Discussion for Mobile Computing Power Network in 5GS

China Telecom

SA WG2 Meeting #146E

16 August - 27 August 2021, Electronic meeting

Content

Background for computing power network

- Draft Recommendation ITU-T Y.2501 (formerly Y.CPN-arch): “Computing Power Network – framework and architecture
- The objective of Recommendation ITU-T Y.2501 is to describe the framework and architecture of computing power network. It is a new type of network that realizes the optimized resource allocation, by distributing computing, storage, network and other resource information of service nodes through network control plane. It combines network context and user requirements to provide the optimal distribution, association, transaction and scheduling of computing, storage and network resources.

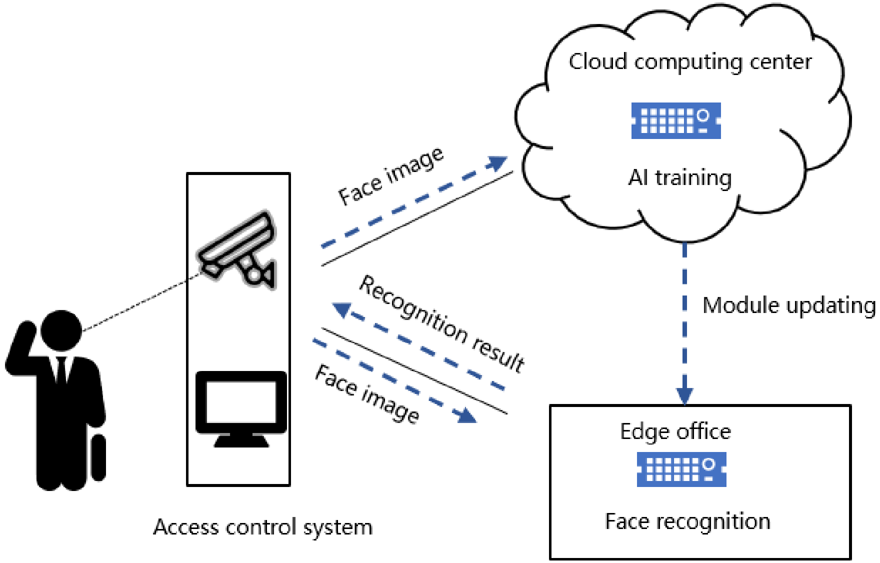
What is Mobile Computing Power Network

- Computing power nodes located in 5GS are selected for 5G users considering the service characteristics, and network conditions.
- Easy accessibility and efficient utilization of different types of computing powers for 5G users
  - Reliable computing power
  - Computing power traceability
  - Computing power measurement
- Jointly optimizing communication and computing resources for 5G user services.

Why mobile computing power network (1/3)

From the perspective of computing power resource consumers

- Services need to consider not only computing power resources but the time for processing, the time for transmitting;
- Centralized computing power resources pool does not satisfy all the requirements, such as latency.



Access control system based on face recognition

- AI training part can be deployed in centralized computing power pools, such as, cloud computing centers, to perform complex computing processes offline.
- In reasoning phase, for instance, recognizing faces to open the door, the latency should be considered.
- Distributed computing power resource pools, such as edge computing nodes, have obvious advantages in latency.

Why mobile computing power network (2/3)

From the perspective of computing power resource providers

- scattered computing resources pools with low utilization
  - The provision of computing power is relatively easy. The existing computing resources pools self-built by enterprises are scattered in location, and its utilization is pretty low.
- super-computing centers not easy to access by individual users
  - The super-computing centers own a large amount of computing power resources, while user applying for these resources is not friendly.

Therefore, it needs to find a new solution and technology system to realize computing power resources sharing.

Why mobile computing power network (3/3)

From the perspective of network technology

- Easy accessibility of mobile computing power resources via 5G networks
  - Network development has laid the foundation for the flexible provision of multi-party computing resources. Users and computing power resources can be efficiently connected through 5G networks.
- With the development of new network technologies such as 5G, mobile computing power resources can be connected in a simple and efficient way.

