Challenges and Opportunities in Space Service Computing

Yan Guo
State Key Lab of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Email: guoyan@bupt.edu.cn

Shangguang Wang
State Key Lab of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Email: sgwang@bupt.edu.cn

***Abstract*—Satellite network has become very popular in recent years. The great development of satellite network motivates the research on satellite-air-ground integrated network to accommodate a variety of services. The concept of space service computing is predicated on moving some services towards space network to harness service capabilities that are currently untapped in space nodes, such as satellites, orbital stations, aircrafts and so on. In this paper, we attempt to sketch a big picture about space service computing with a hope of arousing interest of the research community in this emerging field. We first present a space service computing framework. Compared with the traditional service computing, there are many new challenges for space service computing due to the unique space and aerial environment. Therefore, we also discuss the challenges and opportunities in space service computing.**

***Index Terms*—Space service computing, satellite-air-ground integrated network, satellite, service coordination**

I. Introduction

Satellite network, which can provide global coverage and are not susceptible to natural disasters such as earthquakes, has attracted extensive attention worldwide. In recent years, with the improvement of satellite communication capabilities, the development of recyclable rocket technology and the reduction of manufacturing costs, satellite network has ushered in a second research boom in both academy and industry. For example, SpaceX[1], Amazon[2], Telesat[3], OneWeb[4], and others have proposed satellite programs and started to establish the networks to provide global low-latency broadband Internet. SpaceX proposes the Starlink project and plans to put 12,000 satellites in low Earth orbit [1], [2]. Amazon proposes the Kuiper project and plans to put 3,236 satellites [3], [4]. Telesat plans a constellation with a total of 1,671 satellites [5], and OneWeb plans a constellation with 650 satellites. Until April, 2021, SpaceX and OneWeb has launched 1370 and 182 satellites, respectively. In addition, Google, Facebook, Samsung, and other companies have also joined satellite network research and construction.

As the launch of more satellites, the installment of more ground stations and the improvement of the network infrastructures, the latency, data speed, uptime of communication service

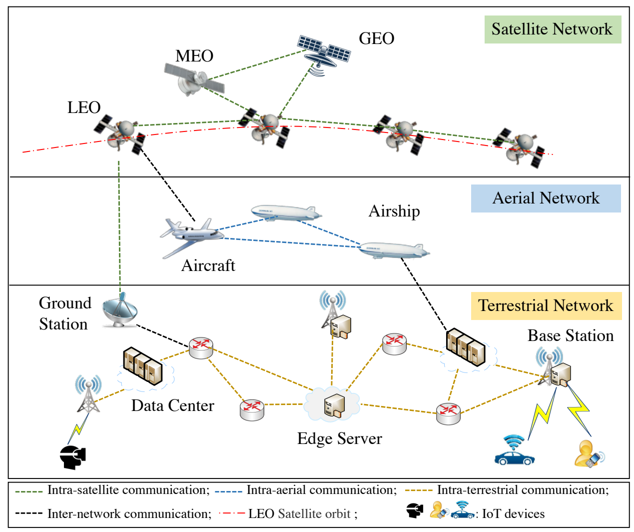

Fig. 1. Network architecture of satellite-air-ground integrated network.

will improve dramatically. Satellite network can deliver high-speed broadband internet to areas of the globe where access is unreliable or completely unavailable. In addition, with the improvement of in-orbit processing and storage, the satellite service capability will also improve.

With the development of Internet of Everything and the recent revolution in wireless devices (e.g., smart wearables, implants and so on), new services are emerging, such as extended reality services, telemedicine, flying vehicles and so on. These services bring higher network requirements on ubiquitous high-speed connectivity, access flexibility, low latency and high reliability. Although the networks have experienced unprecedented growth, there is still a great gap between network requirements and supplements. In addition, the new services in satellite and aerial network are emerging, such as intelligent remote sensing, scientific measurement missions, reconnaissance, emergency rescue and financial transaction, which bring new service requirements. For example, the remote sensing satellite and the weather satellite collect 10TB and 760GB data per day, respectively. To reduce the pressure of too much bandwidth consumption, the on-board data processing is expected. However, due to the insufficient

[1]https://www.spacex.com/webcast
[2]https://www.geekwire.com/2019/amazonprojectkuiperbroadbandsatellite/
[3]https://fcc.report/IBFS/SAT-MPL-20200526-00053/2378318.pdf
[4]https://www.oneweb.world

energy and the limited computing and storage capacities, it is very difficult for satellites to perform sensing, computing and communication simultaneously. The reconnaissance satellites require a second delay which is still a challenge for the present satellite network.

In response to the above trends and challenges, a lot of attention has been paid to satellite-air-ground integrated network (SAGIN). As shown in Fig. 1, SAGIN integrates satellites, aerial platforms and terrestrial communication systems to achieve cross-platform network connectivity [6], [7]. Compared with the terrestrial network, satellite and aerial networks have advantages of wide coverage, long communication distance, large transmission capacity and less dependence on the ground environment. Compared with the satellite and aerial network, terrestrial network has more powerful computing and storage capacities to provide computing-intensive services. Satellite, aerial and terrestrial network can extend and complement each other to satisfy diverse requirements of emerging services. In recent years, more and more researches on SAGIN have been published. For example, they have studied network design [8], [9], cooperative transmission [10], the combination of SAGIN and 5G [11], [12], the combination of SAGIN and mobile edge computing [13], network simulator [14], inter-satellite coordination [15], in-orbit computing [16], ground station design [17]. However, the existing researches on SAGIN are in its primary stage. The most of researchers focus on the network architecture and communication of under layer, rather than the service on the upper layer. There is a lack of a holistic vision for the end-to-end service provision over this integrated network.

