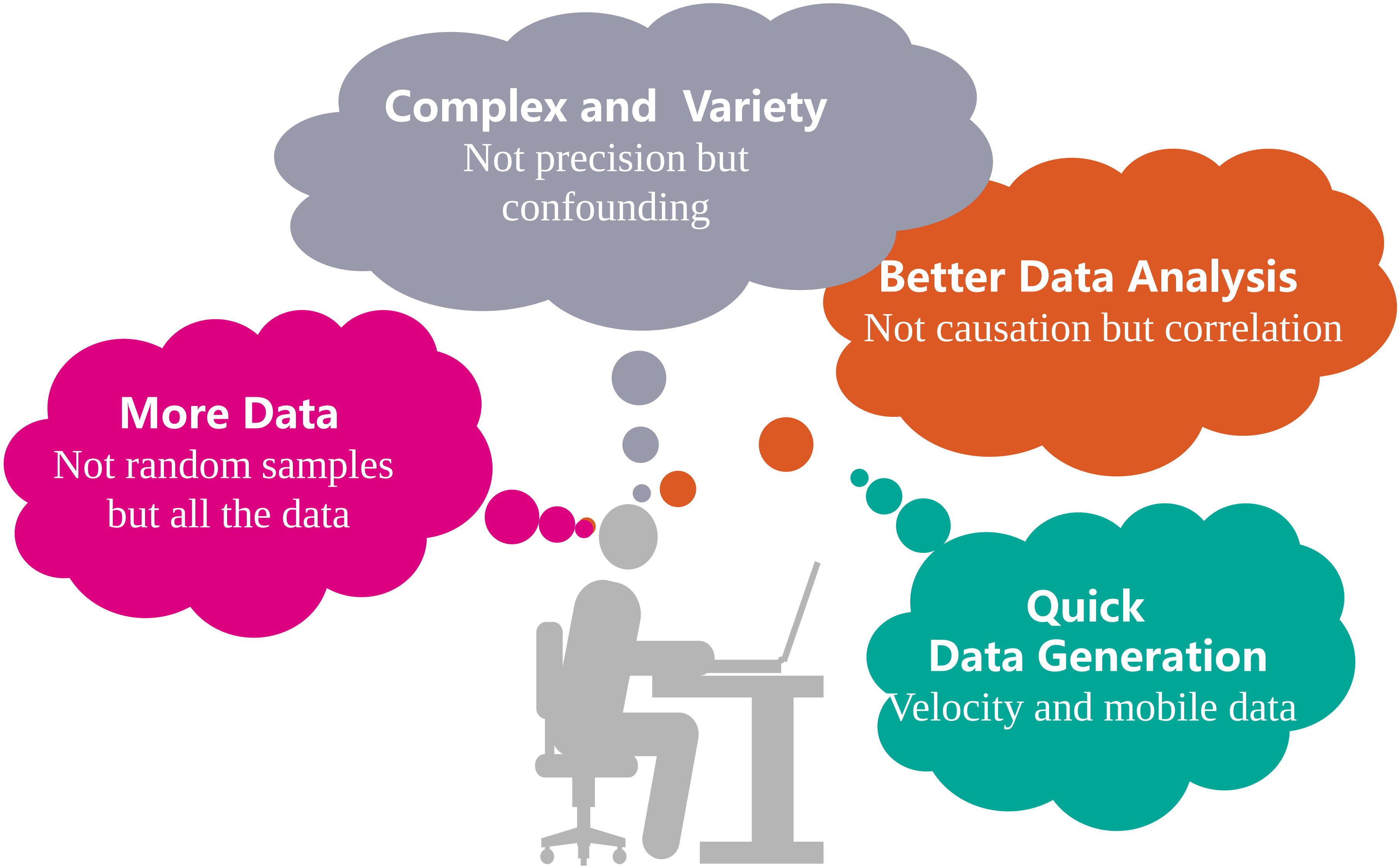
Characteristics of Big Data



Big Data – New Data，New View

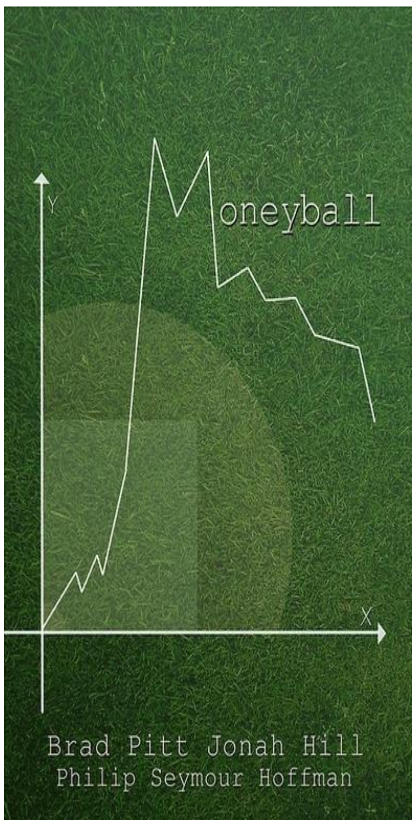
- Big Data
  - old problems, problems that existed when computers were born, and problems to be faced for a long time
- Conceptually,
  - more emphasis on data-based decision making
- Technically,
  - collecting all the data (traditional, existing and unused, external)
  - Through these data, new ways and perspectives can be generated

Thinking change brought by big data



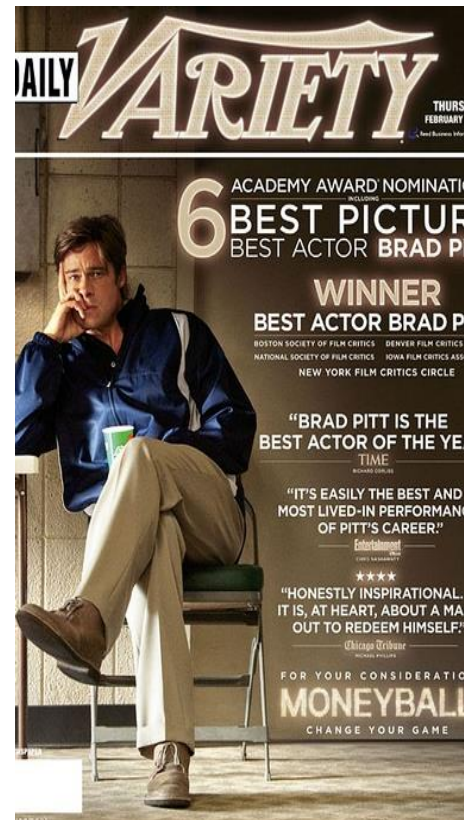
Film《Penalty Kick 点球成金》

Brad Pitt's "Penalty Kick" is an American Oscar-winning film. It tells the story of the baseball team general manager Pitt who uses data analysis to make a radical transformation of the team. The small team can achieve great success.



Based on historical data, use data modeling to quantitatively analyze the characteristics of different players, match them reasonably, and regroup.

Break with traditional thinking, find the most "cost-effective" players by analyzing game data, and use data to achieve success.



Commercial Value of Big Data---Business decision

A certain store sells milk. Through data analysis, it knows that customers who bought milk in the store often go to another store to buy buns. Then the store can consider cooperating with a bun store or directly selling buns in the store.

Commercial Value of Big Data---Personalized marketing

A customer searches mortgage interest rates on the Web. The system recommends the mortgage products. If the customer is really interested in this, the sales department will send the promotion information to the customer. If the customer goes to the bank branch to do business, The business staff will introduce the mortgage products in detail. At the beginning, there are only a few clues, but through multi-channel interaction with the customer, bank's precise and considerate services can be delivered to the customer.

