



Experiential Networked Intelligence (ENI); Space-Ground Cooperative Network Slicing

Disclaimer

The present document has been produced and approved by the Experiential Networked Intelligence (ENI) ETSI Industry Specification Group (ISG) and represents the views of those members who participated in this ISG.
It does not necessarily represent the views of the entire ETSI membership.

Reference

DGR/ENI-0036V411_SGC_NetSlicin

Keywords

network, slicing

ETSI

650 Route des Lucioles
F-06921 Sophia Antipolis Cedex - FRANCE

Tel.: +33 4 92 94 42 00 Fax: +33 4 93 65 47 16

Siret N° 348 623 562 00017 - APE 7112B
Association à but non lucratif enregistrée à la
Sous-Préfecture de Grasse (06) N° w061004871

Important notice

The present document can be downloaded from the
[ETSI Search & Browse Standards](#) application.

The present document may be made available in electronic versions and/or in print. The content of any electronic and/or print versions of the present document shall not be modified without the prior written authorization of ETSI. In case of any existing or perceived difference in contents between such versions and/or in print, the prevailing version of an ETSI deliverable is the one made publicly available in PDF format on [ETSI deliver](#) repository.

Users should be aware that the present document may be revised or have its status changed,
this information is available in the [Milestones listing](#).

If you find errors in the present document, please send your comments to
the relevant service listed under [Committee Support Staff](#).

If you find a security vulnerability in the present document, please report it through our
[Coordinated Vulnerability Disclosure \(CVD\)](#) program.

Notice of disclaimer & limitation of liability

The information provided in the present deliverable is directed solely to professionals who have the appropriate degree of experience to understand and interpret its content in accordance with generally accepted engineering or other professional standard and applicable regulations.

No recommendation as to products and services or vendors is made or should be implied.

No representation or warranty is made that this deliverable is technically accurate or sufficient or conforms to any law and/or governmental rule and/or regulation and further, no representation or warranty is made of merchantability or fitness for any particular purpose or against infringement of intellectual property rights.

In no event shall ETSI be held liable for loss of profits or any other incidental or consequential damages.

Any software contained in this deliverable is provided "AS IS" with no warranties, express or implied, including but not limited to, the warranties of merchantability, fitness for a particular purpose and non-infringement of intellectual property rights and ETSI shall not be held liable in any event for any damages whatsoever (including, without limitation, damages for loss of profits, business interruption, loss of information, or any other pecuniary loss) arising out of or related to the use of or inability to use the software.

Copyright Notification

No part may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying and microfilm except as authorized by written permission of ETSI.

The content of the PDF version shall not be modified without the written authorization of ETSI.

The copyright and the foregoing restriction extend to reproduction in all media.

© ETSI 2025.
All rights reserved.

Contents

Intellectual Property Rights	4
Foreword.....	4
Modal verbs terminology.....	4
Introduction	4
1 Scope	5
2 References	5
2.1 Normative references	5
2.2 Informative references.....	5
3 Definition of terms, symbols and abbreviations.....	6
3.1 Terms.....	6
3.2 Symbols.....	6
3.3 Abbreviations	6
4 Overview	6
4.1 Introduction	6
4.2 Architecture	7
4.2.1 Space-Ground Cooperative Network Slicing Architecture	7
4.2.2 Space-ground Slicing Session Collaboration	7
4.2.3 Intelligent Slice Mapping.....	11
4.2.4 Intelligent slicing technology for space-ground collaborative network resources on demand	14
5 Network Structure	17
5.1 Forward and Backhaul Link	17
5.2 Identification and restriction of satellite access types	17
5.3 UE location Identification	18
6 Mathematical model.....	20
6.1 Mathematical model introduction.....	20
6.2 Mathematical model based on the slices of the optimal weighted graph matching.....	22
6.3 Simulation and verification of network resource slicing generation technology.....	22
6.4 Network resource slicing scheduling demand forecasting technology.....	24
6.4.1 Introduction.....	24
6.4.2 Mathematical model for slice prediction based on spatiotemporal correlation.....	24
6.4.3 Design of network resource slice migration scheme.....	25
7 Conclusion and recommendations.....	26
History	28

Intellectual Property Rights

Essential patents

IPRs essential or potentially essential to normative deliverables may have been declared to ETSI. The declarations pertaining to these essential IPRs, if any, are publicly available for **ETSI members and non-members**, and can be found in ETSI SR 000 314: *"Intellectual Property Rights (IPRs); Essential, or potentially Essential, IPRs notified to ETSI in respect of ETSI standards"*, which is available from the ETSI Secretariat. Latest updates are available on the [ETSI IPR online database](#).

Pursuant to the ETSI Directives including the ETSI IPR Policy, no investigation regarding the essentiality of IPRs, including IPR searches, has been carried out by ETSI. No guarantee can be given as to the existence of other IPRs not referenced in ETSI SR 000 314 (or the updates on the ETSI Web server) which are, or may be, or may become, essential to the present document.

Trademarks

The present document may include trademarks and/or tradenames which are asserted and/or registered by their owners. ETSI claims no ownership of these except for any which are indicated as being the property of ETSI, and conveys no right to use or reproduce any trademark and/or tradename. Mention of those trademarks in the present document does not constitute an endorsement by ETSI of products, services or organizations associated with those trademarks.

DECT™, **PLUGTESTS™**, **UMTS™** and the ETSI logo are trademarks of ETSI registered for the benefit of its Members. **3GPP™**, **LTE™** and **5G™** logo are trademarks of ETSI registered for the benefit of its Members and of the 3GPP Organizational Partners. **oneM2M™** logo is a trademark of ETSI registered for the benefit of its Members and of the oneM2M Partners. **GSM®** and the GSM logo are trademarks registered and owned by the GSM Association.

Foreword

This Group Report (GR) has been produced by ETSI Industry Specification Group (ISG) Experiential Networked Intelligence (ENI).

Modal verbs terminology

In the present document "**should**", "**should not**", "**may**", "**need not**", "**will**", "**will not**", "**can**" and "**cannot**" are to be interpreted as described in clause 3.2 of the [ETSI Drafting Rules](#) (Verbal forms for the expression of provisions).

"**must**" and "**must not**" are **NOT** allowed in ETSI deliverables except when used in direct citation.

Introduction

Space-ground cooperative network includes the mobile communication network on the ground and the satellite network in the space, and the slicing configuration rules of the two networks are different. A slicing adaptation technology connecting mobile communication network and satellite network can effectively support the requirement of the end-to-end slicing service guarantee for space-ground cooperative network. Through the adaptation mapping of data plane and the collaborative management of control plane for Network Slicing, it can improve the customized service capability of space-ground cooperative network for differentiated services.

1 Scope

The present document intends to describe a method of network architecture and slicing mapping for the interconnection between the mobile communication network slicing and satellite network slicing. The detailed plan includes:

- Support identity resolution such as VLAN and IP address on the data plane, support precise identification and control for user services, and realize the slicing adaptation between mobile communication network on the ground and satellite network.
- Exchange the slicing control information, using the control plane of ground mobile communication network and satellite network (5GC and Satellite Network Operation Control Center (SNOCC)), optimize the global service quality of service for the network slicing, and ensure the consistency and continuity of slicing service in space-ground cooperative network environment.
- Leverage Graph Convolutional Networks (GCN) and Gated Recurrent Unit (GRU) to predict traffic patterns and optimize slice-resource mapping in real time.

The present document will deliver research and investigation activities and insights that will further explore the related techniques that can be used to employ connection improvement for space-ground network slicing.

2 References

2.1 Normative references

Normative references are not applicable in the present document.

2.2 Informative references

References are either specific (identified by date of publication and/or edition number or version number) or non-specific. For specific references, only the cited version applies. For non-specific references, the latest version of the referenced document (including any amendments) applies.

NOTE: While any hyperlinks included in this clause were valid at the time of publication, ETSI cannot guarantee their long-term validity.

The following referenced documents may be useful in implementing an ETSI deliverable or add to the reader's understanding, but are not required for conformance to the present document.

- [i.1] ETSI GR ENI 004: "Experiential Networked Intelligence (ENI); Terminology".
- [i.2] ETSI GS ENI 005: "Experiential Networked Intelligence (ENI); System Architecture".
- [i.3] ETSI GR ENI 008: "Experiential Networked Intelligence (ENI); InTent Aware Network Autonomicity (ITANA)".
- [i.4] [NIST Special Publication 800-207](#): "Zero Trust Architecture".
- [i.5] ETSI TS 138 413: "5G; NG-RAN; NG Application Protocol (NGAP) (3GPP TS 38.413)".
- [i.6] ETSI TS 123 008: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; Organization of subscriber data (3GPP TS 23.008)".
- [i.7] ETSI TS 123 501 (V17.8.0): "5G; System architecture for the 5G System (5GS) (3GPP TS 23.501 version 17.8.0 Release 17)".
- [i.8] ETSI TS 123 273: "5G; 5G System (5GS) Location Services (LCS); Stage 2 (3GPP TS 23.273)".

- [i.9] ETSI TS 123 122: "Digital cellular telecommunications system (Phase 2+) (GSM); Universal Mobile Telecommunications System (UMTS); LTE; 5G; Non-Access-Stratum (NAS) functions related to Mobile Station (MS) in idle mode (3GPP TS 23.122)".
- [i.10] ETSI TS 124 501: "5G; Non-Access-Stratum (NAS) protocol for 5G System (5GS); Stage 3 (3GPP TS 24.501)".

3 Definition of terms, symbols and abbreviations

3.1 Terms

For the purposes of the present document, the terms given in ETSI GR ENI 004 [i.1] and ETSI GS ENI 005 [i.2] apply.

3.2 Symbols

Void.

3.3 Abbreviations

For the purposes of the present document, the abbreviations given in ETSI GR ENI 004 [i.1], ETSI GS ENI 005 [i.2] and ETSI GR ENI 008 [i.3] apply.

4 Overview

4.1 Introduction

With the evolution of network services, it has become increasingly challenging to dynamically match network resources to the diverse and simultaneous demands of network capacity. Multiple Network Slicing (NS) technology brings an excellent solution for the mismatch between supply and demand of network capacity. NS virtualizes multiple network slices within a network to provide customized services tailored to diverse performance requirements. It can not only meet the performance requirements of different services, but also maximize the network resource utilization, save the cost of network construction, and improve the profitability of operators. Finally, it achieves the network service and cost-benefit balance.