To fill the research gaps, we propose service computing in SAGIN, called space service computing. Space service computing moves some services towards space network to harness service capabilities that are currently untapped in space nodes, such as satellite, orbital station and aircraft. Space service computing aims to support a wide range of users and a variety of services, realize the on-demand services across all domains for people, things, platforms, environments and data. Compared with the traditional service computing, there are many new challenges due to the unique space and aerial environment:

- long communication delay resulting from the long distance;
- low reliability due to the interference among communication links;
- link disruption caused by the continuous movement of the non-synchronous medium/low earth orbit satellites;
- dynamic network topologies because of the random join/exit of the unmanned aerial vehicles or airships;
- heterogeneous network protocols of different network infrastructures;
- limited in-orbit computing and storage capabilities.

In this paper, we attempt to sketch a big picture about space service computing with a hope to arouse the interest of research community in this emerging and exciting field.

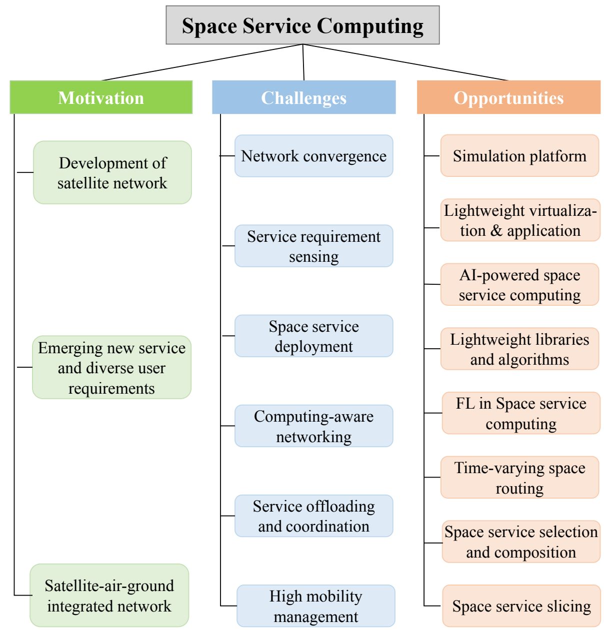

Fig. 2. Motivation, challenges and opportunities in space service computing.

We also discuss some potential research challenges faced and introduce some opportunities for future research in space service computing, which is summarized in Fig. 2.

The remainder of this paper is organized as follows. Section II presents a space service computing framework. Section III presents the research challenges faced in space service computing: network convergence, service requirement sensing, space service deployment, computing-aware networking, service offloading and coordination and high mobility management. Section IV discusses the following eight opportunities for researchers: simulation platform, lightweight virtualization and application, ai-powered space service computing, lightweight libraries and algorithms, FL in space service computing, time-varying space routing, space service selection and composition and space service slicing.

II. Space Service Computing Framework

In this section, we present the framework of space service computing. As shown in Fig.3, the space service computing consists of four tiers. From the top to the bottom, the three tiers are global service plane, service layer, virtualization layer and infrastructure layer, respectively. In this framework, the satellite, aerial and ground network can provide diverse services with different quality, respectively. The composition service across satellite, aerial and ground network is coordinated by the global service plane. The global service plane is also a high level controller to manage all services globally.

The infrastructure layer at the bottom of this framework consists of satellites, high-altitude platforms, aircrafts, airships, unmanned aerial vehicles, base stations, gateways, edge servers, data centers and other device which can provide network, computing and storage resources. Due to the diversified

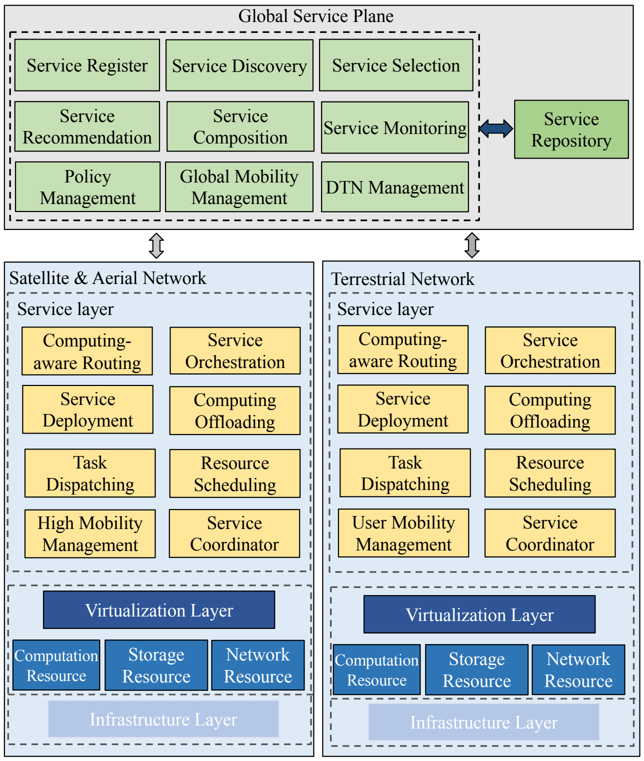

Fig. 3. The framework of space service computing.

and highly distributed infrastructures, the infrastructure layer of space service computing consists of multiple administrative domains. On the top of the infrastructure layer is the virtualization layer. The virtualization layer is based on virtualization technologies. It provides a unified abstraction of the heterogeneous underlying infrastructures [18]. Then the infrastructures are exposed as computation resource, storage resource and network resource, which can be used by the service layer as virtual network service and virtual computing/storage service.