Wind Power 's Prediction Accuracy

GE imagination at work

Vestas

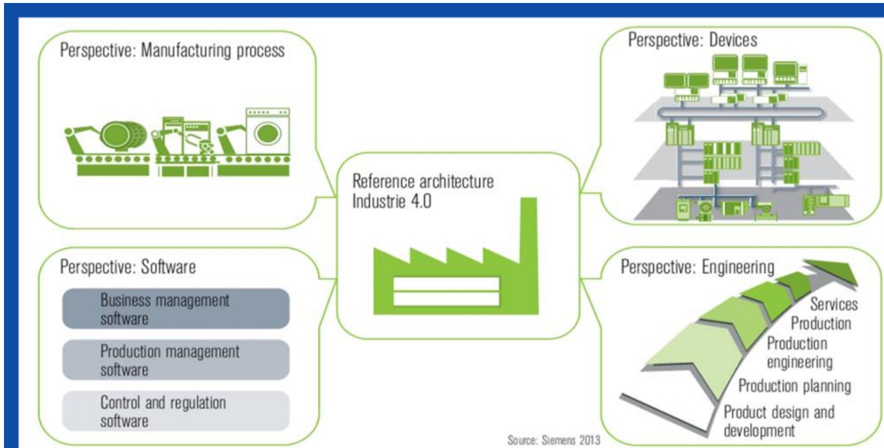
VS

金风科技
GOLDWIND

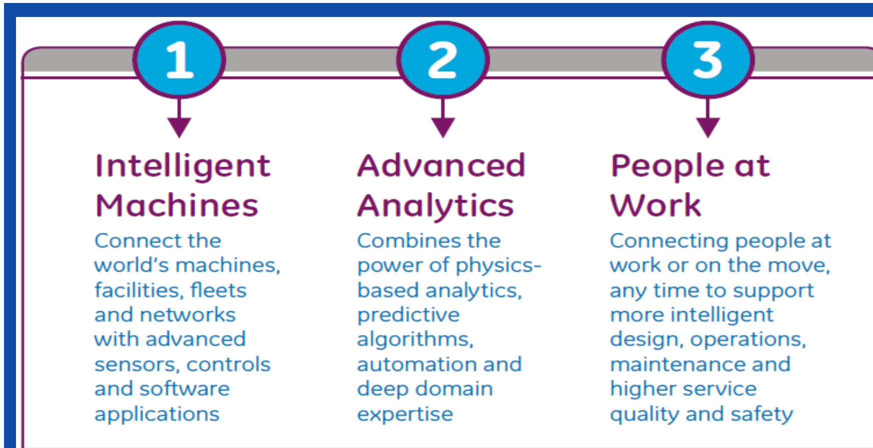
95% | **85%**

- **Big data is an important resource to improve productivity, competitiveness, and innovation.**
- **Mining users' "invisible needs" through big data, and creating more value for users often has low cost and high return.**

Big data has become the core element of manufacturing intelligence



德国工业4.0



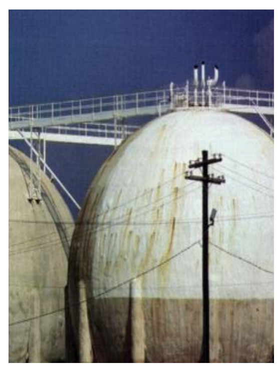
美国工业互联网

"Industry 4.0 is intelligent manufacturing driven by big data." (IBM)
"Big data is the core of Industry 4.0." (Dean of German Leipzig Business School)

"GE was still a manufacturing company yesterday, and has become a software and data company upon waking up." (GE Global Chairman)

Big data is the "new oil of the future"

Data is no longer a "by-product" of social production, but a raw material that can be processed twice or even many times, from which greater value can be explored. Data becomes the means of production.



Mined for only 162 years

Mined for only 45 years

Mined for only 60 years

Non-renewable resources

Data

Data Overload → Lack of Insight

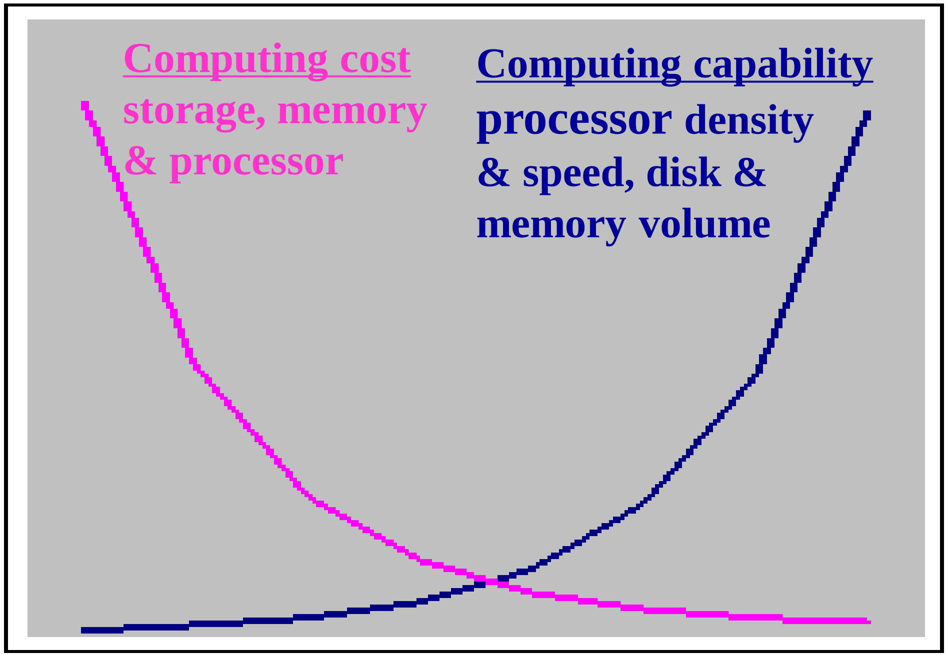
1/2 say they feel overwhelmed by the amount of data their companies manage

1/3 business leaders make decision based on information they don't trust, or don't have.

60% say they need to do a better job capturing and understanding information rapidly.

83% cite "BI & Analytics" as part of their visionary plans to enhance competitiveness.

Moore Law 's Exception



Ubiquitous Computing and Communication
Large Volume of Data

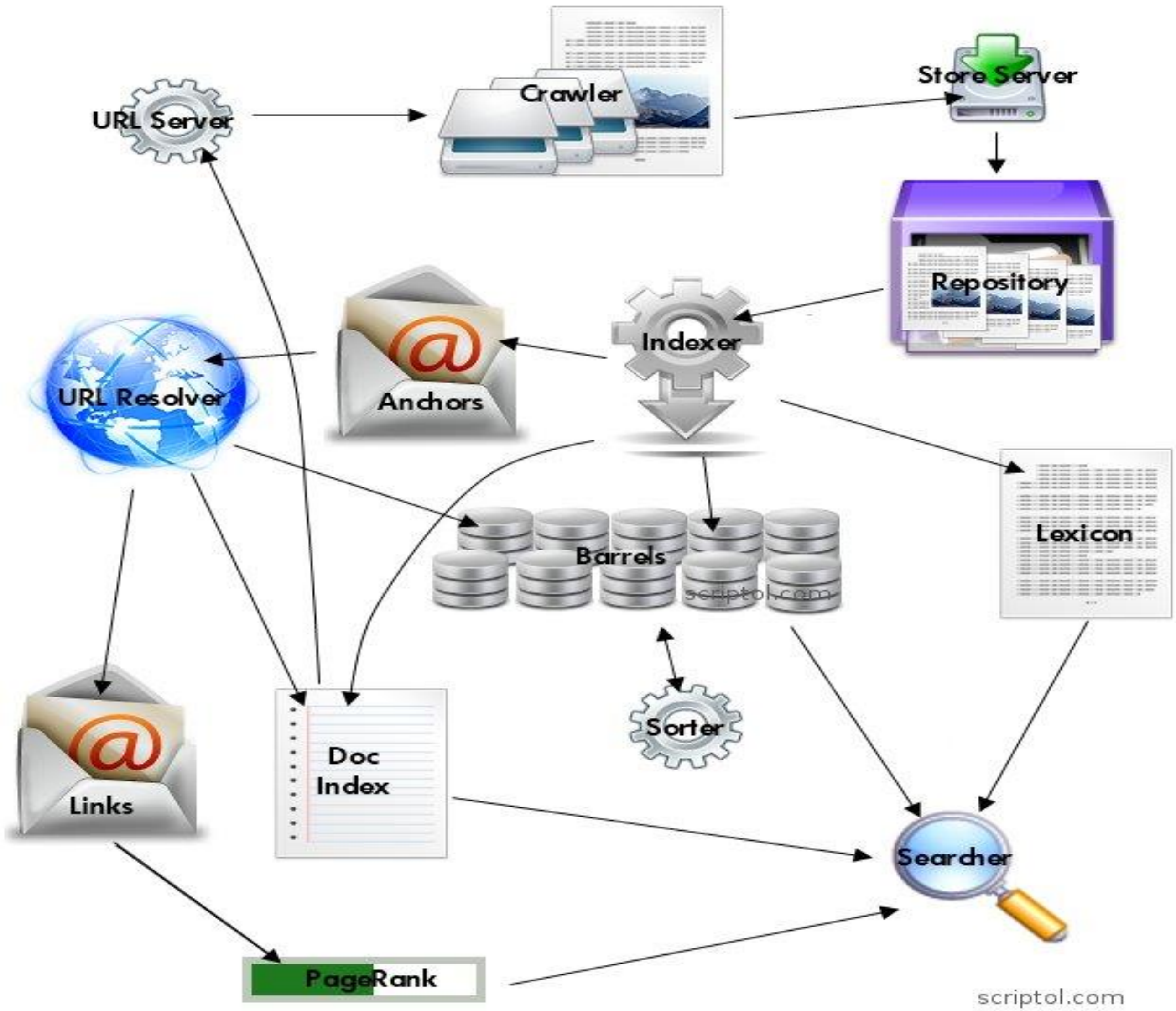
Limited computing and memory capabilities

key problem in CS&T

Challenges for Big Data Processing

- Scale-up
  - Massive data
  - Sponteneous high read/write workloads
- Availability
  - Goal of today's web services:
    As long as the network is on, even if there is a node or communication failure, the service can still be available.
  - The website is most likely unavailable when needed most → Big Loss