To meet diverse user requirements, NS integrates various network elements on a shared physical platform to create independent, end-to-end service subnets. While the underlying architecture is modular and flexible, NS enables tightly integrated service delivery for a wide range of business demands. In recent years, NS technology has advanced rapidly across wireless access, mobile core, IP bearer, satellite and other network environments. However, the interconnection of network slices across heterogeneous domains, such as terrestrial and satellite networks, remains a significant challenge. The configuration and management rules for slicing differ between these environments. Therefore, adaptive slicing technologies that bridge mobile and satellite networks are essential for delivering consistent, end-to-end service guarantees in space-ground cooperative networks. By enabling adaptive mapping at the data plane and the collaborative management at the control plane, these technologies enhance the ability of space-ground networks to deliver tailored, high-quality services.

4.2 Architecture

4.2.1 Space-Ground Cooperative Network Slicing Architecture

- The space-ground cooperative network slicing architecture is shown in Figure 1. The architecture includes the deployment of a programmable slicing gateway and a space-ground cooperative slicing control system, positioned between the terrestrial mobile communication network and the satellite network. The programmable slicing gateway is the transit channel for the slicing service data flows. With definable message parsing (the capability of a gateway to be programmed to identify, interpret, and extract specific information from data packets based on a given configuration policy, enabling precise control and forwarding of different slicing services), processing and forwarding capabilities, the gateway accurately identifies and controls slicing services, and achieves the data mapping between slices according to the configuration policy provided by the control system. It can ensure the service consistency and continuity of service data in space-ground cooperative network slicing and realize the adaptation of heterogeneous network slices. Example of heterogeneous network slices:
- QoS Parameters:** A slice for Ultra-Reliable Low-Latency Communication (URLLC) like autonomous vehicle control would prioritize stringent latency and reliability guarantees, while a slice for massive IoT would prioritize connection density and energy efficiency over low latency.
- Session Management:** A terrestrial mobile network slice might establish a session with a simple handover between cell towers, while a space-ground slice requires complex, predictive session management to handle handovers between satellites moving at high orbital speeds and ground stations.

The space-ground cooperative slicing control system interacts with the space-ground network slicing control planes to open up the slicing session channel between the space and ground network. Taking into account the differences between mobile communication network and satellite network in slicing service classification, slicing quantity and slicing construction form, the control system can optimize the matching mode of service traffic and network resources, and intelligently generate the configuration policy of the programmable slicing gateway, thus improving the end-to-end quality of slicing service in space-ground cooperative network. Optimize the matching mode of service traffic and network resources can be realised by the following methods:

- Dynamic Resource Allocation:** The control system could direct high-bandwidth, non-delay-sensitive traffic (e.g. software updates for ships) to a satellite slice with abundant bandwidth but higher latency, while reserving scarce, low-latency terrestrial resources for real-time video calls.
- Predictive Traffic Steering:** For a user on a high-speed train, the system could predict the route and pre-emptively allocate resources and establish sessions on satellite network slices to maintain service continuity before the terrestrial network connection is lost.

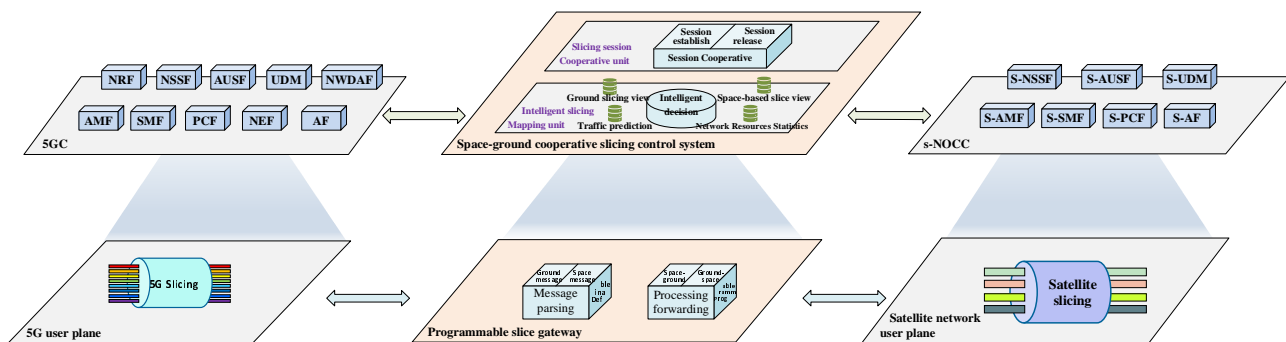


Figure 1: Space-ground cooperative network slicing architecture

4.2.2 Space-ground Slicing Session Collaboration

The main function of slice-session collaboration is to coordinate the management of PDU sessions in mobile communication network and satellite network, and establish PDU session channels from UE to ground-based 5G mobile communication network, space-based satellite network and up to Data Network.

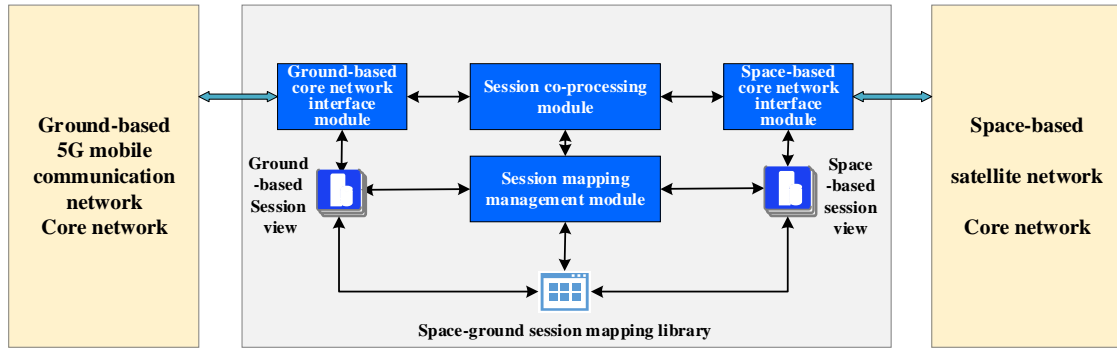


Figure 2: Slicing session collaboration architecture

As shown in Figure 2, the functional modules of slicing session collaboration unit include slicing mapping management module, session collaborative processing module, ground-based core network interface module, and space-based core network interface module. The slice mapping management module is mainly responsible for maintaining the mapping relationship between ground-based PDU sessions and space-based PDU sessions. The session cooperative processing module can cooperate with the process of establishing, modifying and releasing sessions of ground-based and space-based networks, according to the mapping relationship maintained by the slice mapping management module. The interface module of ground-based core network is responsible for the interface with the core network of ground-based 5G mobile communication network. The space-based core network interface module is responsible for the interface with the space-based satellite network core network.

The establishment process of UE-initiated PDU sessions is used as an example to illustrate the slicing session collaboration process. In the following example, assuming that the mapping rule is based on service type, UE1 and UE2 initiate PDU sessions of the same type to access Data Network.

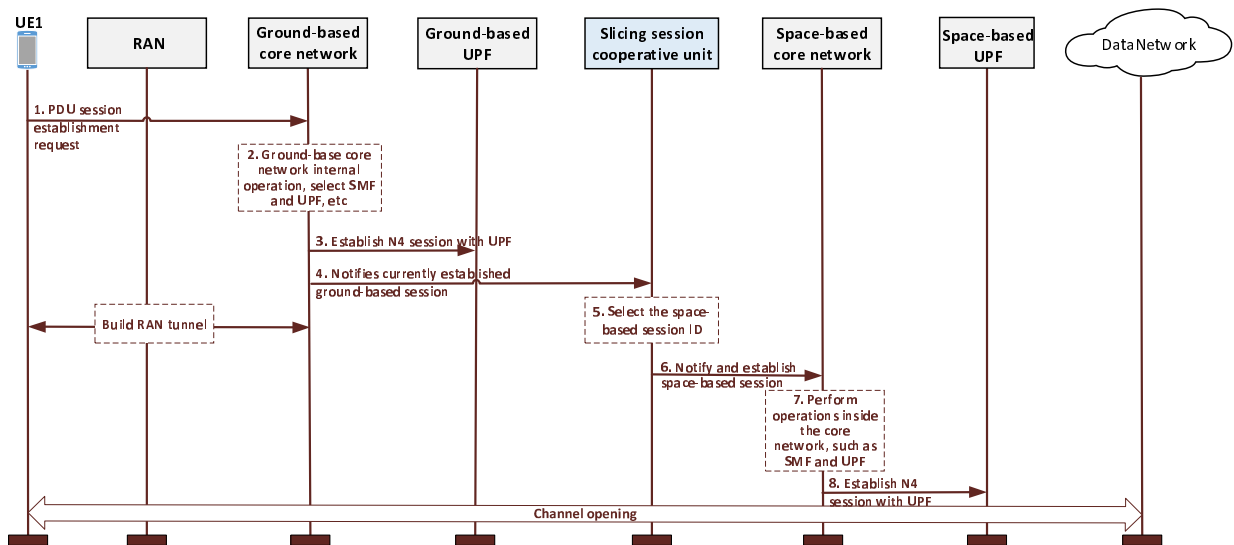


Figure 3: Data channel establishment process of UE1

For the PDU session initiated by UE1, the data path establishment includes three stages, as shown in Figure 3.

The first stage is PDU session establishment process from UE to ground mobile communication network:

- Step1: UE1 initiates a PDU session establishment request.
- Step2: The request is processed by the ground-based core network, and is used to select the ground-based SMF and UPF for the session.
- Step3: The ground-based core network establishes an N4 session with the selected ground-based UPF.
- Step 4: The ground-based core network notifies the space-ground cooperative session management unit of the currently established ground-based session information.

The ground-based core network notifies the information about the current ground-based session to space-ground session management unit. At the same time, the ground core network notifies RAN and users to build RAN tunnels.

The second stage is the slicing session collaborative unit for slicing mapping:

- Step 5: After receiving the notification from the ground-based core network, the slicing session collaborative unit carries out the space-based session mapping.

Since the session of UE1 is a new service type, a new space-based session ID needs to be assigned to the session of UE1.