In service layer, service orchestration is to orchestrate the virtual network service and virtual computing/storage service to form composite network-edge service chains. Computing-aware routing is to perceive the network state and computing state and realize computing-aware networking. It can support the service orchestration. Service deployment is to decide whether/when/where to deploy service. The satellites are source-limited and highly mobile. The dynamic service deployment is needed. The frequent redeployment managed by local module is more efficient than by global module. Based on the service deployment information, computing offloading is responsible for offloading computing tasks for users. Task dispatching plays a role of interface to dispatch tasks coming from computing offloading. Then the resources are scheduled according to the dispatching scheme in resource scheduling module. Both satellite and aerial nodes and terrestrial nodes are all faced with mobility. Specifically, the mobility challenge of terrestrial nodes mainly due to the user movement, such as mobile devices, vehicles and so on. Nevertheless, the mobility challenge faced by satellite and aerial network are more complex due to the high-speed mobility of the satellite and the user movement. Service coordinator is in charge of coordinating with all services during service provision for high quality of services and user experience.

The global service plane is also a high level controller to manage services across network domains and enable the composition service across satellite, aerial and ground network. The global control plane locates at a cloud data center with sufficient resources and serves as a central controller of the space service system. It is composed of service repository and several system control modules. Service repository contains service and application images which has been registered. When the requests from service deployment module are delivered, the corresponding images are pulled to start a service function instance. When a new service is deployed in space service computing, the first step is to register itself to the service register module. Then the services can be discovered, selected and recommended on demand. For the complex user requests, the services in satellite, aerial and ground network are composed in service composition module to meet the functional and nonfunctional requirements. Due to the large scale of services, the efficiency is very important for service composition module. Service monitoring is in charge of monitoring the run state of each service and the success rate of service, especially composed services. Once the success rate of service is low, the corresponding service schemes are adjusted. As for policy management, it runs the control plane of SDN devices and ensure the data transmission across heterogeneous networks. Besides, delay tolerate network management module works to help edge nodes to deal with situations with different latencies. Different from the goal of minimum delay in traditional service computing, how to select and compose services with determined delay is a new problem in space service computing.

III. Challenges

For space service computing, there are many new challenges due to the unique space and aerial environment. In this section, we consider the following six research challenges which will need to be addressed.

A. Network Convergence

The traditional satellite network, the aerial network and the ground network are independent, which leads to poor interconnection and inefficient service provision. Compared with the traditional service computing based on a single ground network, the space service computing is based on heterogeneous networks, i.e., the satellite network, the aerial network and the ground network. A key fundamental issue is network convergence by deeply integrating the independent heterogeneous networks to achieve wide-area coverage, interconnection of networks and efficient services. Different network paradigms are supported by various communication standards, and equipped with different types of network devices. The integrated network is of multi-architecture, multi-protocol and multi-data formats.

The specific network architectures need to be integrated include TCP/IP network, Delay Tolerance Network (DTN), information centric network and smart identifier network. DTN has attracted much attention since it can effectively control delays, and has been used in military battlefield networks, sparse sensor networks and other networks [19]. Recently, DTN has been regarded as a promising solution with service to satellite network [20]. The novel message store-carry-forward mechanism in DTN can reduce intermittent connection, long transmission delay and high bit error rate of space service. In these network architectures, the heterogeneous communication protocols include IP protocol, space packet protocol, space communication protocol and other protocols. Therefore, the integrated network needs to ensure that data interaction adapts to network protocols, by achieving multi-type routing addressing and multi-data format switching [21], [22]. It is challenging to realize seamless data transmission and efficient information interaction between independent networks, and then improve the quality of service and user experience.

B. Service Requirement Sensing

Space service computing covers a wide range of fields and scenes. Different from 5G which divides scenes into three categories: enhance mobile broadband, ultra reliable & low latency and massive machine type communication [23], the scenes in space service computing are more refined and the service requirements are more diverse and complex coupling. For example, there are latency-sensitive service, latency-deterministic service, connection-oriented service, high-bandwidth service, high-reliability service and the convergence service [24].