Google's Dream and Challenge



Write Once,
Many Reads
(millions per seconds)

Massive Web Pages

Low Value Density

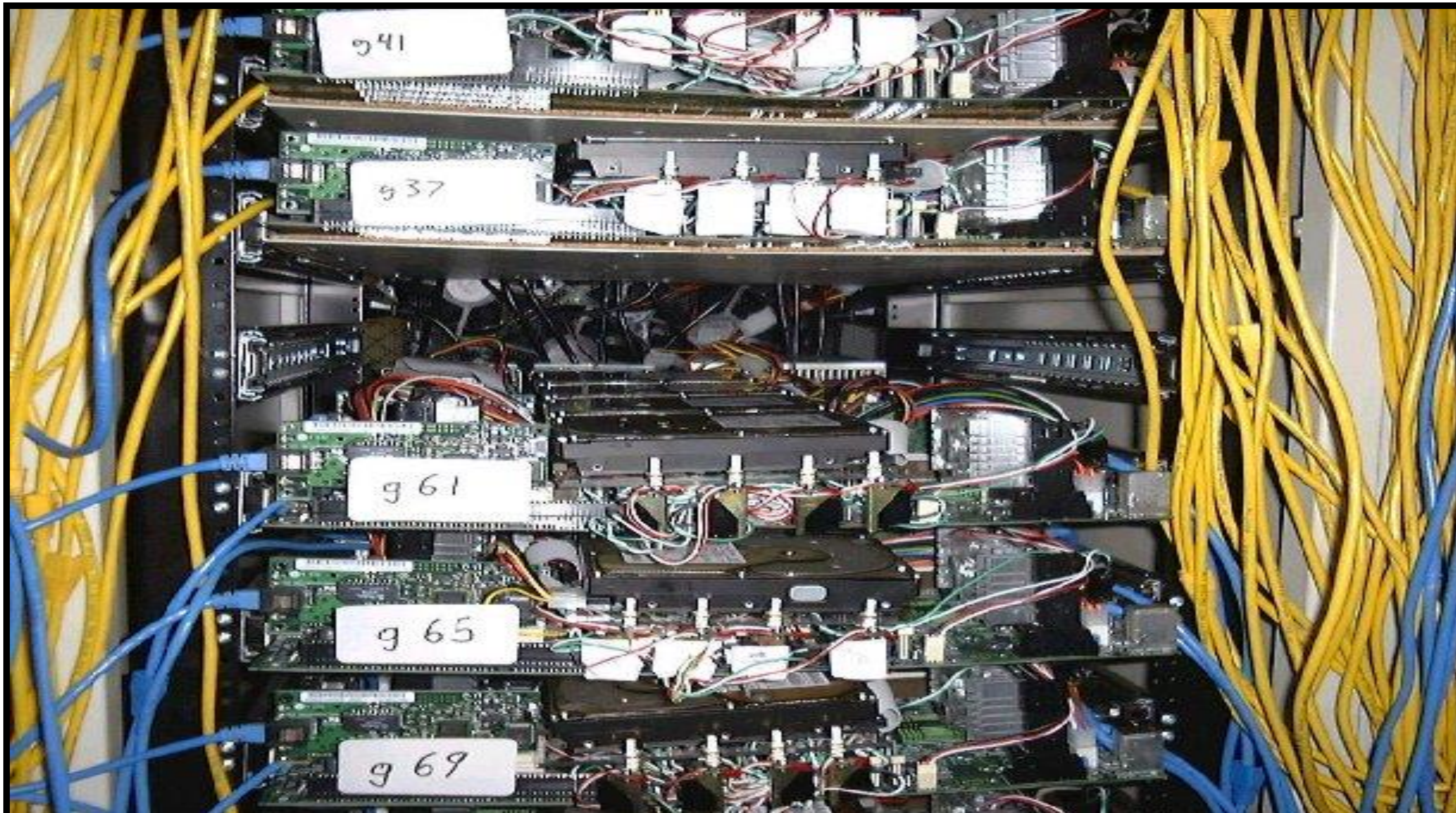
Cost Sensitive
cost per search < 5 cents

Google 's Hardware Solution

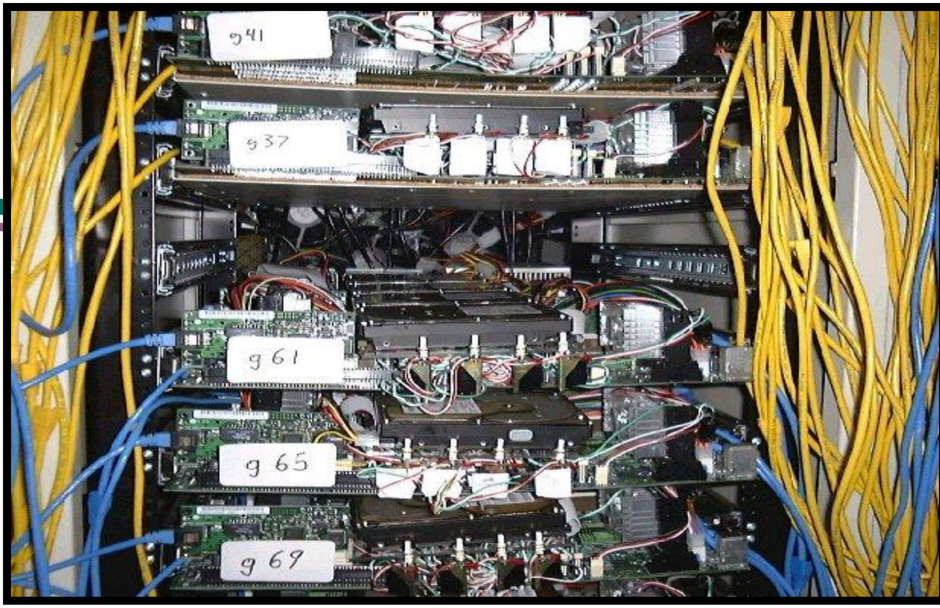
- **Server or PC?**

	Server	PC
Computing Power	High	Low
Data Storgae	Big	Small
Failure Rate	Low	High
Cost	Expensive	Cheap

Google 's DIY Hardware Platform



来源: Mass Data Processing Technology on Large Scale Clusters



- “small machines” → “big system”
- “bad machines” → “good system”

Prepare "distributed data" for "parallel computing"

Contributions of Google

- **Overturn**
  - **inertial thinking – “One Size Fits All”**
  - **the monopoly of industry giants - IBM, Oracle, Microsoft ……**
- **Contribution**
  - **detonate the brutal growth of open sourced big data systems like**

Big Data Landscape

Vertical Apps

- Predictive Policing
- bloomreach
- MYRRIX

Log Data Apps

- splunk>
- loggly
- sumologic

Ad/Media Apps

- rocketfuel
- collective[i]
- bluefin
- Recorded Future
- LuckySort
- Media Science
- TURN
- DataXu

Data As A Service

- kaggle
- factual.
- knoema beta
- GNIP
- DATASIFT
- Windows Azure Marketplace
- INRIX
- LexisNexis
- Space Curve
- LOQATE

