Distributed Cloud Networks Next Gen Cloud Service Platform

Carnegie Mellon University
School of Computer Science
Yudong Liu

The Future of Information Era

- Stronger and more interpretable AI
- Massive Data Analysis
- IoT
- Cloud interaction Platforms

....

Most important factors of Production

- Data
- Algorithms / Models
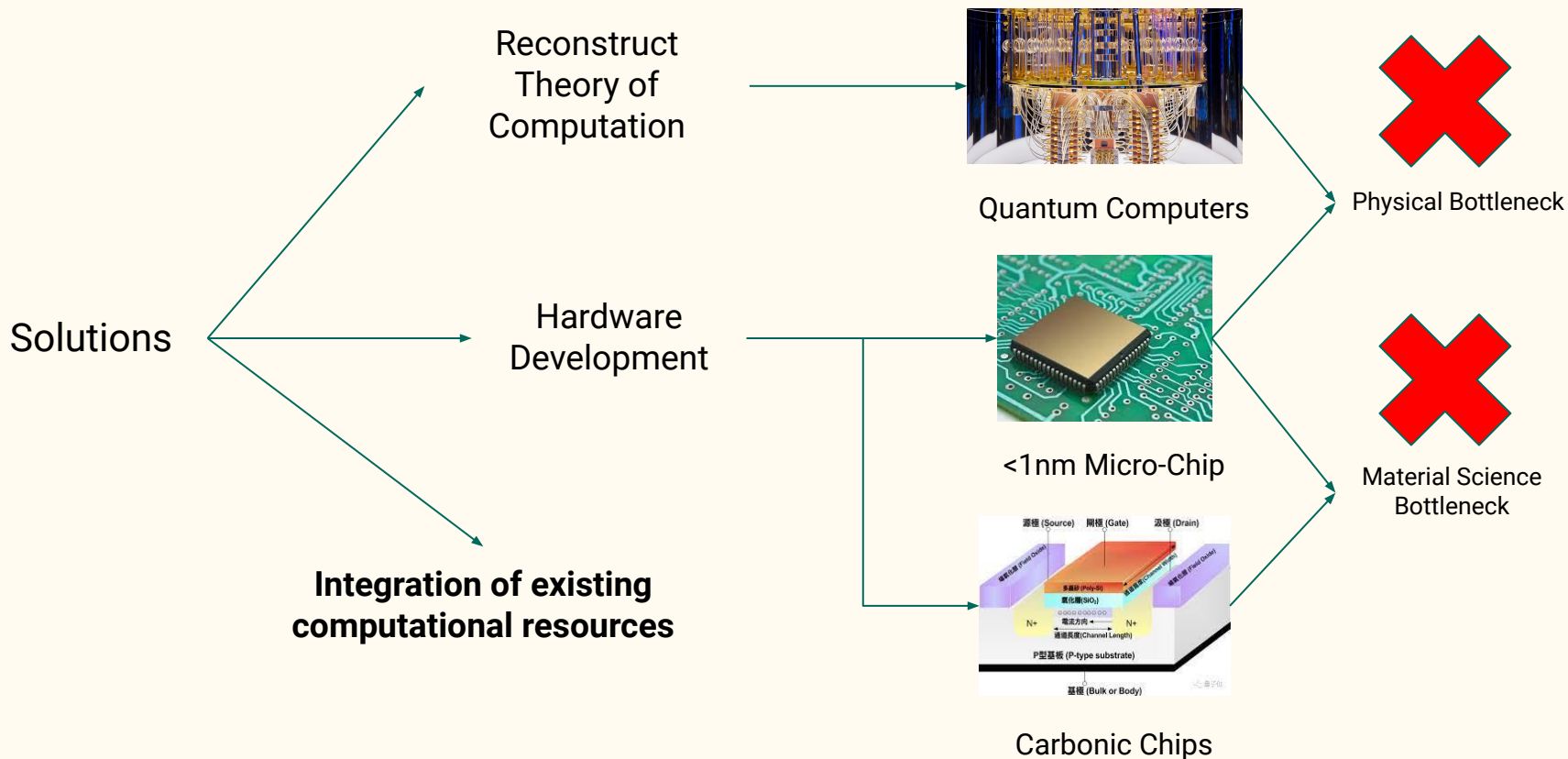- Computational Power