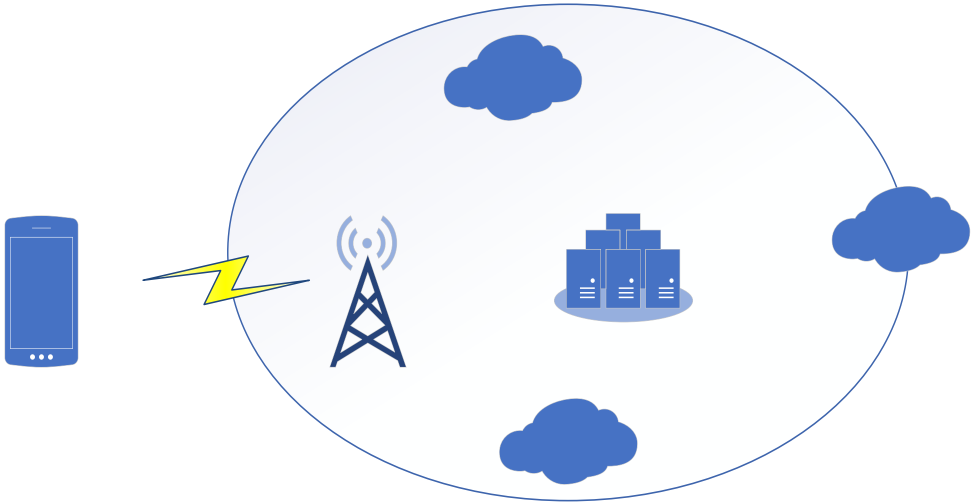
Use Case(1/2)

Characteristics of devices and services

- Limitation of computing and storage capability of mobile devices
- high demands of computing power and storage of new services, such as virtual reality(VR), Augmented Reality(AR), cloud gaming, AI service(such as Federated Learning (FL)), wireless sensing

Service model

- Same service experience of users, no matter high-end phones or low-ends phones the users are using. As users can access and utilize suitable computing power that satisfy their needs via 5G network.
- Computing power nodes are selected according to the service characteristics and network conditions.
- Guaranteed QoE for computing services, both communication QoS and computing QoS are considered



Use Case(2/2)

Group of trusted devices in

- Personal IOT network
- Industry IoT
- V2X scenario

Additional service needs

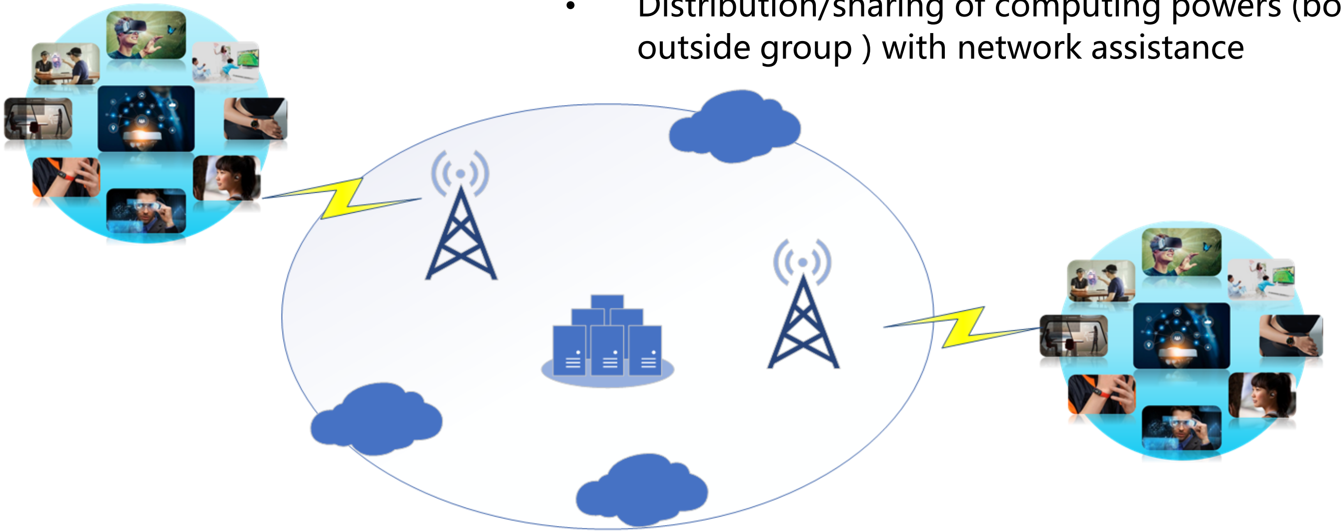
- Computing power sharing among group of trusted devices