The diverse service requirements are created by not only new applications in stand-alone networks, but also the new applications in integrated network. For satellite and aerial network, constrained by technical and economic bottlenecks at the early time, it can only be applied to a number of specific fields such as space exploration, earth observation, emergency communication and broadcasting services. In recent years, with new advanced technologies emerging, satellite and high-altitude platforms can provide high data throughput, processing and storage capability. The new applications are emerging, such as intelligent remote sensing [25], environment monitoring, forest-fire prevention [26], emergency rescue [27] and financial transaction [28], which bring new service requirements. Meanwhile, the new applications in ground network, such as Internet of thing, augmented reality/virtual reality, 4K/8K, also promote the growing diversity of service requirements. On the other hand, with the further integration of networks, there will be many applications beyond expectations, such as holographic communication, advanced intelligent industry and Internet of everything. The service requirements of these new applications are more complex coupling. Due to diverse new service requirements, it is extremely difficult to serve different application scenes flexibly.

C. Space Service Deployment

Although satellite and aerial network is growing fast, the computing and storage resource are still limited. Due to the high mobility of low earth orbit satellite and imbalance distribution of resource, space service deployment faces various challenges in terms of deployment technology, dynamic deployment and coordination deployment.

Specifically, in the aspect of deployment technology, the light-weighted deployment technology is needed for flexibility. The common virtual machine and container based service deployment is not resource efficient even not available. It also leads to high deployment cost due to ongoing maintenance.

Dynamic space service deployment is a promising approach to improve the resource efficiency and quality of service. It is not easy to adjust service deployment under the dynamic spatiotemporal service distribution and the motion of satellites, unmanned aerial vehicles and high-altitude platforms.

The coordination deployment refers to two folds: the coordination among satellites, high-altitude platforms and ground computing nodes and the coordination over satellite constellation. The storage and computation resources are scarce on satellites and the aerial nodes, which makes it difficult to provide intensive computation service. On the other hand, ground network have more resources to provide sophisticated service, but have limited coverage. The coordination service deployment can take full advantage of heterogenous networks by deploying different service in different service. The capacity of a single satellite is limited, while the capacity of the low earth orbit satellite constellation is strong. In addition, the communication condition between satellites in satellite constellation is much better than that between satellite and ground. The coordination over satellite constellation enables to improve the quality of service. However, the connectivity state between satellites and ground network and the inter-satellite topology are constantly changing. Therefore, it is still challenging for coordination service deployment in space service computing.

D. Computing-aware Networking

From the perspective of network, the computing capacity of each satellite is limited, and the satellites lack mutual awareness, the computing tasks can not be scheduled to the optimal satellite for service. In existing service system, computing tasks are generally managed through centralized orchestration. However, the orchestration performance and scalability of the centralized architecture is not guaranteed in the large-scale satellite network and space service computing.

From the perspective of service requirement, the existing application layer and network layer in space service computing are decoupling and it is difficult to perceive varying real-time network status for the application layer. The scheduling strategy only consider the computing capacity without network state may lead to the unbalanced network load and low quality of service. It is difficult to take advantage of satellite network and ubiquitous access. For example, due to the lack of computing-aware networking, the service with the lowest communication delay may not be the optimal due to the long computing delay. The computing tasks can only be scheduled to the optimal service by taking into account the real-time network status and computing status at the same time. Through computing-aware networking, massive users can invoke computing and network resources in different places on demand and in real time to improve the efficiency of computing and network utilization. In addition, it is still essential to differentiate the services and flexibly match the network resources to diverse service requirements. However, it is still challenging to perceive the network state and computing state and realize computing-aware networking in high dynamic large-scale satellite network.

E. Service Offloading and Coordination

In space service computing, it is difficult for a single satellite and aerial node to provide the whole complex service due to the limited computing resources. On the other hand, the distribution of service requests are imbalance in time and space. The computing load of satellite, aerial and ground node is also imbalance in both time and space. For example, some satellites are overload while the other satellites are idle. Therefore, the efficient service offloading and coordination is essential to improve service efficiency and resource utilization, but it also brings many challenges.

The service offloading and coordination contains the service selection, service recommendation and service composition. Service selection is to select the optimal service for offloading. Service recommendation and composition is for service coordination and enable an end-to-end space service for users. Different from the tradition service computing, the service selection, service recommendation and service composition in space service computing face many new challenges due to the unique space environment. For example, the high mobility of satellites calls for dynamic service composition. In the same time, the low delay requirements calls for the efficiency service composition algorithms over a larger scale of candidate services.

The other challenges of service offloading and coordination are in terms of coordination offloading and dynamic coordination scheduling. Firstly, there are multiple options to offload computing for users in space service computing, such as locally processing, satellites, unmanned aerial vehicles, high-altitude platforms, edge nodes and data centers in ground network. Different options have different characteristics. For example, data centers can process computing-intensive tasks with high reliability but longer transmission delay. Edge nodes are closer to users but limited computing capacity and scope. Satellites has greater coverage but longer transmission delay and unreliable computing resource. Moreover, different users have different requirements. For example, some services are delay-sensitive, some require high reliability, some require low cost, and so on. Therefore, when developing a cooperative computation offloading strategy, the user requirements and the characteristics of options need to be considered. Secondly, since the users and satellites are mobile and the computing load are dynamic, dynamic coordination scheduling can further improve the service efficiency. Compared with the edge nodes and data centers in ground network, the coverage of SAGIN is relatively larger, and the centralized scheduling may result in low efficiency and poor resource utilization.