Business Intelligence

- ORACLE | Hyperion
- SAP
- Business Objects
- RJMetrics
- Microsoft | Business Intelligence
- IBM
- COGNOS
- birst
- Autonomy
- MicroStrategy
- QlikView
- bime
- DOMO
- Chart.io
- GoodData

Analytics and Visualization

- tableau
- Palantir
- OPERA
- metaLayer
- dataspora
- METAMARKETS
- centrifuge
- TERADATA ASTER
- sas
- TIBCO
- KARMASPHERE
- panopticon
- pentaho
- Datameer
- platfora
- ClearStory
- CIRRO
- alteryx
- visual.ly
- AYATA

Analytics Infrastructure

- Hortonworks
- VERTICA
- MAPR
- cloudera
- INFOBRIGHT
- PARACCEL
- EMC
- GREENPLUM.
- NETEZZA
- kognitio
- DATASTAX
- EXASOL
- calpont

Operational Infrastructure

- COUCHBASE
- 10gen
- TERADATA
- HADAPT
- TERRACOTTA
- VoltDB
- MarkLogic
- INFORMATICA

Infrastructure As A Service

- amazon web services
- Windows Azure
- infochimps
- Google BigQuery

Structured Databases

- ORACLE
- MySQL
- SQL Server
- PostgreSQL
- IBM DB2.
- SYBASE
- memsql

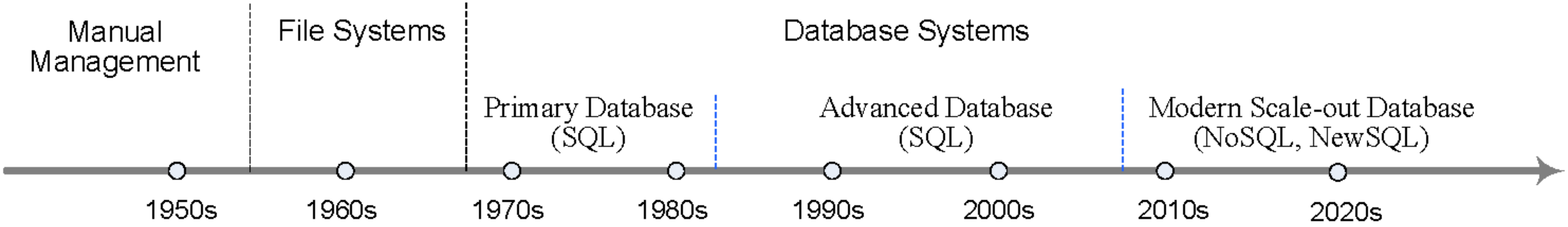
Technologies

- hadoop
- hadoop MapReduce
- mahout
- APACHE HBASE
- Cassandra

dave@vcdave.com
blogs.forbes.com/davefeinleib

Three stages in the development of data management technology

From numerical calculation to data processing (late 1950s), computers are widely used to process data and support decision making in a large scope of enterprises and domains.



- **Manual Management Stage** (before the mid-1950s)
- **File System Stage** (from the mid-1950s to late-1960s)
- **Database System Stage** (after the late-1960s)

Manual Management Stage
(before the mid-1950s)

- The capacities of computers internal memory and external storage were quite limited. The external storage only had magnetic tapes, and no disks.
- Computers were only used for numerical calculation.
- Data processing was mainly conducted manually.

File System Stage
(from the mid-1950s to the late-1960s)

The development of the file systems also progressed from low to high level, from simple to complex, going through two periods.

- **Primary File Systems (in the early-1960s)**
- **Mature File Systems (in the late-1960s)**

Primary File Systems
(in the early-1960s)

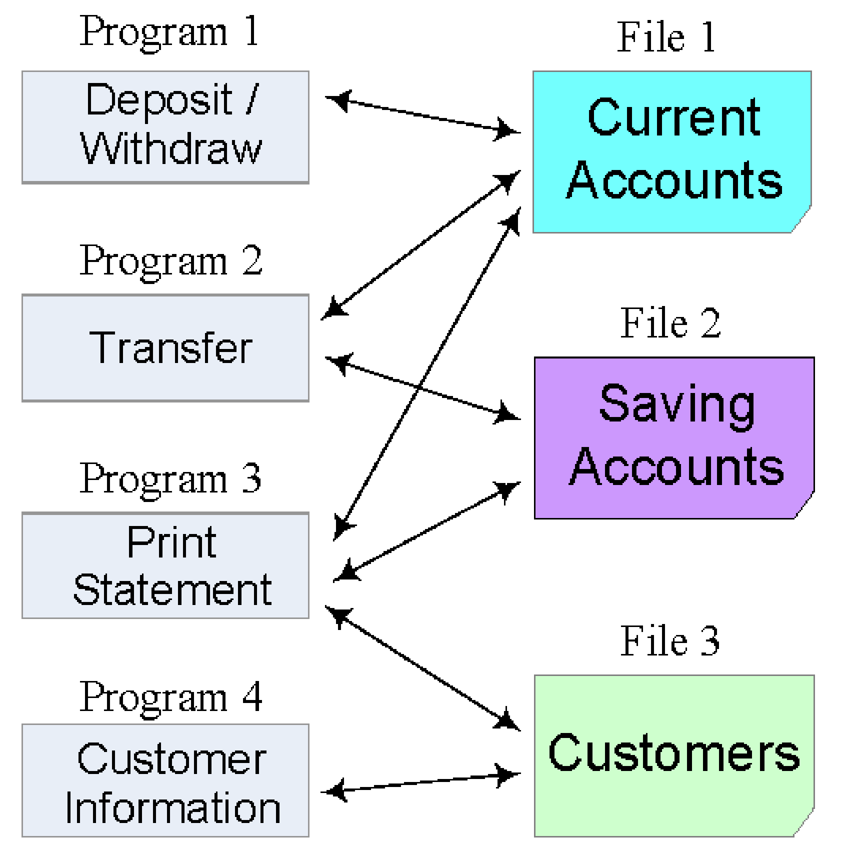
- Only sequential file organization was supported.
- Only batch processing was allowed.
- The physical structure of stored data and the logical structure of users data were the same.
- The software only did input/output processing.
- Changes in data structures led to changes in applications.
- System performance was good when serving only one application.
- Data was highly redundant.
- There was no centralized management of data.

Mature File Systems
(in the late-1960s)

- Sequential file organization and random file organization could be used together.
- Batch processing and real-time processing co-existed.
- There was a simple transformation between the physical structure of stored data and the logical structure of users data.
- The software could not only conduct input/output operations, but also provide data storage and access methods in a logical manner.
- System performance was satisfactory when serving one application.
- There was data redundancy.
- There was no centralized management of data.

Limitations of File Systems

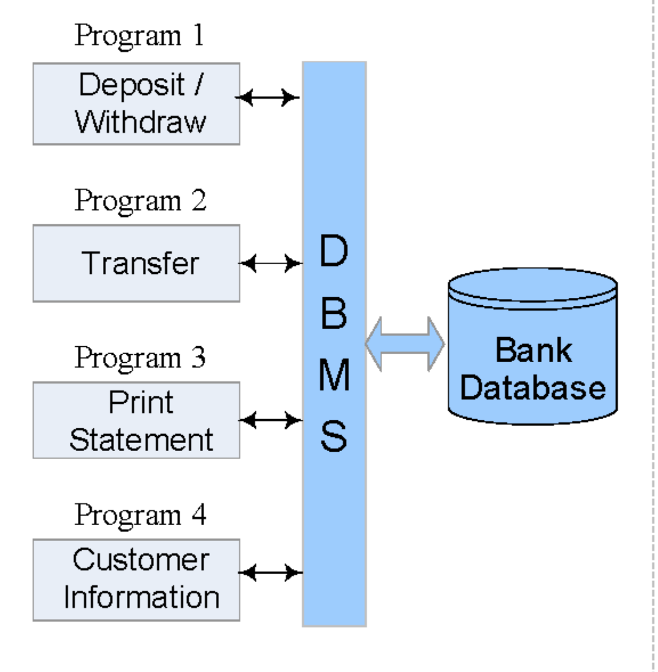
- The file or data is for a single user.
- Data cannot be shared.
- Data independence is poor.
- There is no centralized management of data.