Take ML as an Example

- Data Mining and Processing
  - Oracle
  - Google
- Large Model Development
  - OpenAI
  - Facebook
- Computational Power (Limitation)
  - Amazon (AWS)
  - Alibaba Cloud

Model capabilities limited by the size of its parameters:

- GPT1 117 Million Parameters
- GPT2 1.5 Billion Parameters
- GPT3 175 Billion Parameters
- GPT4 > 1 Trillion Parameters
- Human Brain: >100 Trillion Connections

Computational Power is the limitation Factor

Solutions

Reconstruct
Theory of
Computation

Quantum Computers

Hardware
Development

<1nm Micro-Chip

**Integration of existing
computational resources**

Carbonic Chips

Physical Bottleneck

Material Science
Bottleneck

Cloud Computing Too Expensive

Insufficiency of Integrated Computational Resources

- Tech Giants Monopoly over cloud service resources
- Advanced Hardware unaffordable for individuals
- Cloud Computational Resources Insufficient w.r.t to the population sizes

Personal Computational Devices Wasted

- Billions of family CPU and GPU excluded from production
- Most people unaware of their computational resources as means of production

BluePrint（Decentralization of Computational Power）

- **STEP 1**：Use PC as a family server node, provide sharing of computational power for other personal devices such as phones and laptops
- **STEP 2**：Based on the latest transmission technologies（such as 5G), conduct linking across small servers and PCs in limited scales remotely to test distributed cloud computing
- **STEP 3**：Based on the latest transmission technologies, link all personal computational resources around the globe into a distributed cloud computing network
- **FINAL GOAL**：Use the integrated computational resources of mankind to provide platform for future mass-scale computations
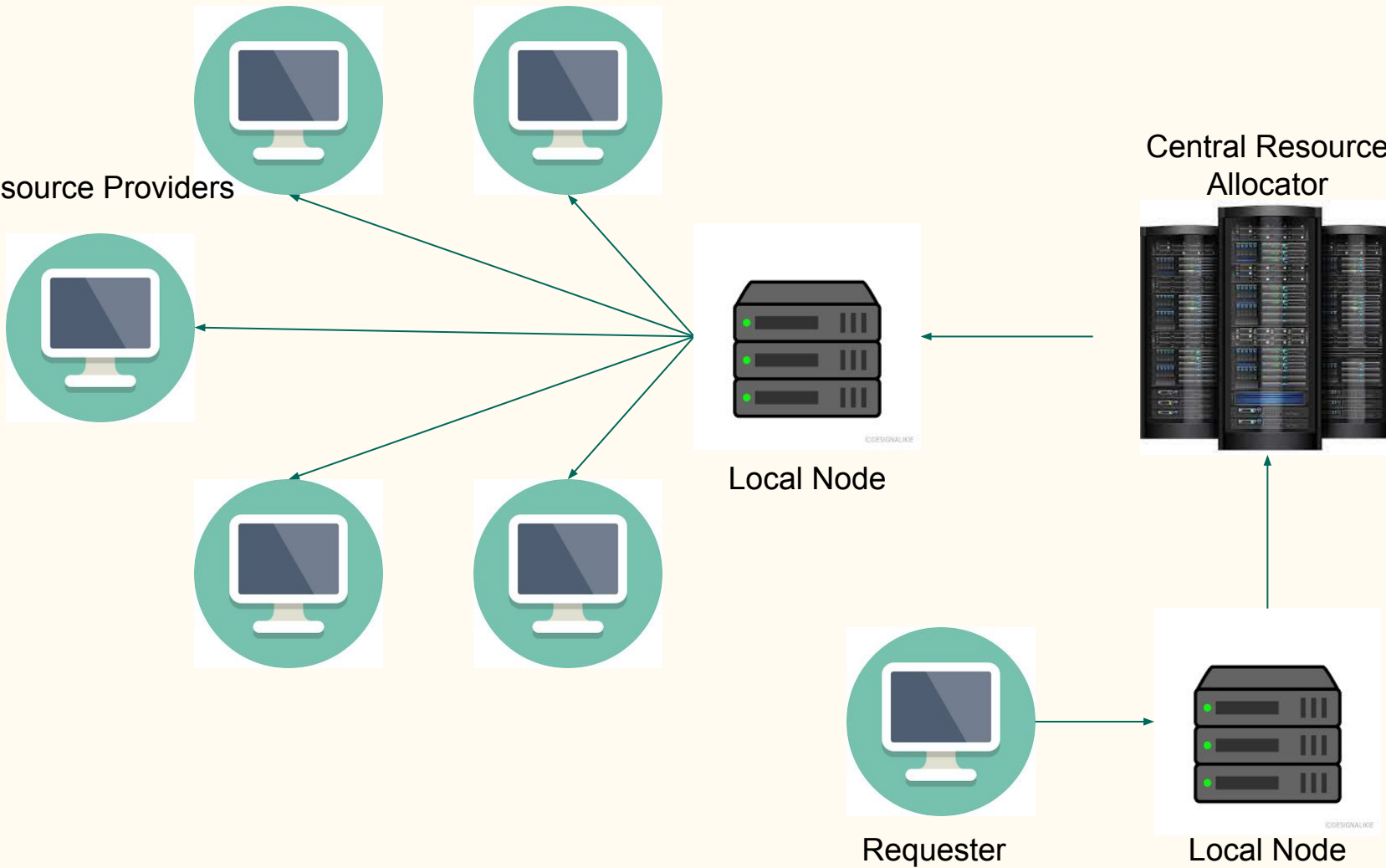
Resource Providers
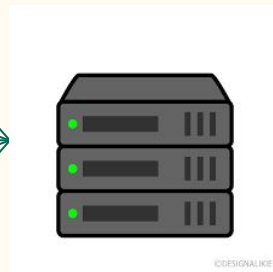
Central Resource Allocator