F. High Mobility Management

Except for the mobile users, the core infrastructures (i.e., satellites and high-altitude platforms) in space service computing are mobile with high speed. The mobility of satellites and end users are intertwined thus making mobility management more challenging. Due to the high mobility of non-synchronous orbit satellites, the connection between satellites and ground network is constantly changing. Not only does the user link change, but also the feeder link changes, which results in service disruption and then reduces the quality of user experience. The frequent switching of satellite-ground links also can lead to increased administrative overhead. Therefore, how to design switching strategy to minimize switching frequency and reduce overhead is a key research. Moreover, the receivers are not limited to fixed ground station, and it may be mobile vehicles, unmanned aerial vehicles, ships and aircrafts. The mobile receivers impose more challenges on the mobility management.

The movement of low earth orbit satellites also brings about changes in the inter-satellite topology and the re-selection of inter-satellite path. In the low earth orbit satellite constellation, the shortest routes between satellites change at least once every 17.753 seconds on average. A change of the shortest route affects an average of 51.528 inter-satellite links. In such dynamic context, how to select the inter-satellite path to ensure the communication reliability is non-trivial. Moreover, routing and service migration are challenging approaches for service continuity under varying network conditions.

IV. Opportunities

Despite challenges that arise when realizing space service computing, there are numerous opportunities for academic research. In this section, we identify eight such opportunities.

A. Simulation Platform

Due to the high cost of satellite manufacturing and launch, it is very necessary to set up a simulation platform of real satellite environment for researchers. The simulation platform for space service computing should contain computing module, power and energy module, communication module and other modules. The computing module is a computing system to support specific application such as remote sensing image processing. The computing capacity may be affected by the space environment such as extreme temperature, sunspot and electromagnetic interference. The power and energy module is responsible for the solar-activity-dependent energy harvesting, energy storage and energy consumption caused by communication and computing. And the communication module is accountable for simulations of the inter-satellite communication links and the ground-satellite communication links.

Nowadays, Systems Tool Kit [29] is a common simulation software in aerospace field. It can simulate the movement of satellites, stars, planets, airplane, ship and other targets. It can also help users to establish orbits, such as geosynchronous orbit, low earth orbit and so on. Then it can calculate the position and attitude of the satellite at any moment, the coverage area of the satellite and the ground station. However, it can not simulate the communication and in-orbit computing. In [30], the authors develop the first orbital edge computing simulator and runtime service, called cote. cote models orbital mechanics and Earth rotation to track ground station and satellite positions over time. It also models the data collection, communication latency and the energy consumption. However, it does not involve the inter-satellite communication and the models are ideal without considering the complex space environment. Therefore, the simulation platform for space service computing is still to be explored.

B. Lightweight Virtualization and Application

Research in lightweight virtualization and application architecture can provide inroads to tackling challenges related to service deployment on satellites and aerial nodes. Unlike large server in ground network, the satellites and aerial nodes do not have substantial resource. The lightweight resource virtualization is needed for high resource utilization. Container technologies are mature lightweight virtualization in ground servers [31]. Serverless is an emerging technology to deploy applications [32]. More research is required to adopt containers and serverless or develop a new lightweight virtualization technology as a suitable mechanism for service deployment on satellites and aerial nodes. The corresponding lightweight virtualized resource management tools, such as Kubernetes[5] and KubeEdge[6] for Docker containers[7], are required.

Based on the lightweight virtualization technology, the quick application deployment is operable. The lightweight application architecture is also needed. Microservice, function as a service are emerging lightweight application architecture [33]. They decompose applications to finer-grained service components which can be deployed and executed independently. However, how to partition applications to a set of microservices or functions, how to deploy dependent microservices and functions dynamically, how to register and manage the highly distributed microservices and functions and how to allocate resource for these large-scale microservices or functions are still to be researched.

C. AI-powered Space Service Computing

In space service computing, there are large-scale heterogeneous networks and a variety of services. The complexity of network management and service orchestration call for the autonomous network and service management [34]. For automation of network and service management, the ETSI proposed zero-touch network and service management framework as a next-generation management system. In this framework, all operational processes and tasks will be executed automatically. Artificial Intelligence (AI) is envisioned as a promising solution to make this zero-touch network and service management framework a reality [35]. Based on machine learning and big data analytics techniques, AI-powered autonomous network and service management can reduce the operational delay and costs, accelerate time-to-value and reduce the risk of human error. Nevertheless, there may be some potential limitations and risks of using AI techniques in the autonomous network and service management framework. And how to leverage AI technologies in this field is still to be explored.

D. Lightweight Libraries and Algorithms

Due to the limited computing and storage capacities of satellites and aerial nodes, they will not support heavyweight softwares. For example, the complex data processing tool Apache Spark[8] requires at least 8 cores CPU and 8 gigabyte memory for good performance. And the limited CPU/GPU and memory are not be sufficient to execute remote sensing data processing using this heavyweight tool. Therefore, the satellites and aerial nodes require lightweight libraries and algorithms for data processing and machine learning.

A lightweight library, Apache Quarks[9], is proposed to be employed on mobile devices for real-time data analytics. However, it can only support basic data processing. With the development of machine learning on device, many lightweight deep learning frameworks for training and inference on device have been proposed, for example, MNN[10] from Ali and NCNN[11] from Tencent. However, the potential of these framework for deep learning on satellites and aerial nodes is still to be explored.

