Consideration for Space-Based Computing Infrastructure Network

draft-wang-space-computing-consideration-00

Jing Wang, China Mobile(Presenter)

Jun Liu, Tsinghua University

Background

- In recent years, the deployment of low-Earth orbit constellations, such as Starlink and OneWeb, has accelerated. The number of satellites has exceeded 10,000, forming the backbone of a global space-based network
- The increasingly sophisticated global satellite network has laid a solid foundation for the expansion of data centers into space

What is the Space-Based Computing Infrastructure Network?

- Stable and reliable satellite links offer efficient interconnection channels for in-orbit data centers
- The widespread deployment of satellites has opened the door for the large-scale deployment of computing infrastructure in space

An infrastructure that enables the interconnection of space-based computing resources via satellite networks

Use Case(1/4)

Emergency Response and Disaster Monitoring

- During natural disasters, such as earthquakes and floods, traditional communication and computing systems are at risk of damage, resulting in delays in the transmission of critical information
  - **Rapidly establish emergency computing and communication nodes to process disaster footage in real time**
  - **Efficiently generate accurate disaster maps and optimal rescue routes**

Use Case(2/4)

Environmental Monitoring and Ecological Management

- The massive raw data collected by satellites needs to be transmitted back to the ground for processing
- Constrained by the bandwidth of satellite–ground communication, less than 10% of the valid data can be transmitted, resulting in extremely low efficiency
    - **By deploying AI models in orbit to perform real-time analysis of remote sensing imagery, only key analysis results are transmitted back to the ground**
    - **Accurately and efficiently identify changes in the status of farmland, forests, water bodies, and glaciers**

Use Case(3/4)

Deep Space Exploration Mission Support

- Deep-space probes often encounter significant communication delays with Earth
  - **Perform in-orbit preprocessing, compression, and intelligent filtering of data collected by satellite detectors**
  - **This can reduce the amount of data that needs to be retransmitted to the ground, thereby effectively improving the response time to satellite detectors**

Use Case(4/4)

In-orbit Training and Inference for Large AI Models

- Training AI models with hundreds of billions of parameters requires a massive amount of electricity
- On-premises data centers face bottlenecks in terms of energy consumption and heat dissipation.
  - **In space, clean energy (solar power) is readily available.**
  - **The low temperatures in space (-270 ℃) facilitate the cooling of in-orbit data centers.**

Requirements(1/2)

Space-Based Computing Resource Monitoring

- The performance of computing devices is constrained by size and power limitations
- In the event of a sudden surge in demand for remote sensing data processing within a region, on-board satellite computing resources are prone to overload
    - **Effective monitoring of on-board computing resources helps ensure their efficient utilization**

Requirements(2/2)

Satellite Routing Considering On-Board Computing Resources Status

- Due to the limited computing capacity of onboard equipment, it is prone to becoming overloaded
- The status of intersatellite links is dynamic
  - **Therefore, on-board traffic scheduling must consider both inter-satellite link status and on-board computing resource status**

Next Steps

- Welcome more people who are interested in this topic!
- Discussions and comments?