(a) File-based Implementation

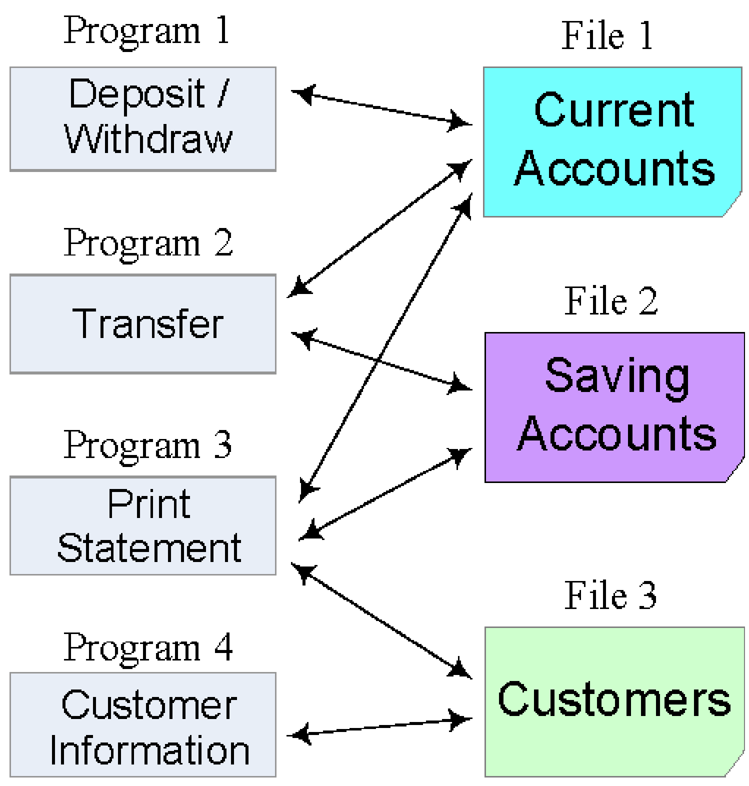
File Database Systems Stage
(after the late-1960s)

- Unlike file systems, users can abstractly and logically store, manipulate, and access data through a software system without being bothered by the low-level physical storage details.
- Analyzing and interacting with the underlying cumbersome physical file structures are left to the software system.

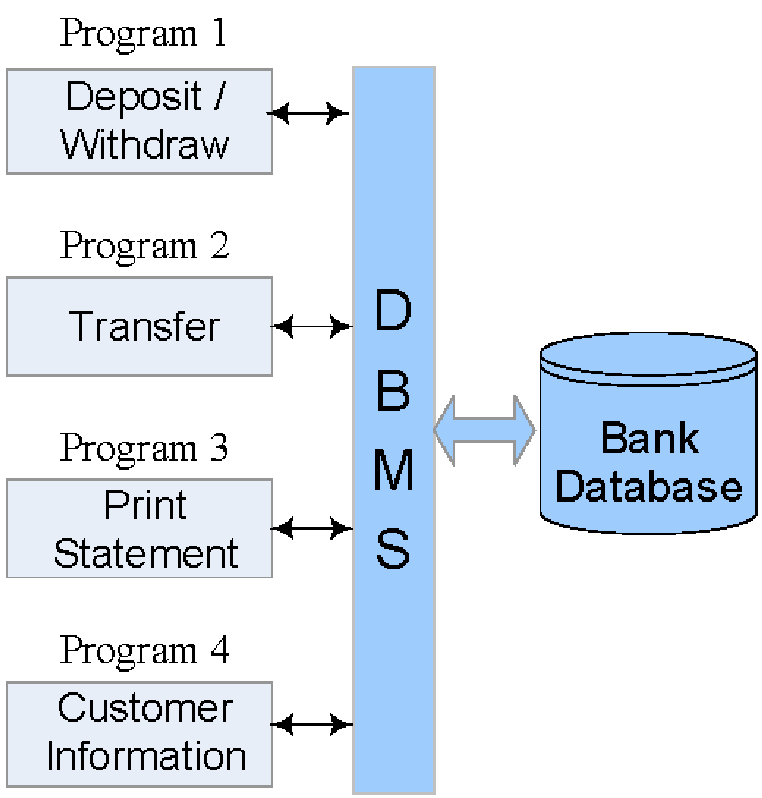


(b) Database-based Implementation

File System vs. Database System



(a) File-based Implementation



(b) Database-based Implementation

Database Systems Stage
(after the late-1960s)

Like the file systems, database systems went through three stages of development:

- **primary database systems**
- **advanced database systems**
- **modern scale-out database systems**

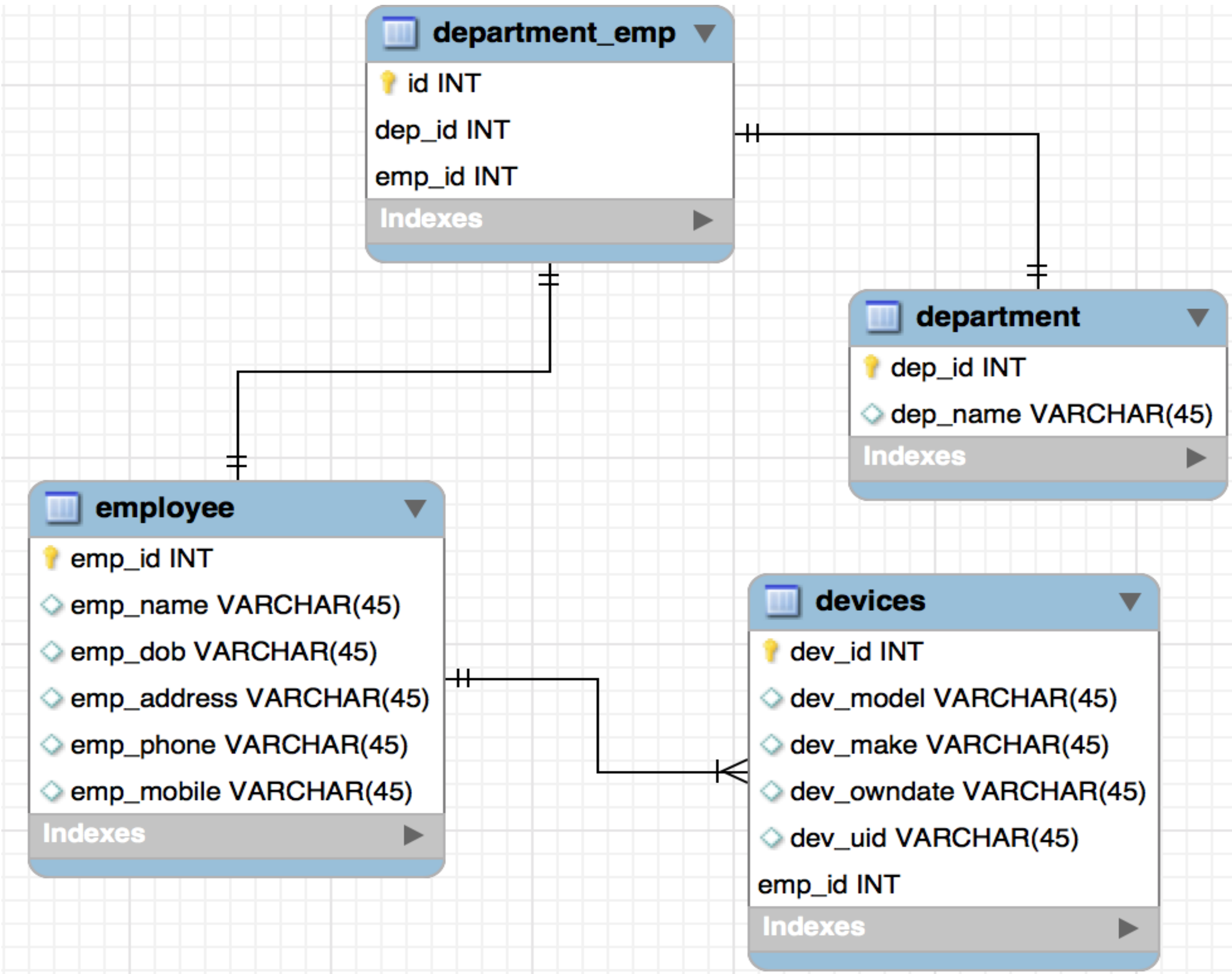
Primary Database Systems
(from the late-1960s to the early-1980s)

- **Event 1: In 1968, IBM developed the world' s first database management system - Information Management System (IMS).**
  - **Hierarchical data model**
- **Event 2: In 1969, the Data Base Task Group (DBTG) of the Conference on Data Systems Languages (CODASYL) published a DBTG report.**
  - **Network data model**
- **Event 3: In 1970, E.F. Codd from IBM published a paper "A Relational Model of Data for Large Shared Data Banks", introducing the relational databases model, and the theory of relational databases.**