E. FL in Space Service Computing

As we all known, the satellites are belong to different companies and organizations, such as SpaceX, OneWeb, Telesat, Amazon, countries and so on. The privacy of satellite data should be considered in AI-powered space service computing. Federated Learning (FL) is a distributed machine learning framework, which was first proposed by Google in 2016. The goal of federated learning is to achieve joint modeling of distributed devices and improve the effect of AI models while ensuring data privacy, security and legal compliance. Federated Learning is wide used in machine learning on mobile devices for training data locally. Similar with the mobile devices in ground network, the satellites are resource-limited and request for data security and privacy. FL seemed like a promising solution for distributed machine learning in space service computing to ensure data privacy, security and legal compliance. Therefore, the FL in space service computing is worth researched.

[5]https://kubernetes.io
[6]https://kubernetes.io/blog/2019/03/19/kubeedge-k8s-based-edge-intro/
[7]https://www.docker.com
[8]http://spark.apache.org
[9]http://quarks.incubator.apache.org
[10]https://github.com/alibaba/MNNKit
[11]https://github.com/Tencent/ncnn

F. Time-varying Space Routing

Routing is a fundamental technology to support data transmission and service. The Routing path in space service computing can be divided into ground path, ground-satellite path and inter-satellite path. The ground-satellite path and inter-satellite path are time-varying due to the dynamics of the satellite network's topology, the intermittence of inter-satellite communication links and the mobility-induced satellite-access switching of mobile terminals. Therefore, the traditional static routing method is not fit for space service computing. In satellite network, the multi-path routing [36], congestion avoidance routing algorithm [37] and secure routing [38] have been proposed.

To support reliable and low-latency data transmission and service, time-varying space routing is required. Specially, the ground-satellite path selection problem can be formulated to the dynamic mapping between ground stations and satellites. The load distribution of ground station, delay, bandwidth, movement of satellites should be considered. The inter-satellite path selection is mainly depend on the time-varying network topology. Many technical problems must be solved to achieve effective inter-satellite path selection, for example, the construction of an accurate network model to present time-varying topologies; efficient time-varying routing algorithms; load balancing among satellites.

G. Space Service Selection and Composition

Whether in traditional service computing or in space service computing, service selection, service recommendation and service composition are important research issues. For service selection in space service computing, we should consider not only the quality of service such as latency, throughput, reliability, cost and other factors, but also consider the service duration time since the satellites or aircrafts may move away from the users when service are running. The dynamic service selection and composition are needed for this situation. In addition, DTN has been regarded as a promising solution with service to satellite network. Different from the goal of minimum delay in traditional service computing, how to select and compose services with determined delay is a new problem in space service computing.

H. Space Service Slicing

Due to the diverse new service requirements, research on serving different application scenes flexibly under the same service system is important. Network slicing based NFV and SDN is the technology to adapt the various QoS requirements [39]. With virtualization technologies, the heterogeneous physical resources from different network segments are abstracted into unified virtual resource pools. Then with network slicing, a physical network could be sliced into multiple network slices. Recently, NFV and SDN based SAGIN system has been proposed [40], where NFV and SDN are applied to manage the abstracted virtual resource flexibly [41].

Similar with network slicing based on NFV and SDN, we propose service slicing based on network slicing and service coordination. Service slicing aims to provides the ability to deliver highly customizable services by utilizing the sharing underlying infrastructure and the microservices or functions deployed across multiple domains. Service slicing may span across multiple technological as well as administrative domains. Specifically, space service slicing includes service deployment, service coordination and dynamic service function chaining. For service deployment in space service computing, how to deploy specific services according to the different characteristics of geosynchronous earth orbit satellites, medium earth orbit satellites and low earth orbit satellites is problem still to be resolved. The high mobility of satellites is a double-edged sword. The high mobility causes the difficulties of space service deployment, meanwhile the high mobility contributes to the deployment of some services and enable the global coverage of some services.

Although service function chaining in SAGIN has been researched [41], [42], many technical problems must be solved to achieve effective inter domain service slicing. For example, how to analyze the service requirement intelligently, how to perceive the network state and computing state and realize computing-aware networking, dynamic space service deployment and coordination, coordination offloading, servicemesh in space service computing, end-to-end network slicing across multiple administrative domains are all opening problems that need further study.

V. CONCLUSIONS

The emerging satellite network and the improvement of in-orbit processing and storage enhance service capability of satellites. Although SAGIN has attracted a lot of attention, the existing researches on SAGIN are in its primary stage. To meet the diverse service requirements of varieties of users, this paper proposes the concept of space service computing. Then we present a space service computing framework which contains global service plane, service layer, virtualization layer and infrastructure layer. And then we presents six research challenges and six rewarding opportunities in space service computing. The six research challenges faced in space service computing are network convergence, service requirement sensing, space service deployment, computing-aware networking, service offloading and coordination and high mobility management. The eight opportunities for academic research are simulation platform, lightweight virtualization and application, AI-powered space service computing, lightweight libraries and algorithms, FL in space service computing, time-varying space routing, space service selection and composition and space service slicing. In our future work, we will also study the service computing in Interstellar Network 1 or Interstellar Network First.

ACKNOWLEDGEMENT

This work was supported in part by Key-Area Research and Development Program of Guangdong Province (No. 2020B010164002) and National Science Foundation of China (Grant 62032003,61922017, and 61921003).