Local Node

Local Node

Requester

Milestones

1. **Local PC service** : PC service for near cellphones, laptops and tablets
2. **Remote PC service** : Remote service from PC to other devices
3. **Small Scale Networking** : Self-sufficient computational power sharing across devices of small enterprises
4. **Large Scale Networking** : Anyone could access the computational power of the whole society

Local PC service：Family Server（PC=Server）

1. Achieve Computational Resource Sharing across Personal Devices

2. Agent: LAN, Bluetooth

3. Partially replace cloud storage and rendering service

4. Scenario：

   a. Access High-GPU demand Computer Games from cell-phones

   b. 3D real-time rendering on PC viewed from cell-phones

5. Less limitation for App developers

   a. C++，python based Apps could be accessed on cell-phone

6. **Use PC as server node, provide basis for distributed networking of PCs**

Remote PC service：Remote Personal Servers

1. Remote Access of PC Computational Resource from other devices
2. Agent： Wireless Network（4G，5G）
3. Target Users： Individual Developers; People without PCs
4. Scienario：
   a. Access your friends' free GPU for ML
   b. Remote access of PC softwares
   c. Use PC for entertainment purposes while traveling
5. Advantages：
   a. Convenient: No need for resetting the environment
   b. Cheap
6. **Basis for Remote Large Scale Sharing of Computational Resources**

Integration of Computational Resources for small Enterprises (Limited Scale)

