

# Week1: Introduction

## Edge Computing

C. García [garsanca@ucm.es](mailto:garsanca@ucm.es)

April 18, 2022

- X. Wang et al. “Edge AI, Convergence of Edge Computing and Artificial Intelligence”,  
[doi://10.1007/978-981-15-6186-3](https://doi.org/10.1007/978-981-15-6186-3)



# Outline

1 Intro

2 Trends

3 Intelligent Edge



# Motivation

- With the proliferation of computing and storage devices:
  - Server clusters in cloud data centers
  - Personal computers and smartphones
  - Wearable and other Internet of Things (IoT) devices
- We are now in an **information-centric era**
  - Computing is ubiquitous

## Some figures

- According with Cisco<sup>1</sup> in 2020 850 Zettabytes (ZB) of data will be generated each year outside the cloud
- Global data center traffic 20.6 ZB

<sup>1</sup>Fog Computing and the Internet of Things: Extend the Cloud to Where the Things Are. <https://>



# Motivation

- Traditional model: **Cloud Computing**
  - Raw data is produced and transferred to the cloud and consumers are sending the request to access the data from the cloud



- This structure is not optimal since a **large amount of data** needs to be transferred



# Motivation

- Data sources for big data are also undergoing a transformation
  - From large-scale cloud data centers to an increasingly **wide range of edge devices**

## Where Edge computing is useful

- 1 Large number of computation need to be delivered to cloud:  
**processing data**
- 2 Many new applications have **strict or tight delay requirements**, cloud computing is not the effective
  - Cooperative autonomous driving

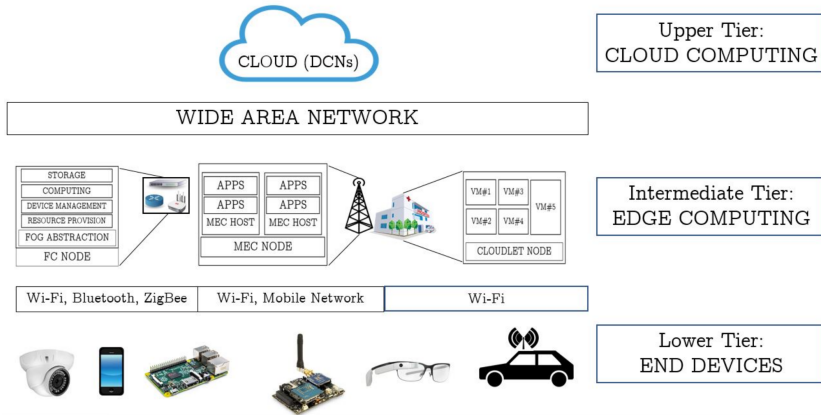


# Advantages Edge

- Compared to Cloud-computing
  - 1 **Backbone network alleviation**, distributed edge computing nodes can handle a large number of computation tasks without exchanging the corresponding data with the cloud, thus alleviating the traffic load of the network
  - 2 **Agile service response**, services hosted at the edge can significantly reduce the delay of data transmissions and improve the response speed
  - 3 **Powerful cloud backup**, the cloud can provide powerful processing capabilities and massive storage when the edge cannot afford



# Advantages Edge



# Evolution computing model

Time	Computing mode
1965-1985	Centralized processing model centered on the mainframe
1986-1990	File-sharing computing mode centered on PC/file server
1990-1996	Distributed computing model centered on C/S architecture
1996-2000-	Distributed computing model with Web and B/S architecture
2000-	Centralized processing model centered on the mainframe
2005-	Distributed computing model with Grid, P2P, Cloud, and other technologies as cores
2015-	"End-edge-cloud" collaborative edge computing model





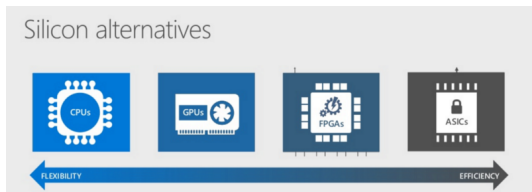
## Trends in Edge Computing

- Technological changes also helps to Edge Computing to a great development
  - From its architectural blueprint to industrial landingT
- The future development of edge computing is going to be integrated with the development of other technologies:
  - Heterogeneous computing
  - Edge intelligence
  - Edge cloud collaboration
  - 5G + edge computing