References

[1] SpaceX FCC update, "Spacex non-geostationary satellite system," 2018, https://licensing.fcc.gov/myibfs/download.do?attachment_key=1569860.

[2] S. Starlink, 2017, https://www.spacex.com/webcast.

[3] A. Boyle, "Amazon to offer broadband access from orbit with 3,236satellite "project kuiper" constellation," 2019, https://www.geekwire.com/2019/amazonprojectkuiperbroadband-satellite/.

[4] Kuiper USASAT-NGSO-8C ITU filing, 2018, https://www.itu.int/ITU-R/space/asreceived/Publication/DisplayPublication/8718.

[5] Telesat, "Application for modification of market access authorization," 2020, https://fcc.report/IBFS/SAT-MPL-20200526-00053/2378318.pdf.

[6] J. Liu, Y. Shi, Z. M. Fadlullah, and N. Kato, "Space-air-ground integrated network: A survey," *IEEE Communications Surveys and Tutorials*, vol. 20, no. 4, pp. 2714–2741, 2018.

[7] N. Cheng, W. Xu, W. Shi, Y. Zhou, N. Lu, H. Zhou, and X. Shen, "Air-ground integrated mobile edge networks: Architecture, challenges, and opportunities," *IEEE Communication Magazine*, vol. 56, no. 8, pp. 26–32, 2018.

[8] N. Zhang, S. Zhang, P. Yang, O. Alhussein, W. Zhuang, and X. S. Shen, "Software defined space-air-ground integrated vehicular networks: Challenges and solutions," *IEEE Communications Magazine*, vol. 55, no. 7, pp. 101–109, 2017.

[9] J. Liu, X. Du, J. Cui, M. Pan, and D. Wei, "Task-oriented intelligent networking architecture for the space-air-ground-aqua integrated network," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 5345–5358, 2020.

[10] X. Zhu, C. Jiang, L. Kuang, N. Ge, S. Guo, and J. Lu, "Cooperative transmission in integrated terrestrial-satellite networks," *IEEE Network*, vol. 33, no. 3, pp. 204–210, 2019.

[11] L. Boero, R. Bruschi, F. Davoli, M. Marchese, and F. Patrone, "Satellite networking integration in the 5G ecosystem: Research trends and open challenges," *IEEE Network*, vol. 32, no. 5, pp. 9–15, 2018.

[12] G. Giambene, S. Kota, and P. Pillai, "Satellite-5G integration: A network perspective," *IEEE Network*, vol. 32, no. 5, pp. 25–31, 2018.

[13] R. Xie, Q. Tang, Q. Wang, X. Liu, F. R. Yu, and T. Huang, "Satellite-terrestrial integrated edge computing networks: Architecture, challenges, and open issues," *IEEE Network*, vol. 34, no. 3, pp. 224–231, 2020.

[14] S. Kassing, D. Bhattacherjee, A. B. Águas, J. E. Saethre, and A. Singla, "Exploring the "internet from space" with hypatia," in *Proceeding of ACM Internet Measurement Conference (IMC)*, 2020, pp. 214–229.

[15] Y. Hauri, D. Bhattacherjee, M. Grossmann, and A. Singla, ""internet from space" without inter-satellite links?" in *Proceedings of the 19th ACM Workshop on Hot Topics in Networks (HotNets)*, 2020, pp. 205–211.

[16] D. Bhattacherjee, S. Kassing, M. Licciardello, and A. Singla, "In-orbit computing: An outlandish thought experiment," in *Proceedings of the 19th ACM Workshop on Hot Topics in Networks (HotNets)*, 2020, pp. 197–204.

[17] D. Vasisht and R. Chandra, "A distributed and hybrid ground station network for low earth orbit satellites," in *Proceedings of the 19th ACM Workshop on Hot Topics in Networks (HotNets)*. ACM, 2020.

[18] Y. Guo, Q. Duan, and S. Wang, "Service orchestration for integrating edge computing and 5G network: State of the art and challenges," in *Proceeding of IEEE World Congress on Services (SERVICES)*. IEEE, 2020, pp. 55–60.

[19] C. Caini, H. S. Cruickshank, S. Farrell, and M. Marchese, "Delay and disruptiontolerant networking (DTN): an alternative solution for future satellite networking applications," *Proc. IEEE*, vol. 99, no. 11, pp. 1980–1997, 2011.

[20] C. Caini and R. Firrincieli, "DTN for LEO satellite communications," in *Proceedings of the third International Conference on Personal Satellite Services (PSATS)*, vol. 71, 2011, pp. 186–198.

[21] S. Claypool, J. Chung, and M. Claypool, "Comparison of tcp congestion control performance over a satellite network," in *Proceedings of Passive and Active Measurement (PAM)*, 2021, pp. 499–512.

[22] M. Luglio, M. Quadrini, S. P. Romano, C. Roseti, and F. Zampognaro, "Enabling an efficient satellite-terrestrial hybrid transport service through a quic-based proxy function," in *Proceedings of International Symposium on Networks, Computers and Communications (ISNCC)*, 2020, pp. 1–6.

[23] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5G wireless networks: A comprehensive survey," *IEEE Communications Surveys and Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.