The third stage is the PDU session establishment process of the satellite network:

- Step 6: The slicing session collaboration unit notifies the establishment of a new space-based session to the space-based core network.
- Step 7: The space-based core network selects the space-based SMF and UPF for the session after receiving a session establishment notification.
- Step 8: The space-based core network establishes the N4 session with the selected space-based UPF.

At this point, for the PDU sessions of UE1, the channel from UE1 to ground-based 5G mobile communication network, space-based satellite network and up to Data Network has been established and opened.

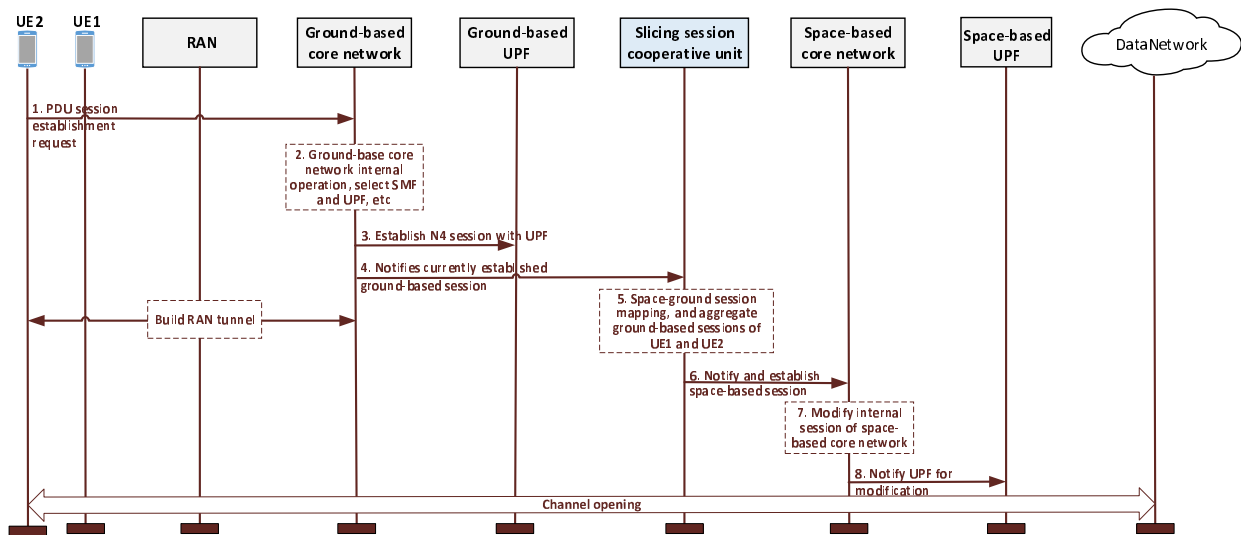


Figure 4: Data channel establishment process of UE2

As shown in Figure 4, after UE1 has established the channel to the Data Network, and when UE2 intends to access the Data Network, the establishment of the data channel also includes three stages, as shown in Figure 4.

The first stage is the process of PDU session establishment from UE to ground-based mobile communication network.

- Step1~Step4: The process of ground-based network for the session establishment request of UE2 is the same as that of UE1.

The second stage is the slicing session collaborative unit for slicing mapping:

- Step 5: The slicing session collaborative unit carries out the space-ground session mapping after receiving the notification of the ground-based core network. Based on resource allocation, the control unit determines that UE2 and UE1 sessions can be aggregated into the same space-based session (e.g. if they share the same service type and QoS requirements).

The third stage is the PDU session establishment process of the satellite network:

- Step 6: The slicing session collaborative unit notifies the space-based core network to modify the space-based session, and the modification can be for QoS parameters.

- Step 7: The space-based core network performs a modification operation for the session after receiving the session modification notification.
- Step 8: The space-based core network notifies the corresponding space-based UPF to perform session modifications.

At this point, for the PDU sessions of UE2, the channel from UE2 to ground-based 5G mobile communication network, space-based satellite network and up to Data Network has been established and opened. PDU sessions of the same service type in UE1 and UE2 are allocated to the same slice, and the slicing sessions terminates (this refer to the end of the process).

4.2.3 Intelligent Slice Mapping

In the space-ground cooperative network, there are many types of service requirements. The performance requirements of services such as real-time voice, data transmission, control signalling, and short message have different performance requirements, and the service delay, bandwidth, and security requirements all change in real time. To meet the differentiated application requirements of wide-area information networks, the space-ground cooperative network needs to dynamically construct differentiated network slices involving different service characteristics, accurately match the resource requirements of different service data, and realize multi-service converged application.

The mechanism achieves this through the intelligent, real-time decision-making capabilities of the **Space-Ground Cooperative Slicing Control System** introduced in clause 4.2.1. This system acts as the central orchestrator:

- **Reconciling Conflicting QoS:** The control system can prioritize and reconcile conflicting requirements by leveraging the heterogeneous resources of both terrestrial and satellite networks, as described in the "Optimize the matching mode" examples (clause 4.2.1). For instance, a service with strict low-latency requirements (e.g. real-time voice) would be matched to a terrestrial network slice or a Low-Earth Orbit (LEO) satellite link with minimal delay. A service demanding high bandwidth but tolerating higher latency (e.g. data transmission for software updates) would be directed to a geostationary (GEO) satellite slice with abundant bandwidth. The control system intelligently makes this choice based on its global view of all network resources and service requirements.
- **Orchestration:** The orchestration is performed by this control system. It is responsible for the end-to-end lifecycle management of slices across both domains. It "intelligently generates the configuration policy of the programmable slicing gateway" and, by interacting with both ground and space core network control planes, it orchestrates the overall resource matching and slice mapping process.

The space-ground cooperative network proposes an intelligent slice mapping mechanism based on spatial-temporal correlation.

The intelligent slice mapping mechanism is the **brain**, while the programmable gateway and session collaboration unit are the **executing limbs**:

- **Connection to the Programmable Slicing Gateway (clause 4.2.1):** The control system (which houses the intelligent mapping logic) generates the **configuration policy** for the gateway. Based on the mapping decisions, it commands the gateway on how to perform "data mapping between slices" - for example, which specific service flows to steer onto which network paths (terrestrial or satellite) to meet their QoS requirements. The gateway executes these policies using its "definable message parsing" capability.
- **Connection to the Session Collaboration Unit (clause 4.2.2):** The session collaboration unit is a key functional component that **implements** the mapping decisions for session management. The intelligent mapping mechanism likely provides the **mapping rules and logic** (e.g. "based on service type" or "based on resource allocation") that the slice mapping management module within the session collaboration unit follows. This is shown in the UE1/UE2 example where the session unit decides to create a new space-based session for a new service type (UE1) or aggregate sessions into an existing one (UE2) based on these rules.

Traffic prediction is used to establish the prediction model of resource demand of network services:

- **Model Type:** The spatial-temporal correlation depend on the use of advanced ML models capable of analyzing patterns across both **time** (e.g. time of day, network usage cycles) and **space** (e.g. user location, satellite coverage, ground network congestion). Models used including the following:
 - **Time Series Forecasting models** (e.g. LSTMs - Long Short-Term Memory networks) were used to predict traffic load fluctuations.
 - **Reinforcement Learning** models, where the control system learns optimal mapping and resource allocation policies through continuous interaction with the network environment.
 - **Graph Neural Networks** were used to model the complex topology of the space-ground network and optimize resource paths.

This enables the space-ground cooperative network to respond to the service characteristics and the transformation of access node in real time. Thus, the slices of network resources can be matched as needed with the fluctuating traffic in the space-ground cooperative network.

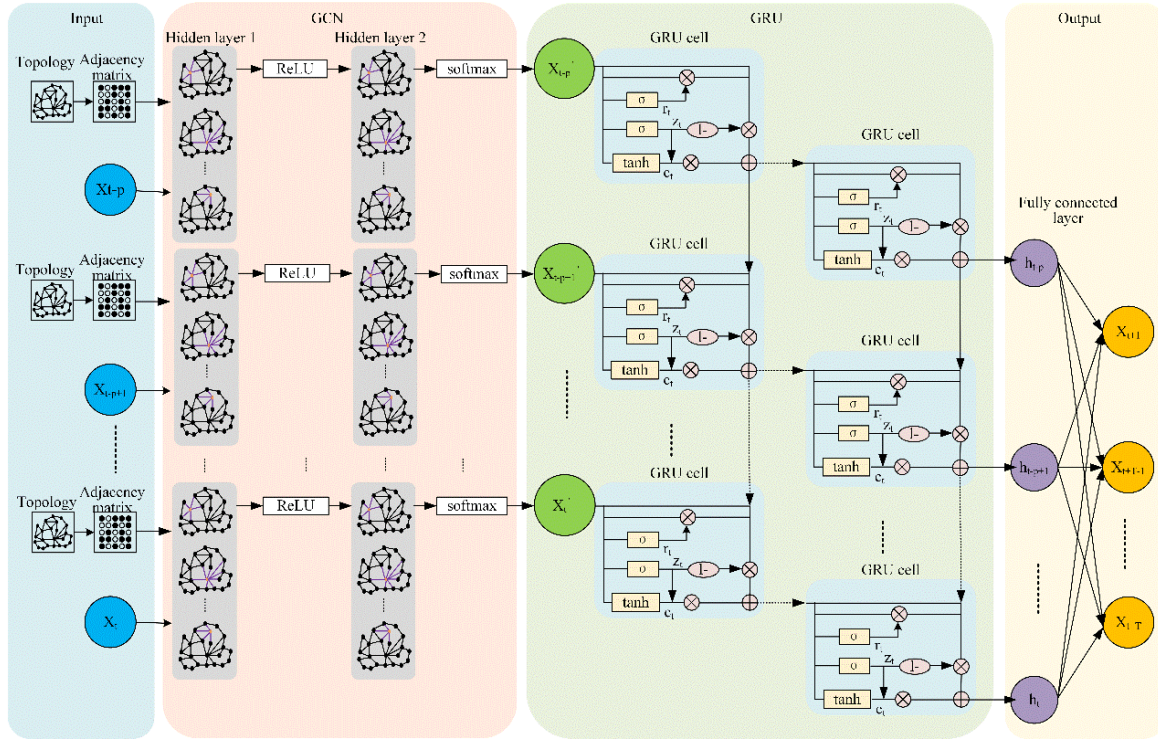


Figure 5: Intelligent slice mapping based on spatial-temporal correlation

Figure 5 shows the smart slice mapping diagram based on spatial-temporal correlation. Graph Convolutional Network (GCN) and Gated Recurrent Unit (GRU) are used to extract the temporal-spatial characteristics of the historical traffic load of each node in the space-ground cooperative network slicing, which is to provide a decision basis for slice mapping. Firstly, the network topological features are captured by GCN to obtain the spatial dependence. Secondly, the dynamic changes of node attributes are captured by GRU to obtain the local time trend of traffic load. Finally, the multi-output fully connected layer of artificial neural network is used to realize the transformation from traffic load to resource demand, and output the predicted result. The system monitors the network resource status in real time, slices are allocated network resources based on the predicted results of slicing service requirements to complete slicing adaptation decisions.