# Heterogeneous Computing

- Use machines with different performance and structure to meet different **computing needs**
- Can obtain the maximum overall performance on **heterogeneous platforms** through **algorithms**



# Edge intelligence

- Push artificial intelligence technology to the edge
  - Deploying on the edge nodes can obtain richer data faster
    - Saves communication cost and reduces response delays
  - Optimizing edge-side resource scheduling decisions
    - More efficient services



# Edge cloud collaboration

- Edge computing is an extension of cloud computing...
- ... it can complement each other with cloud computing
- Cloud computing
  - Is good at global, non-real-time, long-cycle big data processing and analysis
- Edge computing
  - Does well in field-level, real-time, and short-cycle intelligent analysis

## Collaboration

- Deploy **compute-intensive** tasks in the cloud vs **fast response** are placed on the edge
- Edge can also preprocess the data sent to the cloud to further



# 5G + Edge Computing

- 5G characteristics:
  - ultra-high speed
  - large connection
  - ultra-low latency

## Collaboration

- Edge computing is an important component of the 5G network and effectively alleviate the problem of data explosion in the 5G era
- 5G provides a good network foundation for the industrial deployment and development of the edge computing industry



# Artificial Intelligence

- Deep learning-based intelligent services have evolved
  - Great advantages of Deep Learning (DL)
  - such as Computer Vision (CV) and Natural Language Processing (NLP)
- However there are some factors to think if Cloud Computing services are the best solution:
  - What about **costs**, **latency**, **reliability** or **privacy**

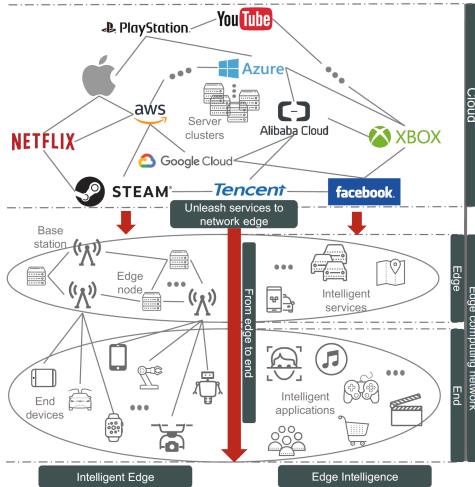


## Factors to take into account

- **Costs:** training and inference of AI models in the cloud requires to transmit massive amounts of data to the cloud
- **Latency:** delay to access cloud services is generally not guaranteed and might not be short enough to satisfy the requirements of many time-critical applications
- **Reliability:** most cloud computing applications relies on wireless communications and backbone networks...
- **Privacy:** the data required for AI might carry a lot of private information, and privacy issues are critical



# Intelligent Edge





## Factors to take into account

- Since the edge is closer to users than the cloud
- Edge computing is expected to solve many of these issues

### Combination

- Edge computing is gradually being combined with Artificial Intelligence (AI)
- Edge intelligence and intelligent edge are not independent of each other
  - Edge intelligence is the goal, and the AI services in intelligent edge are also a part of edge intelligence
  - Intelligent edge can provide higher service throughput and resource utilization for edge intelligence.



## Best choice

- 1 AI services are deployed close to the requesting users
  - Cloud service only when additional processing is required
  - It reduces latency and cost of sending data to the cloud
- 2 Since the raw data required for AI services is stored locally on the edge instead of the cloud
  - Protection of user privacy is enhanced
- 3 Hierarchical computing architecture provides more reliable AI computation
- 4 Edge computing can promote the pervasive application of AI
  - Realize the prospect of “providing AI for every person at everywhere”
- 5 Diversified and valuable AI services can broaden the commercial value of edge computing