Challenges

- Heterogeneous operating systems and computing powers of devices from different vendors

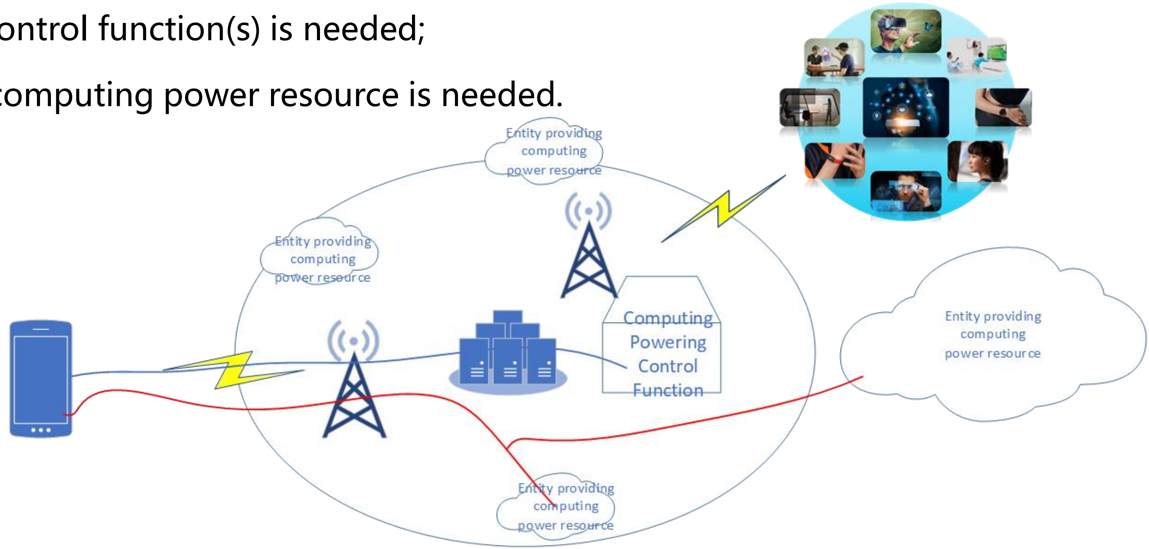
New service model

- Distribution/sharing of computing powers (both insider group and outside group) with network assistance