The real-time monitoring of network resource status is achieved through a comprehensive telemetry data collection framework. The Space-Ground Cooperative Slicing Control System acts as a central brain, continuously gathering and analyzing real-time performance and status metrics from all key network elements across both terrestrial and satellite domains. This is done through integrated agents and standardized interfaces:

- **Terrestrial Network (Ground-based UPF & RAN):** The system collects real-time telemetry on Key Performance Indicators (KPIs) such as **bandwidth utilization, port statistics, packet loss, and session load** from User Plane Functions (UPFs), and **radio resource block usage, connected User Equipment (UE) count, and handover events** from the Radio Access Network (RAN).
- **Satellite Network:** The system interfaces with the satellite network's core and management systems to monitor **per-beam capacity utilization, Signal-to-Noise Ratio (SNR) metrics for different geographic areas, satellite transponder load, and propagation delay** characteristics.
- **Unified View:** This constant stream of spatial-temporal data provides the control system with a unified, real-time view of the entire integrated network's health, capacity, and congestion points, forming the ground truth against which AI predictions are validated and resource allocation decisions are executed.

Figure 5 shows the smart slice mapping diagram based on spatial-temporal correlation. Graph Convolutional Network (GCN) and Gated Recurrent Unit (GRU) form the core of a predictive AI model that extracts the temporal-spatial characteristics from the historical traffic load of each node (e.g. base stations, satellite gateways, UPFs) in the space-ground cooperative network. This model provides the intelligent decision basis for proactive slice mapping and resource allocation. The process, also illustrated in Figure 5, works as follows:

- 1) **Spatial Feature Extraction via GCN:** The **GCN layers** process the physical and logical **network topology**. Nodes represent network elements (e.g. gNBs, Satellites, UPFs), and edges represent the links between them. The GCN effectively captures the **spatial dependence** and correlations between neighboring nodes, understanding how traffic congestion in one cell might impact adjacent cells or a satellite beam.
- 2) **Temporal Feature Extraction via GRU:** The time-series data of historical traffic load for each node is fed into the **GRU layers**. The GRU is adept at capturing **dynamic changes and local time trends**, learning patterns such as daily usage cycles, periodic bursts of traffic, and long-term growth trends.
- 3) **Prediction and Decision Output:** The combined spatio-temporal features from the GCN and GRU are then fed into a **multi-output fully connected layer** of an artificial neural network. This layer performs the non-linear transformation from abstract features into a concrete **predicted resource demand** (e.g. required Gbps of bandwidth, number of required network slices) for each segment of the network, outputting the final forecast result.
- 4) **Closed-Loop Resource Allocation:** These predicted outputs are fed directly into the resource allocation decisions of the control system. The system proactively instructs the programmable slicing gateway and session collaboration unit to pre-emptively configure slices and allocate resources before the predicted demand arrives.
- 5) **The Feedback Loop for Real-Time Monitoring:** Crucially, this is a closed-loop system. The system **monitors the network resource status in real-time** via the telemetry framework described above. This real-world data on **actual resource utilization, Bit Error Rate (BER), and achieved data rates** is continuously fed back as new input into the AI model. This allows the model to compare its predictions with reality, automatically learn from any discrepancies, and refine its future predictions, creating a self-optimizing loop. Slices are thus dynamically adapted and network resources are allocated based on this continuous cycle of prediction and observation, ensuring efficient resource utilization and meeting stringent end-to-end Quality of Service (QoS) requirements.

In figure 5, the desired system performance can benefit from an adaptive mechanism of slicing. A well-designed slicing algorithm takes into account the (unavoidable) tradeoffs between bandwidth and power efficiency). Haiyuan Li et al. (arXiv:2310.17523) proposed a Multi-Agent Deep Reinforcement Learning (MADRL) Approach, highlighted a state-of-the-art method that addresses the trade-offs between bandwidth and power efficiency through intelligent, adaptive slicing algorithms that can be applied in the present document.

The slicing scheme here was initially intended to mitigate the data rate drop in a generic link, where interference is present. The candidate switching schemes have been chosen based on the merit of combined power and bandwidth efficiency. More realistic models will also be addressed in this. For model simplicity and discussion continuity, the following assumptions are made:

- a) The channel is a direct link channel.
- b) Synchronization is maintained.
- c) The system operates at 2 possible data rates (moderate and high).

The direct-link assumption simplifies initial validation; future iterations will incorporate 3D channel models to account for Doppler, polarization and atmospheric effects.

In the test model two sets of BER test data will be used. The first set will range from 10^{-9} to 10^{-7} , to represent moderate degradation and second set will range from 10^{-8} to 10^{-6} , to represent severe degradation. The data rate will be chosen as a realistic 4 Mbps and a higher rate of 250 Mbps.

Table 1: Data rate degradation in Space Ground network

Moderate Degradation (BER) $10^{-9} \rightarrow 10^{-7}$		Severe Degradation (BER) $10^{-8} \rightarrow 10^{-6}$	
Initial Rb (Mbps)	Final Rb (Mbps)	Initial Rb (Mbps)	Final Rb (Mbps)
4	3	4	2,9
250	187,9	250	179,4

Assumptions (a-c) represent a controlled test case; future work will address multipath and mobility, 4 Mbps (narrowband IoT) and 250 Mbps (broadband video) reflect 3GPP NTN use cases.

In Table 1, while in the moderate degradation, the data rates fell for 4 Mbps and 250 Mbps respectively to 3 Mbps and 187,9 Mbps, which is about 25 % of throughput loss; for severe degradation, the data rate fell for 4 Mbps and 250 Mbps respectively to 2,9 Mbps and 179,4 Mbps, which is about 28 % of throughput loss.

Comparing data rate loss for systems operating at 4 Mbps, the difference between the moderate and severe model is not significant. In contrast, for a system operating at 250 Mbps, the data rate loss between the moderate and severe model indicates a large departure from the operating speed. This suggests that systems operating at a higher speed are more susceptible to environmental change than their lower speed counterparts. The comparisons between these two data rate bands can be further demonstrated.

4.2.4 Intelligent slicing technology for space-ground collaborative network resources on demand

This clause explores intelligent slicing technology for on-demand space-ground collaborative network resource allocation, focusing on efficient slice generation to meet diverse application requirements and a mathematical model using optimal weighted graph matching to optimize the mapping process. The need to dynamically construct differentiated network slices tailored to specific business characteristics, ensuring accurate alignment between resource capabilities and service demands were also addressed. The importance of timeliness in slice mapping algorithms, particularly when handling large volumes of real-time service requests, while also balancing node load and link bandwidth to prevent network overload were emphasised. Additionally, a mathematical approach grounded in adjacency matrix feature vector decomposition to formulate slice mapping as an optimal weighted graph matching problem was introduced. This model translates resource slicing requirements into a virtual topology representation, enabling efficient, simultaneous processing of multiple slice requests and enhancing the scalability and responsiveness of space-ground collaborative networks:

1) Efficient generation technology of network resource slicing:

In order to meet the requirements of differentiated wide-area information network applications, the generation and scheduling of intelligent slices of space-ground collaborative network. This generation needs to dynamically construct differentiated network slices involving different business characteristics. It also, accurately matches the resource requirements of different business data requirements, and realizes multi-service integration applications. In the process of providing a wide variety of network services, the space-ground synergy network needs to deploy a large number of network resource slices with different functional sequences to meet the real-time needs of users, and the timeliness of the slice mapping algorithm is very important in the face of a large number of real-time service requests. In addition, due to the limited processing power and link bandwidth of nodes in the network, once some nodes and links are overloaded, the carrying capacity of the network will be reduced. Therefore, the efficient generation technology of network resource slicing needs to complete the rapid mapping of a large number of service requests while balancing the node load and link bandwidth.

The "differentiated wide-area information network applications" refers to distinct types of data services that span large geographical areas (wide-area) and have vastly different performance requirements (differentiated). These applications cannot be served by a one-size-fits-all network and instead require dedicated, customized logical networks (slices).

Examples of Differentiated Wide-Area Information Network Applications:

1) Ultra-Reliable Low-Latency Communication (URLLC) for Critical Control Signaling:

- Application: Remote control of autonomous cargo ships or Unmanned Aerial Vehicles (UAVs) operating in remote oceans or airspace. This requires continuous, real-time transmission of control commands and telemetry data.

- Differentiated Requirements: This application is not about moving large amounts of data but about extremely high reliability (e.g. 99,999 %) and very low latency (e.g. < 10 ms). A delay or a lost packet could lead to a catastrophic failure. The network slice for this application would be engineered to prioritize these metrics above all else, potentially using dedicated satellite beams with robust error correction and pre-empting other traffic.
- 2) Enhanced Mobile Broadband (eMBB) for High-Throughput Video Transmission:
- Application: Live video broadcasting from a major international sporting event (e.g. the Olympics, World Cup) or providing high-speed internet connectivity to passengers on commercial flights or cruise ships.
 - Differentiated Requirements: This application is defined by its need for very high bandwidth and data rates to support HD/4K video streams for thousands of users simultaneously. Latency can be tolerated to a much greater degree than in URLLC. The network slice for eMBB would be allocated a large portion of the available spectrum and channel bandwidth on both terrestrial and satellite links to maximize throughput.
- 3) Massive Machine-Type Communications (mMTC) for Large-Scale Sensor Networks:
- Application:
 - Environmental Monitoring: A network of sensors deployed across a continent to monitor soil moisture, air quality, or seismic activity.
 - Global Logistics: Tracking the location and status (e.g. temperature, humidity) of millions of shipping containers worldwide.
 - Differentiated Requirements: The key requirement is connection density - the ability to support a massive number of devices (e.g. millions per square kilometer) transmitting very small amounts of data infrequently. These devices are often battery-powered, so energy efficiency is paramount. Latency and bandwidth are minimal concerns. A slice for this would be optimized to handle many small, sporadic connections with minimal power consumption.
- 4) Generate a mathematical model that uses the optimal weighted graph matching the slices:

Based on the principle of adjacency matrix feature vector decomposition, the present document uses the optimal weighted graph matching algorithm to calculate the function and the mapping relationship of links, and describes the resource slicing requirements of space-ground collaborative network as a virtual topology map. The map then supports the simultaneous processing of multiple slice requests.