Relational Database – Table Example

ID	Name	City	Country
1	Espen	Oslo	Norway
2	Harald	Munich	Germany
3	Sam	San Jose	USA

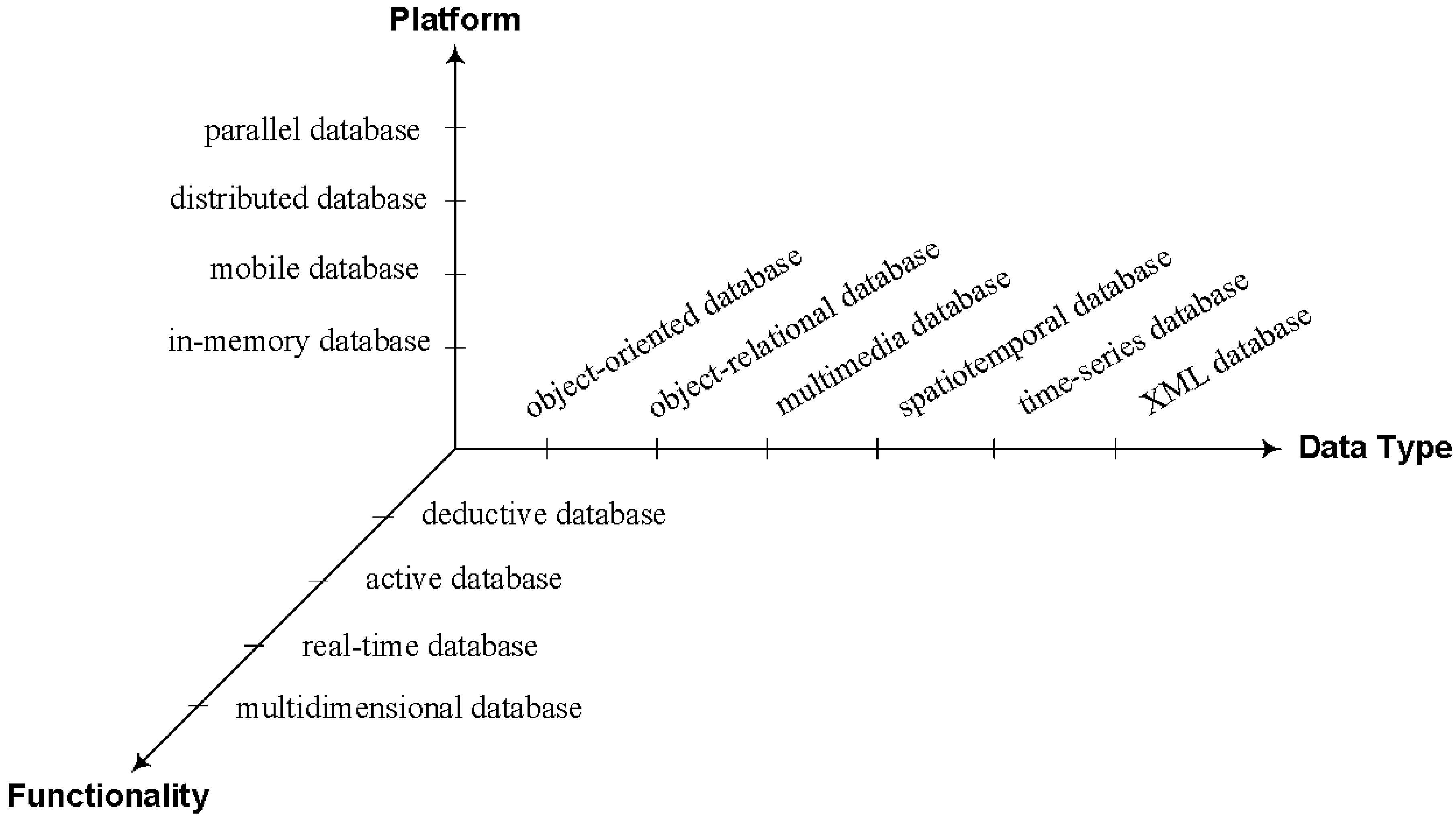
Linkage of Relational Tables



Advanced Database Systems
(from the early-1980s to the late-2000s)

- **Handling diverse data types,** including object-oriented database, object relational database, multimedia database, spatiotemporal database, time series database, XML database, etc.
- **Providing advanced functionalities,** including deductive database, active database, real-time database, multidimensional database, etc.
- **Enforcing diverse running platforms,** including in-memory

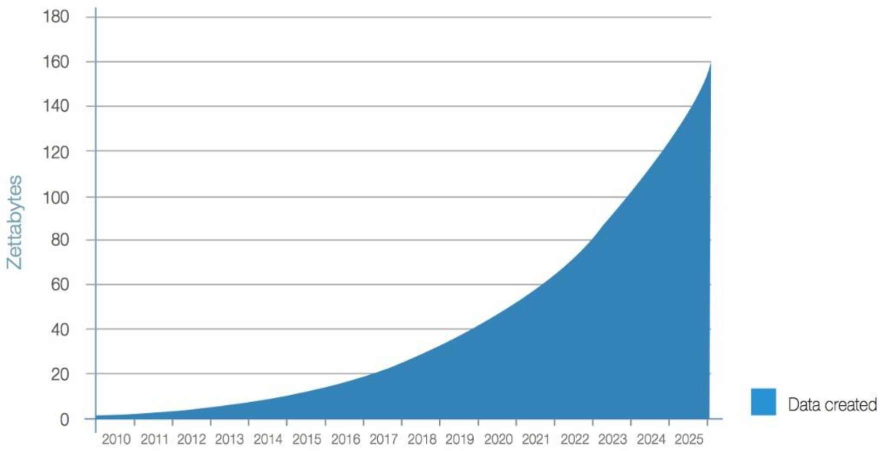
Some representatives of advanced database systems



Modern Scale-out Database Systems (after the late-2000s)

Huge amounts of mobile and real-time data have been generated from various embedded systems and Internet of Things (IoT).

- **From Data to Big Data**



Annual size of the global data sphere (Source: IDC's Data Age 2025 study)

- **Challenges for Big Data Management**
  - **Scalability and Availability**

Modern Scale-out Database Systems
(after the late-2000s)

- **Limitations of Traditional SQL Databases**
  - **ACID-based strong data consistency, i.e., every database transaction is:**
    - Atomic (transactions either succeed or fail);
    - Consistent (only valid data is saved);
    - Isolated (concurrent transactions do not affect each other);
    - Durable (committed data is persisted on disk).
  - **But can only support two of three desired characteristics:**
    - consistency, availability, and partition tolerance.
  - **Pre-defined rigid database schema**