[24] J. Hu, L. Cai, C. Zhao, and J. Pan, "Directed percolation routing for ultra-reliable and low-latency services in low earth orbit (LEO) satellite networks," in *Proceedings of the 92nd IEEE Vehicular Technology Conference (VTC)*, 2020, pp. 1–6.

[25] Y. Li, Y. Iwamoto, L. Lin, and Y.-W. Chen, "Parallel-connected residual channel attention network for remote sensing image super-resolution," in *Proceedings of Asian Conference on Computer Vision (ACCV) Workshops*, I. Sato and B. Han, Eds., 2020, pp. 18–30.

[26] L. Shumilo, B. Yailymov, and A. Shelestov, "Active fire monitoring service for ukraine based on satellite data," in *Proceedings of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2020, pp. 2913–2916.

[27] T. Lei, H. Cheng, A. Li, W. Qu, Z. Pang, J. Fu, L. Li, X. Li, and J. Lu, "Cooperative emergency monitoring and assessment of flood disasters based on the integrated ground-air-space remote sensing," in *Proceeding of IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, 2019, pp. 9733–9736.

[28] D. Bhattacherjee, W. Aqeel, G. Laughlin, B. M. Maggs, and A. Singla, "A bird's eye view of the world's fastest networks," in *Proceeding of ACM Internet Measurement Conference (IMC)*, 2020, pp. 521–527.

[29] Y. Liu, Q. Zhang, S. Han, and D. Zou, "A GEO satellite positioning testbed based on labview and STK," in *Proceedings of the 8th International Conference on Communications, Signal Processing, and Systems (CSPS)*, ser. Lecture Notes in Electrical Engineering, vol. 571, 2019, pp. 1892–1899.

[30] B. Denby and B. Lucia, "Orbital edge computing: Nanosatellite constellations as a new class of computer system," in *Proceedings of Architectural Support for Programming Languages and Operating Systems (ASPLOS)*, 2020, pp. 939–954.

[31] W. Wang, L. Zhang, D. Guo, S. Wu, H. Cui, and F. Bi, "Reg: An ultra-lightweight container that maximizes memory sharing and minimizes the runtime environment," in *Proceeding of IEEE International Conference on Web Services (ICWS)*, 2019, pp. 76–82.

[32] J. R. Gunasekaran, P. Thinakaran, N. C. Nachiappan, R. S. Kannan, M. T. Kandemir, and C. R. Das, "Characterizing bottlenecks in scheduling microservices on serverless platforms," in *Proceedings of 40th IEEE International Conference on Distributed Computing Systems (ICDCS)*, 2020, pp. 1197–1198.

[33] S. Wang, Y. Guo, N. Zhang, P. Yang, A. Zhou, and X. Shen, "Delay-aware microservice coordination in mobile edge computing: A reinforcement learning approach," *IEEE Transactions on Mobile Computing*, vol. 20, no. 3, pp. 939–951, 2021.

[34] C. Ding, A. Zhou, X. Ma, and S. Wang, "Cognitive service in mobile edge computing," in *Proceedings of IEEE International Conference on Web Services (ICWS)*, 2020, pp. 181–188.

[35] C. Benzaid and T. Taleb, "Ai-driven zero touch network and service management in 5G and beyond: Challenges and research directions," *IEEE Network*, vol. 34, no. 2, pp. 186–194, 2020.

[36] P. Liu and T. Zhang, "Constrained multipath routing algorithm based on satellite network," in *Proceedings of Communications and Networking (ChinaCom)*, H. Gao, P. Fan, J. Wun, X. Xiaoping, J. Yu, and Y. Wang, Eds., 2021, pp. 381–391.

[37] Y. Kuang, X. Yi, and Z. Hou, "Congestion avoidance routing algorithm for topology-inhomogeneous low earth orbit satellite navigation augmentation network," *International Journal of Satellite Communications and Networking*, vol. 39, no. 2, pp. 221–235, 2021.

[38] Y. Yan, G. Han, and H. Xu, "A survey on secure routing protocols for satellite network," *Journal of Network and Computer Applications*, vol. 145, 2019.

[39] J. Ordonez-Lucena, P. Ameigeiras, D. López, J. J. Ramos-Muñoz, J. Lorca, and J. Folgueira, "Network slicing for 5G with SDN/NFV: concepts, architectures, and challenges," *IEEE Communications Magazine*, vol. 55, no. 5, pp. 80–87, 2017.

[40] H. Wu, J. Chen, C. Zhou, W. Shi, N. Cheng, W. Xu, W. Zhuang, and X. S. Shen, "Resource management in space-air-ground integrated vehicular networks: SDN control and AI algorithm design," *IEEE Wireless Communications*, vol. 27, no. 6, pp. 52–60, 2020.

[41] G. Wang, S. Zhou, S. Zhang, Z. Niu, and X. Shen, "SFC-based service provisioning for reconfigurable space-air-ground integrated networks," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 7, pp. 1478–1489, 2020.

[42] B. Feng, G. Li, G. Li, Y. Zhang, H. Zhou, and S. Yu, "Enabling efficient service function chains at terrestrial-satellite hybrid cloud networks," *IEEE Network*, vol. 33, no. 6, pp. 94–99, 2019.