The algorithms designed to find a matching between two graphs that minimizes the total cost or maximizes the total weight include:

- The Hungarian Algorithm (Kuhn-Munkres Algorithm): A classic combinatorial optimization algorithm that solves the assignment problem in polynomial time. It is highly suitable for finding the one-to-one optimal mapping between slice requests (virtual network functions, VNFs) and physical network resources (nodes, links) based on a cost matrix (e.g. latency, available CPU, bandwidth).
- The Auction Algorithm: Known for its parallelizability and speed in certain scenarios, which will be advantageous for real-time slicing requests.

Balancing Efficiency and Load: The algorithm itself would not inherently balance load; this is achieved through the design of the cost function. The "weighted" aspect is key. The cost matrix used for matching incorporate multiple weighted factors:

- Efficiency Metrics: Latency, bandwidth, proximity to user.
- Load Metrics: Current CPU utilization of a node, current bandwidth consumption of a link. By assigning a high cost to overloaded nodes and links, the algorithm would naturally avoid them, seeking a mapping that fulfills the service requirement while distributing load across the network. This transforms the load-balancing requirement into an optimization constraint within the graph-matching problem.

The mathematical model's cost function shall include decomposed SLA parameters as constraints. For instance:

- The SLA's latency budget would be split into a terrestrial latency budget and a satellite latency budget.
- The graph-matching algorithm would then only consider node and link mappings that can meet their portion of the decomposed SLA. A satellite beam with high propagation delay might be rejected for a URLLC slice even if it has available bandwidth.

How the architecture handles session failures, handover interruptions, or resource contention:

- **Session Failure & Handover Interruptions:** The system managed by the Space-Ground Cooperative Slicing Control System, upon detecting a satellite beam handover or a node failure (e.g. via loss of heartbeat), it would:
 - 1) Trigger a rapid re-calculation of the optimal mapping using the graph-matching model, excluding the failed resource.
 - 2) Instruct the Programmable Slicing Gateway to re-route existing sessions according to the new mapping.
 - 3) Utilize the Session Collaboration Unit (from clause 4.2.2) to re-establish or modify PDU sessions on the new path with minimal disruption.
- **Resource Contention:** In scenarios where resources are insufficient to meet all requests, a prioritization policy can be realized by the following:
 - **Slice Priority:** Pre-defined slice priorities (e.g. URLLC for emergency control > eMBB for video streaming).
 - **Admission Control:** The control system would reject new, lower-priority slice requests to protect the performance of existing, high-priority slices.

Security measures and compliance requirements:

- **Authentication & Authorization:** Every network element (gateway, control system, satellite modem) shall mutually authenticate before exchanging data or control commands, using robust protocols like TLS 1.3 or protocol-specific secure variants (e.g. HTTPS, NETCONF over SSH).
- **Encryption:** All user plane traffic and control plane signaling need to be encrypted End-to-End (E2E) or at least hop-by-hop. This is non-negotiable for protecting sensitive data like control signals for UAVs or industrial sensor data.
- **Slice Isolation:** This is a core security requirement. Slices shall be strictly isolated from one another to prevent a breach or traffic explosion in one slice (e.g. mMTC) from impacting another (e.g. URLLC). This aligns with standards like 3 ETSI TS 123 501 [i.7], which mandates security isolation between network slices as a fundamental principle.
- **Compliance:** The architecture should be designed to comply with broader zero-trust frameworks like NIST SP 800-207 [i.4], which emphasizes "never trust, always verify," moving beyond traditional perimeter-based security - a crucial concept for open, wide-area networks.

5 Network Structure

5.1 Forward and Backhaul Link

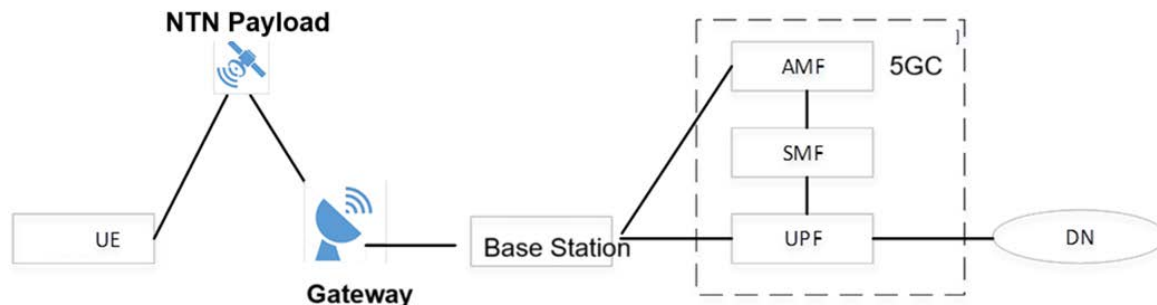


Figure 6: Architecture diagram of satellite access

Figure 6 shows that satellite access, from the perspective of network elements, NTN payloads, satellite gateway information customs stations, 5G core networks, etc., and from the link point of view, it includes the service link between the terminal and the NTN payload, and the feed link between the NTN payload and the information customs station system architecture that supports satellite backhaul.

Figure 6 illustrates the architecture for satellite access from two perspectives:

- 1) **Network Elements:** This includes NTN payloads, satellite gateways, customs information stations, and the 5G core network.
- 2) **Communication Links:** This encompasses the service link between the terminal and the NTN payload, and the feeder link between the NTN payload and the gateway. This architecture supports satellite backhaul.

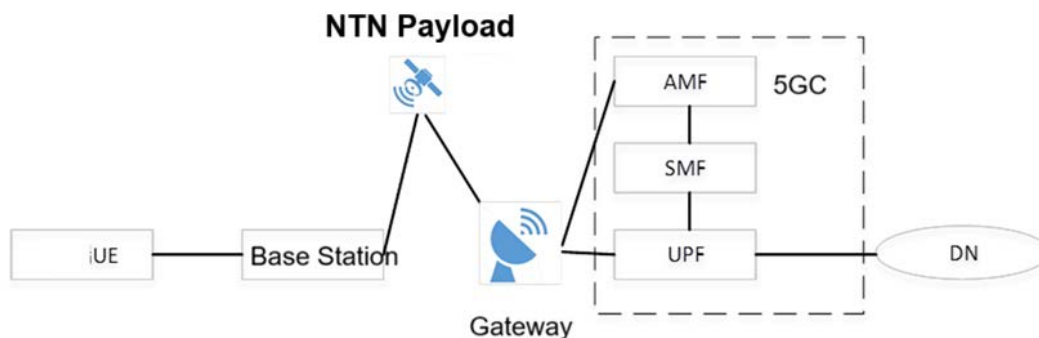


Figure 7: Satellite link as backhaul

Satellite backhaul is used between the core and terrestrial access networks to provide transmission to N1/N2/N3 reference points. The satellite system transparently carries the communication payload of the 3GPP reference point.

5.2 Identification and restriction of satellite access types

In the case of New Radio (NR) satellite access, the Radio Access Technique (RAT) type values "NR(LEO)", "NR(MEO)", "NR(GEO)" and "NR(othersat)" are used in 5GC to distinguish different NR satellite access types.

When a UE connects to the network via satellite, the AMF determines the NR satellite access type (e.g. NR (LEO), NR (MEO), NR (GEO) and NR (othersat)) based on information provided by the radio access network over the N2 interface, as specified in ETSI TS 138 413 [i.5].

Serving PLMN may provide the AMF with information about the user's NR access restrictions as defined in ETSI TS 123 008 [i.6] in accordance with ETSI TS 123 501 [i.7], clause 5.3.4.1. For example, if a PLMN's license does not permit NR(LEO) as primary access in a given region, the AMF will reject registration attempts for that RAT type in that area.

5.3 UE location Identification

The network should ensure that UEs only receive service in areas permitted by the PLMN's license, and should promptly de-register or deny service when this cannot be assured.

To meet regulatory requirements, during the mobility management and session management processes, the network can be configured to verify the UE location to determine if the selected PLMN is allowed to provide services in the current UE location. In this case, when the AMF receives an NGAP message containing the user's location information, this is used to determine the use of NR satellites to access the UE, the AMF can decide to verify the UE's location. If the AMF determines that services are not allowed at the current UE location based on the PLMN ID selected and the ULI (including the Cell ID) received from the gNB, the AMF should reject any NAS request with the appropriate reason value. If the UE is already registered with the network, when AMF determines that it is not allowed to run in the current UE location, AMF may initiate the UE to de-register. Unless the AMF has sufficiently accurate UE location information to determine that the UE is located in a geographic area where PLMN does not allow operation, the AMF should not reject the request or register the UE.

NOTE: The regions in which UE is allowed to operate can be determined based on the regulatory areas that PLMN's licensing conditions allow it to operate.

If, based on the ULI, the AMF is not able to determine the location of the UE with sufficient accuracy, the AMF should continue with the Mobility Management or Session Management process and may initiate the UE location process after the completion of the Mobility Management or Session Management process, as specified in ETSI TS 123 273 [i.8], clause 6.10.1, to determine the location of the UE. If the message received from the LMF indicates that the UE is registered with a PLMN that is not allowed to run in the UE location, the AMF should be prepared to register the UE. In the case of a NAS flow, the AMF should reject any NAS request with the destination of a PLMN that is not allowed to run in a known UE location and indicate an appropriate reason value, or accept the NAS request and initiate a de-registration process once the UE location is known. In the de-registration message sent to the UE, the AMF should contain an appropriate reason value. For UE processing to indicate that PLMN is not allowed to run at the current UE location, see ETSI TS 123 122 [i.9] and ETSI TS 124 501 [i.10].