1. Agent：Wired or Wireless Networks

2. Target Users：Enterprises without their own Cloud Computing Clusters

3. Core Advantages：

   a. Each enterprise can access the computational resources of involved enterprises combined

   b. Provide cloud service of a much lower cost for all involved enterprises and personnel

   c. Service Provider no longer needs to maintain cloud computing hardwares

4. Scienario：

   a. Hosting high data-flow websites for specific events such as Expo

   b. Reduce Technological research costs for small enterprises

5. **Provide basis for Large Scale Distributed Cloud Networking**

Universal Computational Resources Sharing

1. Agent：Wired and Wireless Networks

2. Targe Users：Everyone

3. Based on Previous Steps, integrate all PCs and servers as nodes in the distributed cloud computing service

4. Advantages：

    a. Access to the computational resources of the whole society

    b. Break monopoly of Cloud computing service by Tech Giants

    c. Much lower maintenance fee for platform provider

    d. Significantly lower cloud computational costs

Technical Challenges

- **Wireless Transmission Throughput**
  - Solution：Next Gen Wireless Networks
  - Feasibility：High since developement of wireless networks is progressing fast. 6G is expected to be two magnitudes faster than current wireless networks
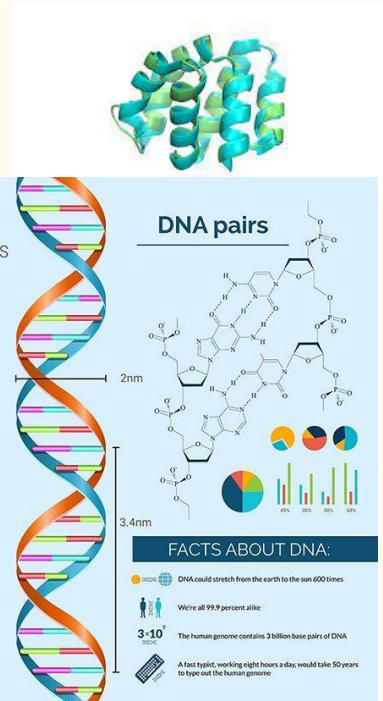- **Personal Data Security**
  - Solution：
    - Requesting installation of a small child system inside users' PC for distributed cloud computation (Isolated from their main system).
    - Remove all the temp data after computation to prevent data leakage

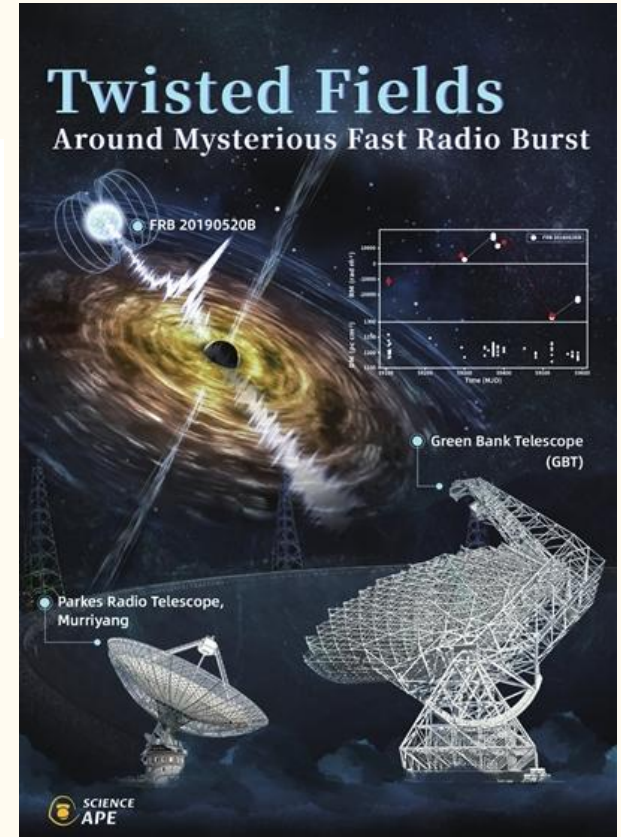# Why it matters: Technological + Social Revolution

Gather all computational resources of mankind

Strong AI

Mass Scale Bioinformatics Analysis

Mass Scale Astronomical Data Interpretation

The Foundation for the True Era of Intelligence

IoT

Smart City

Thank You

—