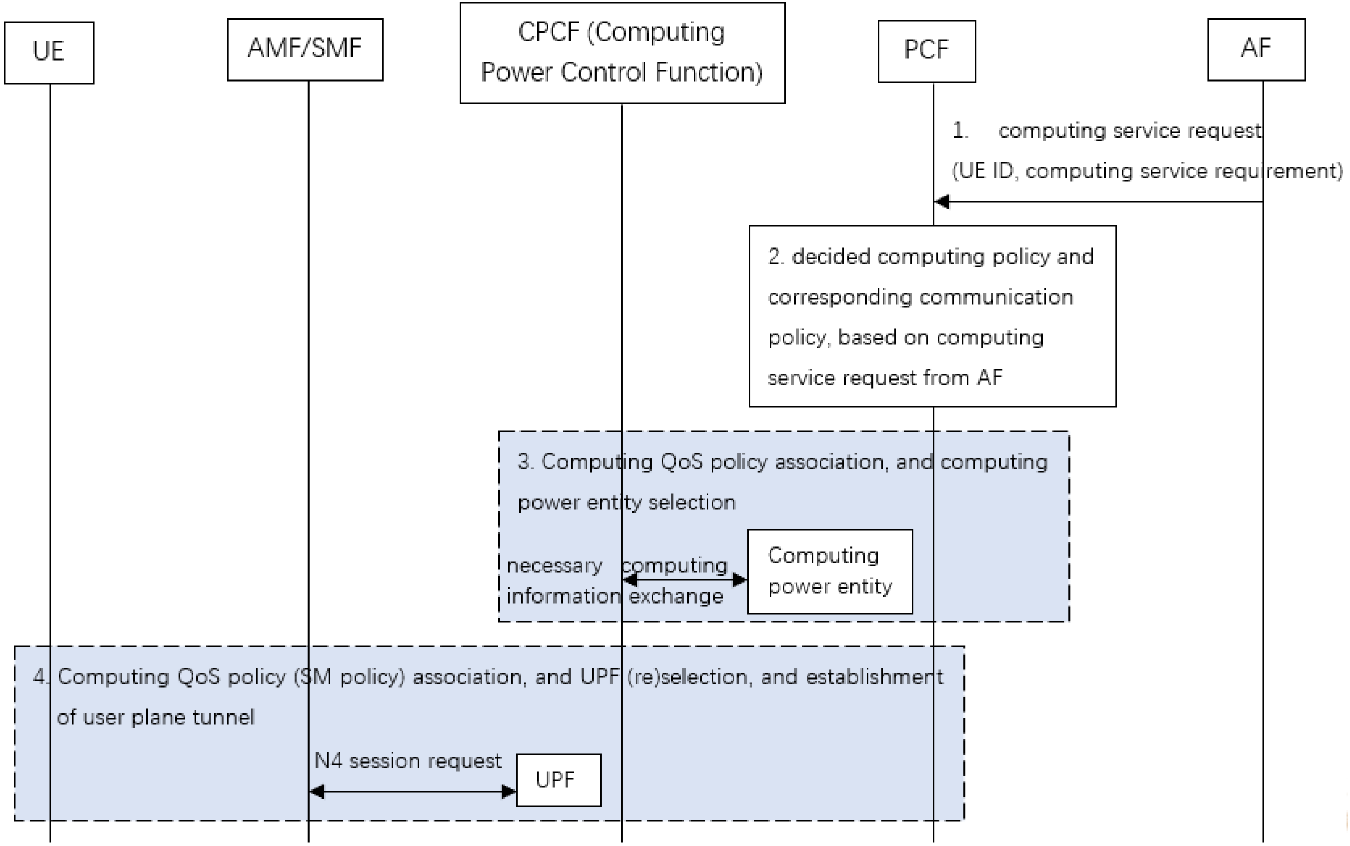


Objective

- Identification and standardization of computing power;
- How to support QoS for computing power；
- How to negotiate and distribute computing power per the service requirements of 5G users, network conditions;
- How to manage the function which jointly optimizes communication resource and computing resources
- Potential impacts on 5GC
  - Whether new computing power control function(s) is needed;
  - Whether new entity(s) providing computing power resource is needed.


Entity providing computing power resource
Entity providing computing power resource
Computing Powering Control Function
Entity providing computing power resource
Entity providing computing power resource

Example



UE

AMF/SMF

CPCF (Computing Power Control Function)

PCF

AF

1. computing service request (UE ID, computing service requirement)

2. decided computing policy and corresponding communication policy, based on computing service request from AF

3. Computing QoS policy association, and computing power entity selection

necessary computing information exchange

Computing power entity

4. Computing QoS policy (SM policy) association, and UPF (re)selection, and establishment of user plane tunnel

N4 session request

UPF

Schematic diagram for call flows

Timeline for study

- Expect to start with other R18 studies depending on SA2 time plan
  - Estimated 8 TU for study phase
- Sent to TSG for information in 2022 March
- Sent to TSG for approval in 2022 June

THANKS

感谢聆听