During the handover process, if the (target) AMF determines that it is not allowed to provide services to the UE at the current location, the AMF can reject the switchover, or accept the switchover and initiate a deregistration process for the UE.

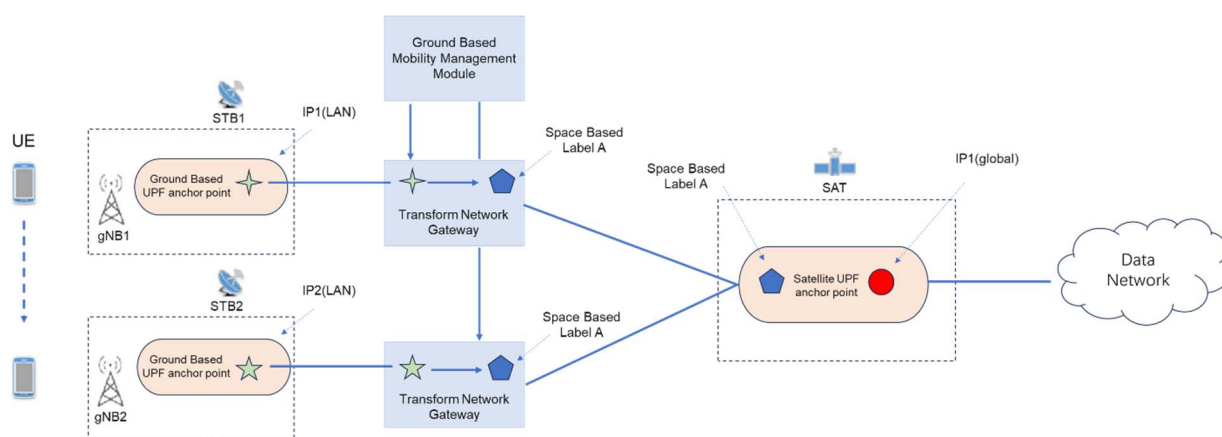


Figure 8: The process flow of the ground-based mobility management module

Figure 8 shows the UE Mobile Event Management process. The ground-based mobility management module carries out unified management of the network identification used by the ground-based and space-based network. The module realizes the collaborative distribution and mapping of the network identity. These network identifiers include the label of the space-based network and the global IP address of the access data network. The transform gateway that is deployed after the ground-based UPF anchor of the satellite communication base station is managed by the ground-based mobility management module. The identity conversion operation is performed, and the data is processed by the transform gateway, then arrives at the satellite SAT through the satellite-ground link, reaches the data network via a satellite UPF anchor.

The identity mapping is performed by the Transform Gateway acting as a sophisticated Network Address and Port Translation (NAPT) device or a Mapping Function between two addressing domains.

Here is a step-by-step breakdown of the process:

1) Registration & Mapping Establishment:

- When a User Equipment (UE) registers with the network, the Ground-based Mobility Management Module (e.g. a unified AMF/SMF) assigns it a globally routable IP address for the Data Network.
- Simultaneously, the module coordinates with the space-based network to assign a unique, internal label for the space-based network (e.g. a satellite-specific temporary identifier or a non-routable IP address).
- The Mobility Management Module creates a binding record that maps the UE's global IP address to its space-based network label. This binding record is then provided to the Transform Gateway.

2) Uplink Data Flow (UE → Data Network):

- A data packet originates from the UE, destined for the internet. This packet has a source IP address set to the UE's global IP address.
- The packet travels through the satellite (SAT) and arrives at the Transform Gateway.
- The Transform Gateway consults its mapping table. It performs a reverse mapping: it identifies the global IP address and rewrites the packet's source address to the corresponding space-based network label. This step is crucial for the satellite network's internal routing to function correctly.
- The packet, now with the space-based label as its source, is routed through the satellite UPF anchor and onward to the Data Network.

3) Downlink Data Flow (Data Network → UE):

- A response packet arrives from the Data Network at the satellite UPF anchor, destined for the UE's global IP address.
- The packet is routed to the Transform Gateway.
- The Gateway performs a forward mapping: it uses the destination address (the global IP) to look up the corresponding space-based network label.
- It rewrites the packet's destination address to this space-based label.
- The packet is then sent back over the satellite-ground link to the satellite, which uses the internal label to correctly route the packet down to the specific UE.

In essence, the Transform Gateway maintains a dynamic lookup table that binds a publicly routable identity (Global IP) to a privately routable, network-specific identity (Space-based Label). It performs a bidirectional translation on all passing data packets, ensuring seamless end-to-end connectivity while hiding the internal addressing scheme of the satellite network from the core data network, and vice versa.

As shown in Figure 8, the UE accesses through gNB1 in Satellite Terminal Building 1 (STB1) as follows:

- 1) The UE obtains an IP address of type LAN, which is recorded as IP1. At the same time, the ground-based mobility management module assigns a space-based label A and an IP address of type global to the UE, which is denoted as IP(global).

- 2) When UE service data is exported from the UPF anchor of STB1, NAT is not performed, and the source address of the output packet is IP1.
- 3) The packet arrives at the translation gateway, maps the source address to space-based label A, and outputs the packet.
- 4) The packet arrives at the UPF anchor of the satellite network, which is initially space-based label A, maps the source address to IP (global), and outputs the IP packet to the data network.

The space-ground collaborative network mobility management model realizes the smooth switching of slice channels when users move between different STBs. Taking Figure 8 as an example, when the UE moves from STB1 to STB2, the access process through gNB2 in STB2 is as follows:

- 1) The UE obtains a new IP address of type LAN, which is recorded as IP2; At the same time, the mobility management module delivers the previously assigned space-based tag A to the corresponding transform gateway according to the UE access situation.
- 2) When UE service data is exported from the UPF anchor of STB2, NAT is not performed, and the source address of the output packet is IP2.
- 3) The packet arrives at the translation gateway, maps the source address to space-based label A, and outputs the packet.
- 4) The packet arrives at the UPF anchor of the satellite network, which is initially space-based label A, maps the source address to IP (global), and outputs the IP packet to the data network.

6 Mathematical model

6.1 Mathematical model introduction

Based on the principle of adjacency matrix feature vector decomposition, the optimal weighted graph matching algorithm is used to calculate the mapping relationship between the function and the link, and the resource slicing requirements of the space-ground collaborative network are described as virtual topology maps, which support the simultaneous processing of multiple slice generation requests. The optimal weighted graph matching problem is modelled as follows.

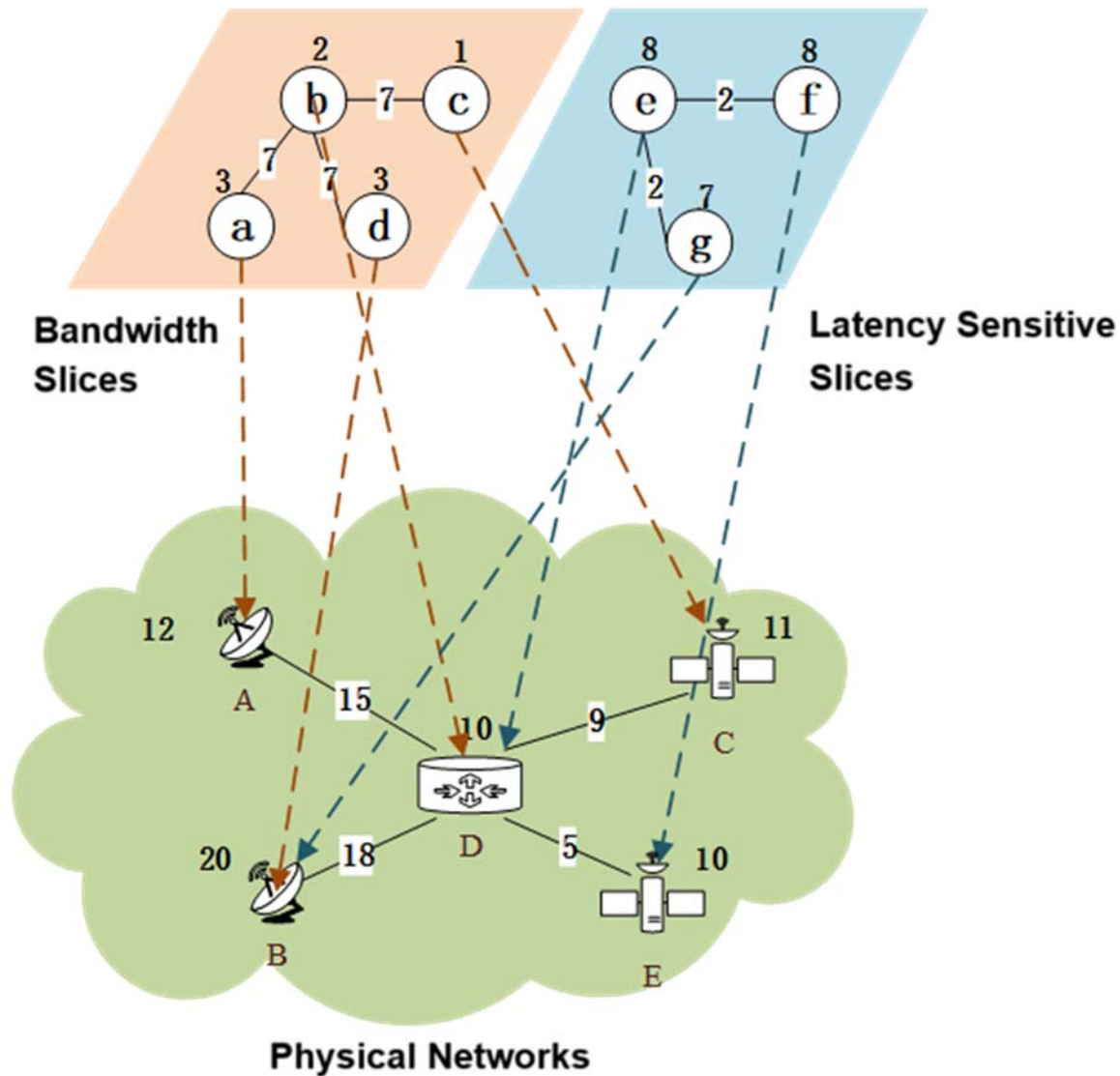


Figure 9: Schematic diagram of network resource slice generation

Figure 9 shows the process of generating network resource slices. Firstly, based on the principle of adjacency matrix feature vector decomposition, the fast matching of network weighted graphs is carried out to reduce the computational complexity of network resource slicing. Secondly, the adjacency matrix element substitution mechanism is used to reduce the minimum bandwidth requirement of network resource slicing. Finally, the balance of load and bandwidth is further optimized through the virtual network mapping optimization based on the hill climbing algorithm.