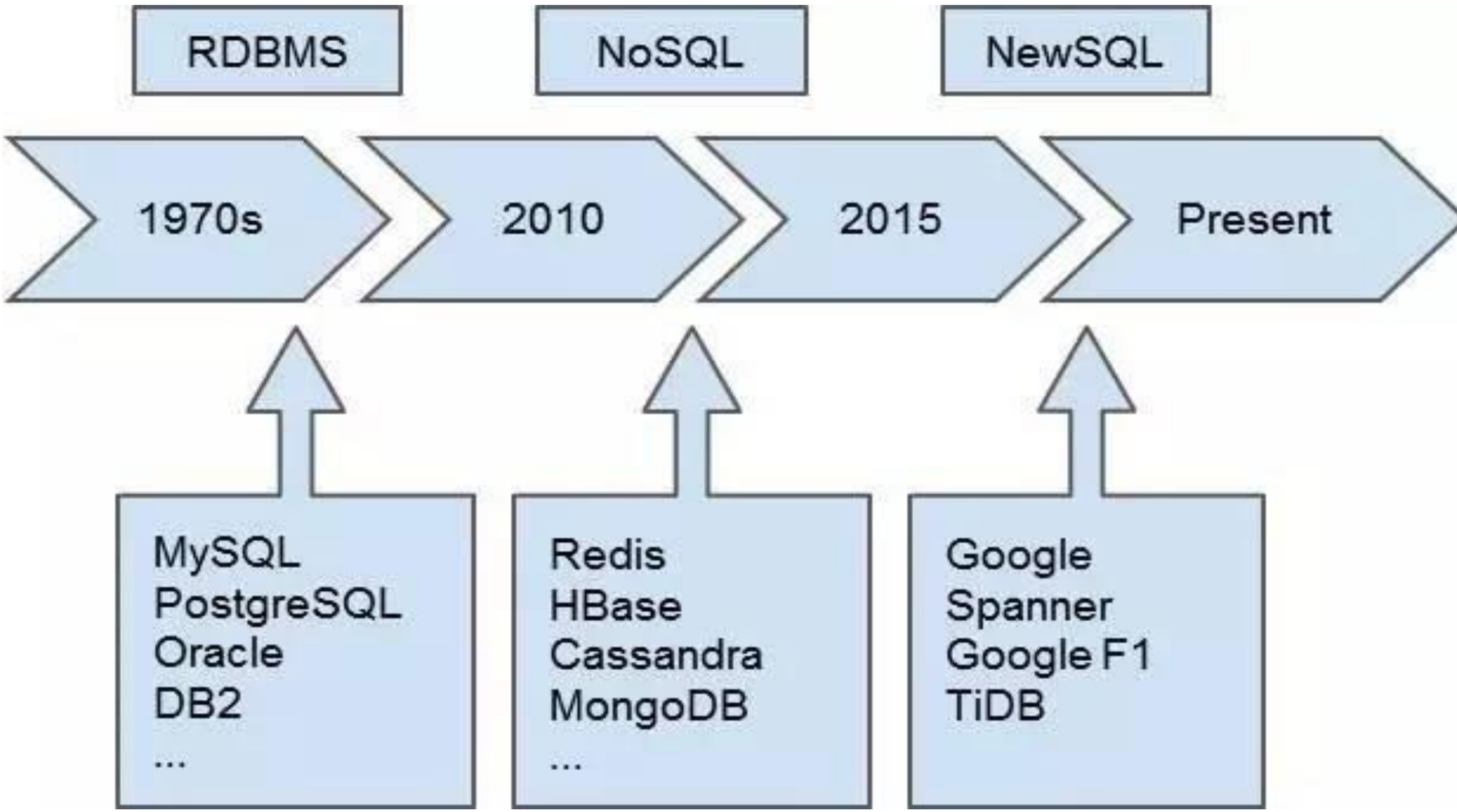
From SQL to NoSQL Databases
(in the late-2000s)

- Address Scalability and Availability
  - Relax one or more of the ACID properties
  - Emphasize availability and partition tolerance
  - Support eventual data consistency
- Three Seeds of NoSQL Databases
  - Seed 1: “Towards Robust Distributed Systems” (Eric A. Brewer, PODC 2000)
  - Seed 2: “Bigtable: A Distributed Storage System for Structured Data” (Google, OSDI 2006)
  - Seed 3: “Dynamo: Amazon’s Highly Available Key-value Store” (Amazon, SOSP 2007)

From NoSQL to NewSQL Databases
(in the early-2010s)

- **Advantages of NewSQL databases**
  - provide the same scalability and availability of NoSQL databases, while still maintaining the ACID guarantees of traditional SQL relational databases.
  - offer the best of both SQL and NoSQL worlds, i.e.,
    - 1) the relational data model and ACID transactional consistency of traditional SQL databases, and the primary interactivity of SQL;
    - 2) the scalability and availability of NoSQL databases.