During the mapping of virtual links, neighbouring functions can be matched to both neighbouring and non-neighbouring nodes. In order to meet the above requirements in the process of link mapping, it is necessary to replace the matrix elements representing the link bandwidth in the physical topological adjacency matrix, select an optimal path for each pair of nodes, and update the matching weight. In the present document, a path is calculated for any pair of nodes, so that the ratio of the minimum bandwidth to the transmission hop number of the path is maximized, and the bandwidth hop ratio is updated to the link matching weight.

In the present document, the hill-climbing algorithm is used to optimize the matching results, and a pair of vertices are tried to be exchanged in each iteration, and a smaller exchange method is found under the premise of satisfying the constraints, until there is no better exchange method.

6.2 Mathematical model based on the slices of the optimal weighted graph matching

Based on the principle of adjacency matrix feature vector decomposition, the present document uses the optimal weighted graph matching algorithm to calculate the function and the mapping relationship of links, and describes the resource slicing requirements of the space-ground collaborative network as a virtual topology map, which supports the simultaneous processing of multiple slice generation requests. The optimal weighted graph matching problem is modelled as follows.

$G_1 = (V_1, E_1)$ and $G_2 = (V_2, E_2)$ denote two weighted graphs with a number of n nodes, where V_1 and V_2 denote the set of vertices, E_1 and E_2 denote the set of edges, and if there is a mapping function P such that the G_1 difference between the newly generated graph and the vertices and edges in G_2 is minimized after the function is mapped, then P is called the optimal weighted graph matching function of G_1 and G_2 . The optimal weighted graph matching problem can be expressed as equation (1), where A_1 and A_2 denote the adjacency matrices of the two graphs, respectively.

$$\text{Minimize } J(P) = \|PA_1P^T - A_2\| \quad (1)$$

Eigenvalue decomposition of equation (1) and equivalence to equation (2), the weighted graph matching problem is transformed into an assignment problem with $\bar{U}_2\bar{U}_1^T$ as the efficiency matrix.

Finally, the optimal mapping function can be obtained by processing the $\bar{U}_1\bar{U}_2^T$ by the Hungarian algorithm, where \bar{U}_1 and \bar{U}_2 are the absolute value matrices of the eigenvectors of A_1 and A_2 , respectively, $tr(x)$ represents the trace of matrix x , $H(x)$ indicates that the efficiency matrix x is solved by the Hungarian algorithm.

$$\text{Maximize } tr(P^T\bar{U}_2\bar{U}_1^T) \quad (2)$$

$$P = H(\bar{U}_2\bar{U}_1^T) \quad (3)$$

6.3 Simulation and verification of network resource slicing generation technology

Figure 10 shows the relationship between the average response time of different schemes to service requests and the size of network slices, and the red line is the proposed scheme. Since the computational complexity of this scheme is only related to the number of nodes in the physical topology, the average response time does not change with the increase of network slice size, and the Hungarian scheme is used to solve the efficiency matrix optimally, and the mapping function is modified by the hill climbing algorithm.

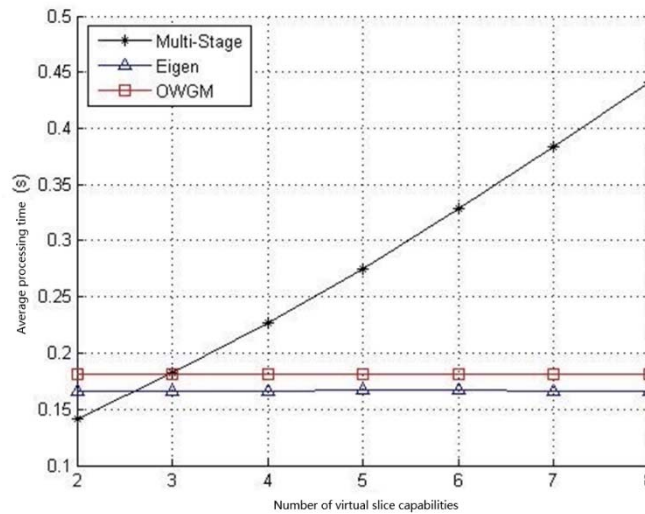


Figure 10: Response time of service requests for each solution

Figure 11 and Figure 12 show the balance performance of each scheme on node load and link bandwidth, respectively, and the evaluation indexes are the variance of the remaining processing capacity of the node and the remaining bandwidth of the link, respectively. Because the adjacency matrix is isomorphised in the matching process, the optimal matching can be achieved with a high probability, and the load balancing and bandwidth balancing are significantly better than those of other schemes, and the mapping function is modified by the hill climbing scheme to further optimize the load and bandwidth balancing effect.

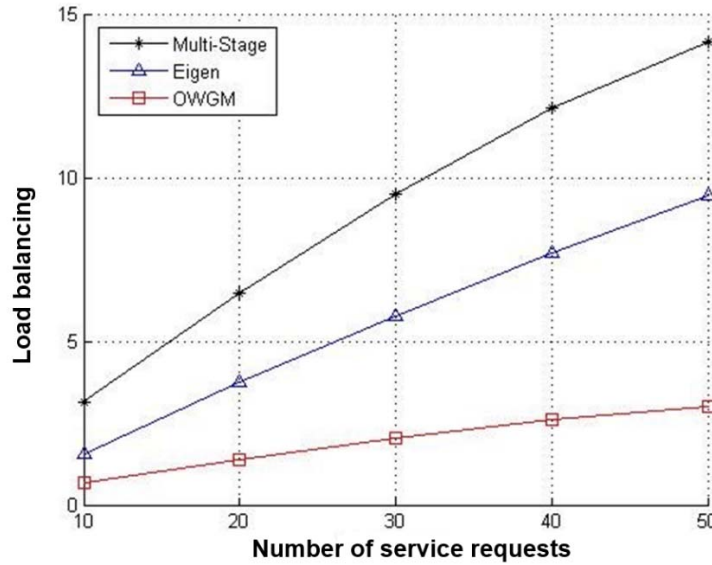


Figure 11: Load balancing performance of nodes in each solution

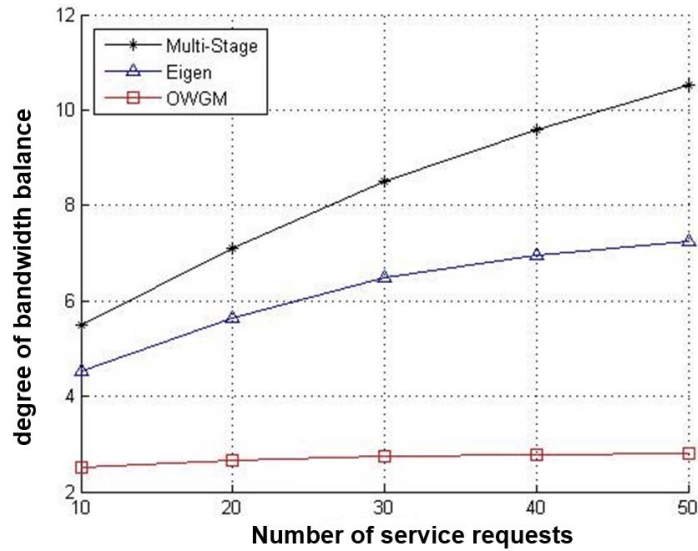


Figure 12: Balanced performance of link bandwidth in each scheme

Figure 13 shows the simulation results of the maximum number of service requests that can be supported by the three schemes, and it can be seen that the proposed scheme can support significantly more service requests due to the reduced bandwidth loss and better load and bandwidth balancing performance. In addition, the larger the network slice scale, the more obvious the advantages of this solution, because the larger the network slice, the higher the load and bandwidth balancing performance. In addition, the larger the network slice scale, the more obvious the advantages of this solution become, as its superior load and bandwidth balancing capabilities provide greater gains in larger, more complex environments

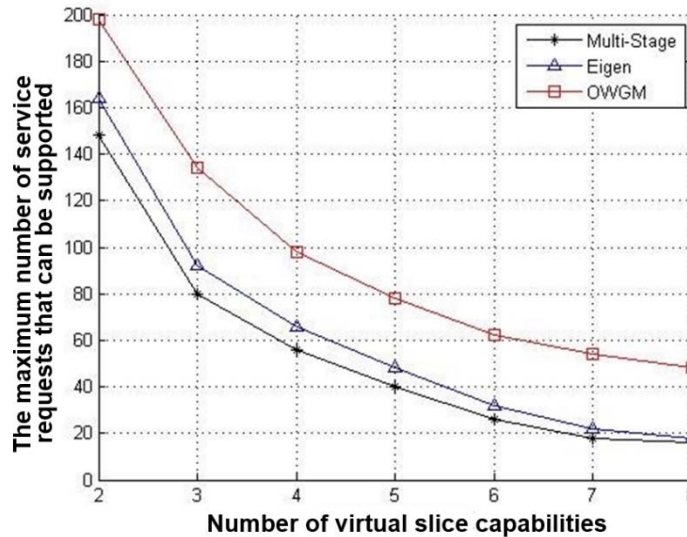


Figure 13: Maximum number of service requests per scenario

6.4 Network resource slicing scheduling demand forecasting technology

6.4.1 Introduction

In order to cope with the dynamic fluctuation of network traffic, network administrators need to dynamically and flexibly schedule the number of slice service instances, that is, slice capacity, based on actual resource requirements. For any network slice, the number of Virtualised Network Functions (VNFs) deployed on physical nodes depends on the real-time traffic demand, so it is important to accurately predict the immediate demand of the slice traffic and functional instances, and allocate network resources on demand in advance. The existing network slice traffic prediction methods model these two features separately and ignore the potential interaction between spatial and temporal features, resulting in limited prediction accuracy and insufficient basis for resource scheduling. Therefore, how to capture the spatiotemporal dependence of traffic at the same time is an urgent problem to be solved in the demand forecasting of slice scheduling.

6.4.2 Mathematical model for slice prediction based on spatiotemporal correlation

As shown in Figure 14, in the scenario where only a single network slice is considered, the slice capacity prediction problem is to predict the number of VNF instances on the corresponding virtual node at a certain time in the future based on the historical information of the data flow flowing through each virtual node on the slice at different times, and generate a view of the VNF resource capacity demand.

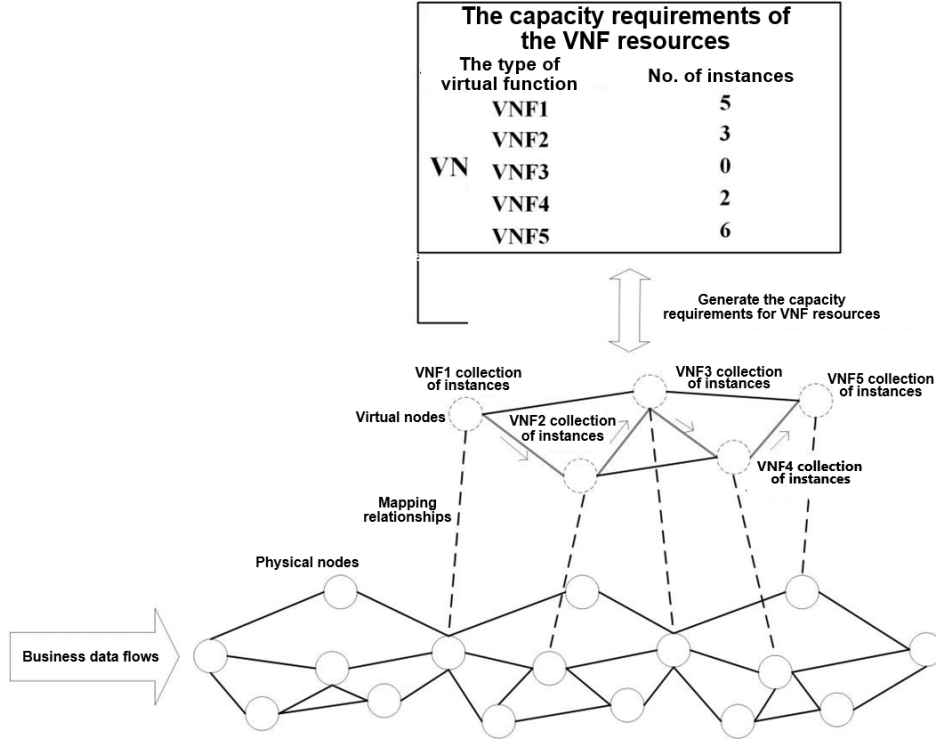


Figure 14: Capacity prediction model in a single-slice scenario

The graph in Figure 14 counts the number of nodes of the virtual network in the slice as n , the number of VNF instances on a node is counted as a characteristic of the node. The number of VNF instances on a node at time i is recorded as X_i^n , then the characteristics of all nodes of the whole network at time ii can be represented by the n -dimensional vector $X_i = (X_i^1, X_i^2, \dots, X_i^n)$, $X_i \in X_i^n$. The length of the historical sequence of the number of VNF instances on each node in the network is denoted as l . Where l represents the number of features entered into the model by a single node. Then the node characteristics of the whole network are represented by matrix $X^{n \times l}$, denote $f(\cdot)$ as the predictive model to be learned, T is denoted as the length of the time series for which forecasting is required, when $T = 1$, only the number of VNF instances at the next time of the node is predicted. The network slice capacity prediction problem can be represented as follows:

$$[X_{t-l}, \dots, X_{t-1}, X_t] \xrightarrow{f(\cdot)} [X_{t+T}] \quad (4)$$

$$[X_{t-l}, \dots, X_{t-1}, X_t] \xrightarrow{f(\cdot)} [X_{t+1}, \dots, X_{t+T}] \quad (5)$$

6.4.3 Design of network resource slice migration scheme

The output action of Dynamic Migration Mapping (DMM) calculation is the network slice capacity mapping policy, that is, to determine which node the VNF instance is placed on. Since the action corresponds to the resource margin of the different nodes, the output action is a continuous action. The PPO algorithm based on the Actor-Critic framework is good at the processing of continuous action output, the DRL framework is implemented by the PPO algorithm, and the Multilayer Perceptron (MLP) is used as the neural network of the agent.

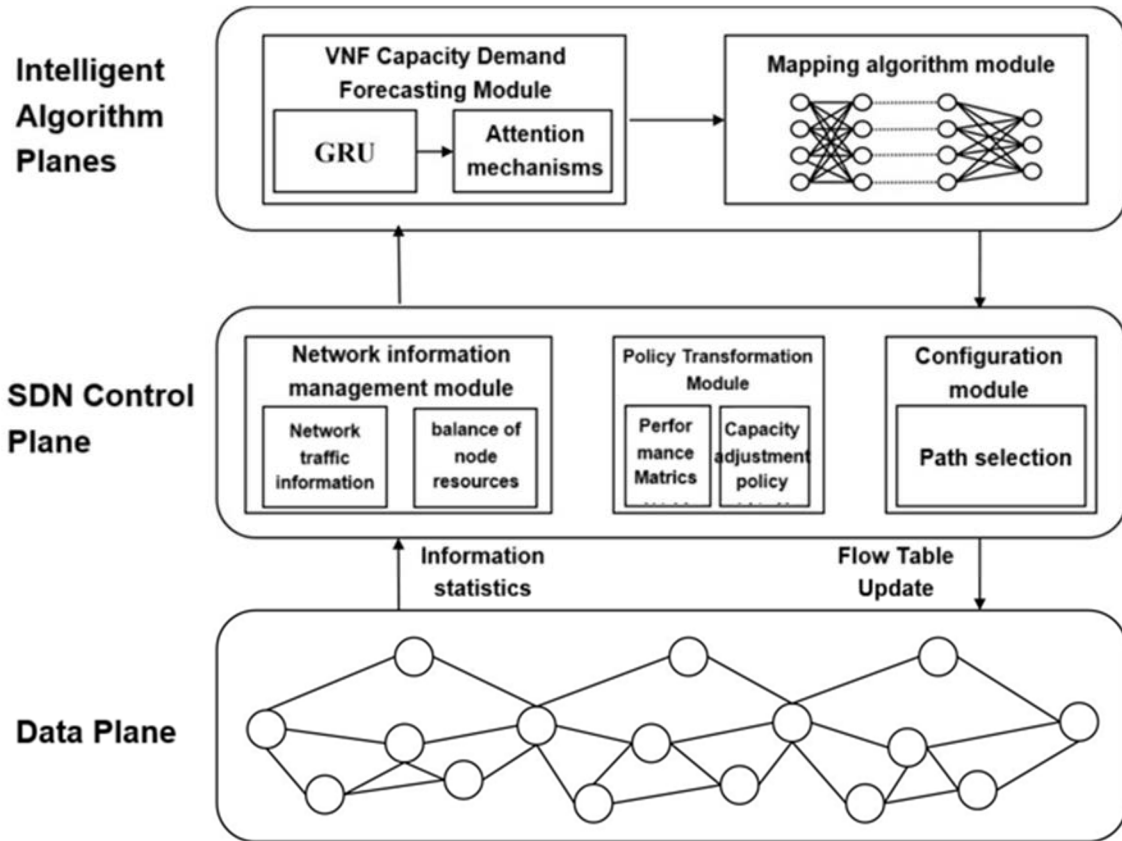


Figure 15: Dynamic migration mapping for network slicing

The DMM method will collect the resource margin information of network nodes, and uses this information as the state information of the algorithm environment together with the network slice capacity prediction information. At the same time, in the policy network (actor network), collects the output action generated by MLP, according to the mapping relationship between input and output. After the action the DMM method intervenes in the state information, and the reward value after the network slice capacity mapping is calculated. The agent updates the strategy based on the reward.

7 Conclusion and recommendations

Conclusion

Space-ground cooperative network slicing represents a transformative paradigm for achieving seamless, scalable, and reliable connectivity across terrestrial, aerial, and satellite networks. By integrating Non-Terrestrial Networks (NTNs) with 5G/6G infrastructures, this approach addresses critical challenges such as coverage gaps in remote areas, dynamic resource allocation for heterogeneous services, and Ultra-Reliable Low-Latency Communication (URLLC) for mission-critical applications. ETSI standards, particularly those developed by the Network Functions Virtualisation (NFV) and Multi-access Edge Computing (MEC) groups, provide foundational frameworks for orchestrating slices across hybrid domains. However, the dynamic nature of satellite orbits, spectrum sharing complexities, and latency asymmetries between space and ground segments introduce unique challenges. Current ETSI specifications need to evolve to address multi-domain slice lifecycle management, cross-layer security, and interoperability between NTN and terrestrial Radio Access Network (RAN)/core components. Furthermore, the lack of standardized interfaces for real-time coordination between satellite operators and terrestrial providers hinders end-to-end Service-Level Agreement (SLA) compliance [i.4].

Recommendations

To realize the full potential of space-ground slicing: collaboration with 3GPP, ITU-T, and satellite standardization bodies to define unified architectures is needed. Key priorities include:

- 1) **Interoperability Standards:** Develop APIs and protocols for seamless interaction between satellite payloads, ground stations, and 5G core networks, ensuring slice continuity during satellite handovers.

- 2) **Dynamic Slice Orchestration:** Enhance Management and Orchestration (MANO) to support predictive resource allocation using AI/ML, accounting for satellite mobility and intermittent connectivity. This includes integrating federated learning for distributed SLA assurance.
- 3) **Spectrum Harmonization:** Advocate for regulatory alignment on shared spectrum usage, prioritizing Ka/Q/V bands for high-throughput satellites and guard bands to mitigate interference.
- 4) **Security-by-Design:** Embed zero-trust principles into slice isolation mechanisms, leveraging quantum-resistant encryption for space links and blockchain for auditable slice lifecycle transactions.
- 5) **Sustainability:** Promote energy-efficient slicing algorithms to minimize satellite power consumption and orbital debris.

By bridging standardization gaps, a cohesive ecosystem where space and ground networks operate as a unified, SLA-driven fabric can be catalysed.

History

Document history		
V4.1.1	August 2025	Publication