SQL → NoSQL → NewSQL

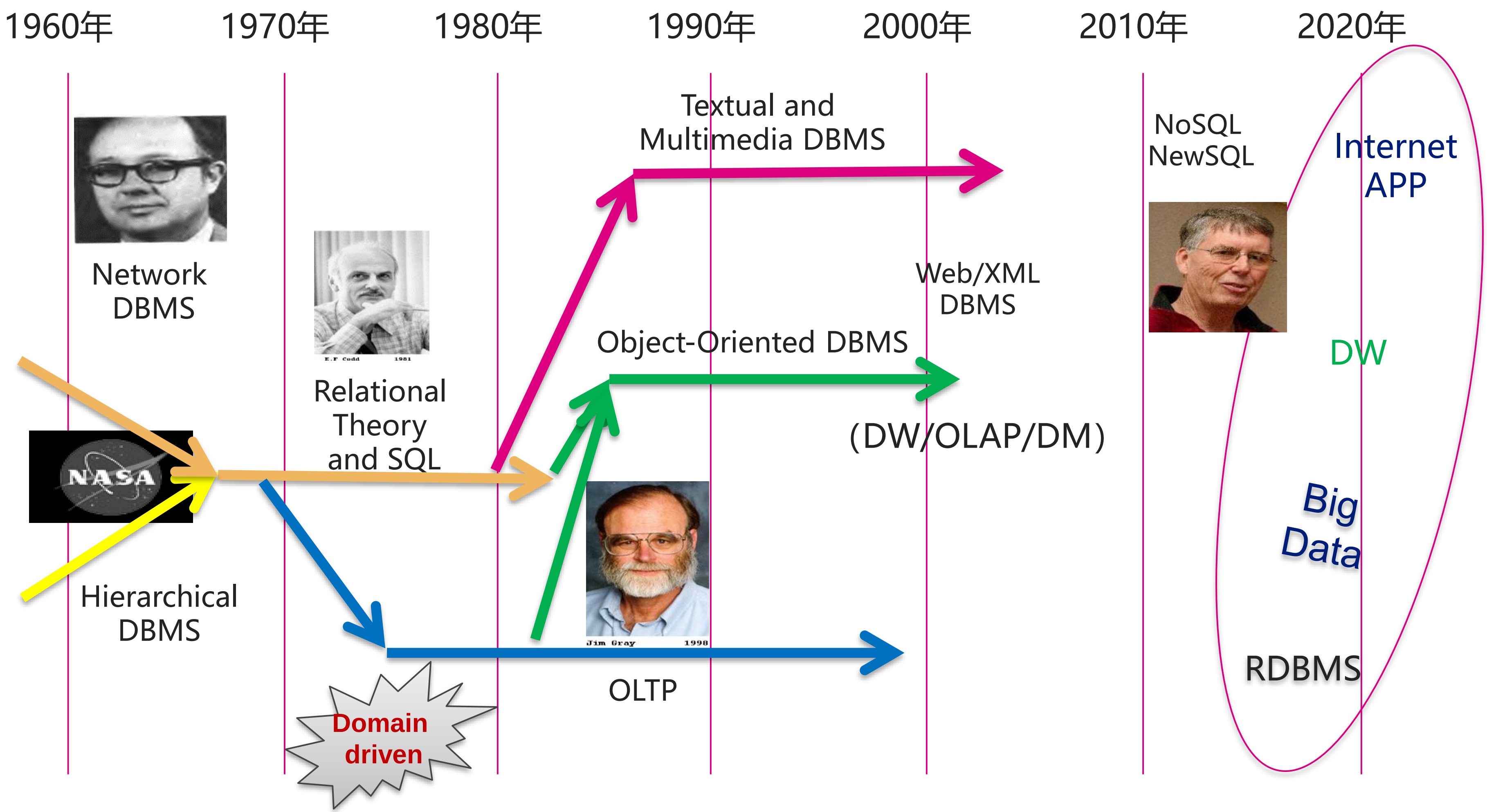


Describe/Store/Process Data
Data Integrity/Consistency

Scale-Up
Available

Both

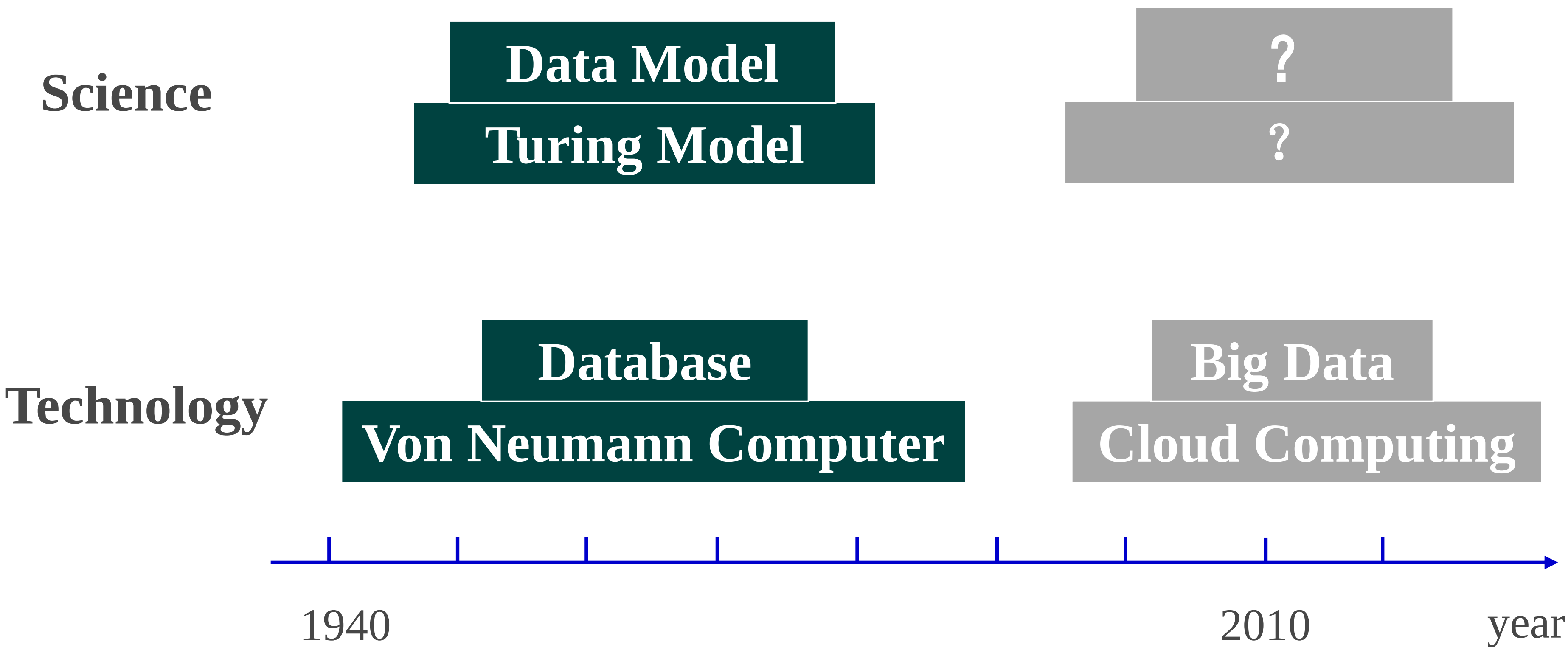
Overview of Data Management



From "One size fits all"
To "One size fits none"

Some Promising Solutions

1. On the Cloud
2. Multi-tenants
3. Combine OLAP and OLTP
4. Failure recovery re-usable and autonomous


Science
Data Model
Turing Model
?
?
Technology
Database
Von Neumann Computer
Big Data
Cloud Computing
1940
2010
year

So, we have the course.
What’s your course expectation?

1. Course Overview

- Course Background
- → Course Goal
- Course Content
- Course Assessment

Course Goal

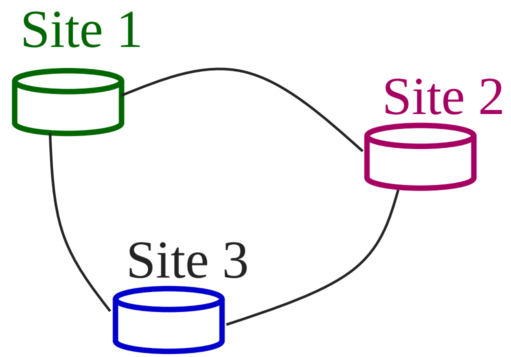
1. Review the historic development of data management technologies;
2. Enhance the previous knowledge of database systems by deepening the understanding of the theoretical and practical aspects of the database technologies;
3. Understand the need for distributed database technology to tackle deficiencies of centralized database systems;
4. Introduce basic principles and implementation techniques of distributed database systems;
5. Expose active and emerging issues in distributed database systems and application development;
6. Apply theory to practice by building a data center in a distributed context.

1. Course Overview

- Course Background
- Course Goal
- → Course Content
- Course Assessment

Course Content – Theory

- Histroric development of data managment technologies
- Distributed DataBase Systems (DDBS)
  - Architecture
  - Design (fragmentation and allocation)
  - Query processing and optimization
  - Transaction management and concurrecy control
  - Failure recovery and reliability
- State-of-Art big data management
  - SQL→ NoSQL → NewSQL
  - parallel and streaming data management
  - data warehousing and OLAP



Course Content - Practice

- Teamwork of 2 members
- Design and implement a data center in a distributed context
- Demonstration, presentation, and documentation

1. Course Overview

- Course Background
- Course Goal
- Course Content

→ **Course Assessment**

Course Assessment

1. Individual Assignments (30%)
   - distributed database architecture and design
   - distributed database querying
   - distributed transaction management and concurrency control
2. Group Project (60%)
   - Demo, Presentation, Mannual, and Report
3. Course Attendance (10%)

Scientific Fostering 科学培育

- basic concepts
- problem formulation
- identification of key technical issues and scientific challenges
- possible solutions
- algorithm presentation
- system level design and implementation
- testing
- maintenance

Prerequisites

- Can know or learn the knowledge of Database and Computer Network in the next two weeks.
- Can think, write, compile, run, and debug computer programs.
- Have the time and the will to work hard for a semester to achieve a worthy goal.

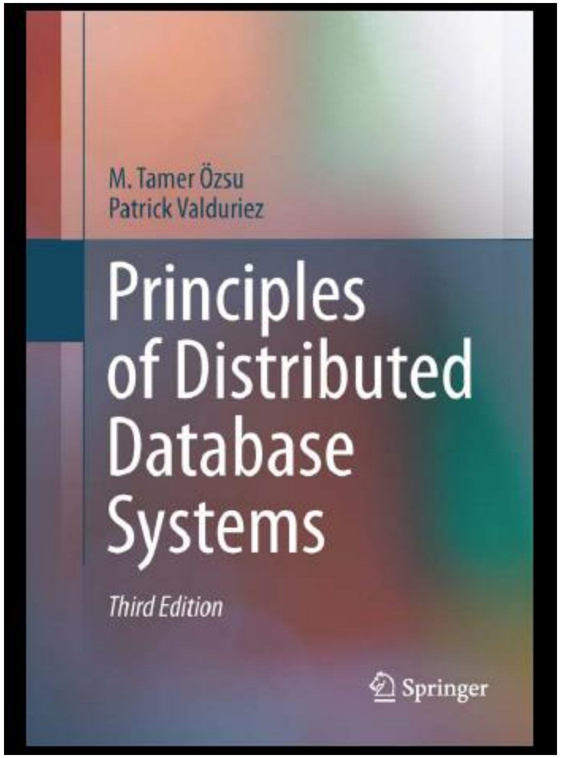
After the Course

- get familiar with the currently available models, technologies for and approaches to building distributed database systems;
- have developed practical skills in the use of these models and approaches to be able to select and apply the appropriate methods for a particular case;
- be aware of the current research directions in the field and their possible outcomes;
- be able to carry out research on a relevant topic, identify primary references, analyse them, and come up with meaningful conclusions;
- be able to apply learned skills to solving practical database related tasks.

Main Textbook

Principles of Distributed Database Systems

M. Tame Özsu
Patrick Valduriez
Prentice-Hall, 2011



Course Lecturer

- Ling Feng (冯铃)
  - Office: East Main Building #10-208
  - Phone: 62773581, 15901369009
  - Email: fengling@tsinghua.edu.cn

How to Succeed?

- Where there is a will, there is a "good"!
- Lectures are designed to convey information, not to be mechanically memorized.
- LISTEN to what is said.
- THINK about what is said.
- ASK questions when you have doubts.
- WRITE down only what is important.
- PRACTICE what has learned.

Tip #1

Don’t wait until the last minute to get help.

Tip #2

Bad things happen while learning a new programming skill. You will probably crash and burn on some programs. Start early; give yourself time for mistakes.

Tip #3

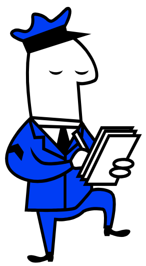
Don’t be too ambitious with your course load. You CANNOT slack off in this class, even for a few days.

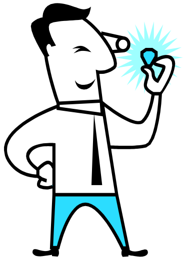
Tip #4

- Critical Thinking
- Research
- Problem Identification and Solving
- Speaking
- Writing
- Work Ethic
- Number Crunching
- Physical Performance
- Influencing People
- Teamwork

Ten things employers want you to learn in college.

Tip #5





Learn Question Practice Enjoy