NIDA-TS-005-2025

NIDA/TC/WG01/TF11

Date: 2026-2-5

Technical Requirements of Metropolitan Area Networks for Computing Services

NIDA copyright office
Email: contact@nida-alliance.com
Website: www.nida-alliance.com

Published in Shenzhen

Contents.

Foreword

The Network Innovation and Development Alliance (NIDA) is a voluntary international, industry-specific, non-profit social organization comprised of industry organizations, universities, research institutes, companies, and other entities from around the world, dedicated to promoting fixed network technology innovation and industrial upgrading.

The work of preparing NIDA Standards is normally carried out through the NIDA Technical Committee (TC), its Working Groups (WGs), and Task Forces (TFs). The procedures used to develop this document and those intended for its further maintenance are described in the **NIDA Standards Development Guidelines**.

Drafting organizations: China Telecom Research Institute, Zhongguancun Ultra Cross Connection New Infrastructure Industry Innovation Alliance, Internet Association of Kazakhstan, Huawei Technologies, ZTE Corporation.

Principal drafters: Yongqing Zhu, Bo Yuan, Zehua Hu, Shavkat Sabirov, Jie Dong, Haidong Zhu, Xiaowei Ji, Ruijuan Li, Fan Zhang, Jiayuan Hu, Tanxin Pi and Jun Zhang.

Copyright Notice:

This document is the property of the NIDA and is protected by copyright laws and international treaties. Unless otherwise stated, no part of this document may be reproduced, modified, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the NIDA.

This document was developed and is maintained by the **NIDA**. The content of this document is made available solely for use by authorized members and users for purposes consistent with the goals of the NIDA. For authorization, contact: contact@nida-alliance.com.

Patent Statement:

NIDA draws attention to the possibility that the implementation of this document may involve the use of (a) patent(s). NIDA takes no position concerning the existence, evidence, validity or applicability of any claimed patent rights in respect thereof. As of the date of publication of this document, NIDA may receive notice of (a) patent(s) which may be required to implement this document. However, implementers are cautioned that this may not represent the latest information, which may be obtained from the relevant patent database. NIDA shall not be held responsible for identifying any or all such patent rights.

Any trade name used in this document is information given for the convenience of users and does not constitute an endorsement.

For an explanation of the voluntary nature of standards, the meaning of NIDA specific terms and expressions related to conformity assessment, please refer to the relevant NIDA IPR documents.

This document was prepared by the Technical Committee (TC), Working Group WG 01, Network Evolution working group, Task Force TF ##, Technical Requirements of Metropolitan Area Networks for Computing Services.

Any feedback or questions on this document should be directed to the NIDA TC Secretariat.

Technical Requirements of Metropolitan Area Networks for Computing Services

1 Scope

This document specifies NIDA's technical requirements for Computing service Oriented Metropolitan Area Networks(hereinafter referred to as COMAN).

It applies to the planning, design, construction and acceptance of COMAN.

2 Normative references

All normatively referenced materials in the document listed as normative references are considered part of this document. For a dated reference, only the edition corresponding to that date applies. For an undated reference, the latest edition of the referenced documents (including all amendments) applies.

3 Terms and definitions

3.1 Definitions of Computing Power Metrics

3.1.1

Effective Computational Power Rate

Effective Computational Power Rate refers to a metric used in intelligent computing clusters to measure the utilization efficiency of the actual computing power of a DC or computing platform. It usually refers to the ratio of computing power actually available for computational tasks to the total available computing power.

3.1.2

Computing Power Availability Rate

Computing Power Availability Rate refers to the ratio of the time during which computing resources are actually available for computational tasks to the total time during a training task. It reflects the availability of computing resources and is an important metric for evaluating the resource utilization of a DC or computing platform.

3.1.3

Collective Communication Efficiency

Collective Communication Efficiency refers to the efficiency of communication operations in which multiple compute nodes (or processes) in a parallel computing environment. These communication operations include Broadcast, Gather, All-Gather, Scatter, Reduce, All-Reduce, and All-to-All et al. Collective Communication Efficiency directly affects the performance of parallel computational tasks.

3.2 Definitions of Network Terms

3.2.1

Network Communication Efficiency

Network Communication Efficiency refers to the efficiency of data transmission in an AIDC network environment. It is generally used to evaluate the rate, accuracy, and resource utilization of data transmission over the network, covering multiple aspects such as throughput, packet loss ratio, latency, and bandwidth utilization.

3.2.2

Throughput

Throughput refers to the amount of data successfully transmitted over a network per unit time, typically measured in bits per second (bps).

3.2.3

Latency

Latency in this document refers to the E2E delay that a network is required to guarantee for different computational tasks.

3.2.4

Packet Loss Ratio

Packet Loss Ratio refers to the ratio of packets that fail to reach their destination to the total number of packets transmitted.

3.2.5

Bandwidth Utilization

Bandwidth Utilization refers to the ratio of used network bandwidth to the total network bandwidth. A high bandwidth utilization indicates that network resources are effectively used.

4 Abbreviations

The following acronyms and abbreviations apply to this document.

AI: Artificial Intelligence

AIDC: Artificial Intelligence Data Center

ASIC: Application–Specific Integrated Circuit

CaaS: Computing as a Service

COMAN: Computing service Oriented Metropolitan Area Network

CPU: Central Processing Unit

DC: Data Center

E2E: End–to–End

ECN: Explicit Congestion Notification

EDN: Enhanced Deterministic Network

FLOPS: FLoating–point Operations Per Second

FPGA: Field Programmable Gate Array

GPU: Graphic Processing Unit

IB: InfiniBand

MAN: Metropolitan Area Network

MTTR: Mean Time To Repair

NPU: Neural network Processing Unit

O&M: Operation and Maintenance

PUE: Power Usage Effectiveness

PFC: Priority–based Flow Control

POD: Point Of Delivery

POP: Point Of Presence

QoS: Quality of Service

RDMA: Remote Direct Memory Access

SLA: Service Level Agreement

TB: Terabyte

ZB: Zettabyte

5 Overview of Computing service Oriented Metropolitan Area Network(COMAN)

5.1 Development Trend of the Computing Infrastructure

With the rapid development and widespread application of general-purpose computing, intelligent computing, and supercomputing technologies, the demand for computing power is experiencing explosive growth. Currently, the computing infrastructure is developing toward diversification, intensification, intelligization, and popularization, and is becoming a key engine for the digital transformation across industries.

—In the past, computing power mainly relied on CPUs for general-purpose computing and was widely used in office, database, and basic service scenarios. With the development of AI, especially deep learning and large language model technologies, traditional CPUs can no longer meet the performance requirements of highly concurrent, low-latency matrix operations. GPUs, due to their strong parallel processing capabilities, are therefore widely adopted as the mainstream computing carriers for AI training and inference. Meanwhile, dedicated chips such as NPUs, FPGAs, and ASICs are continuously emerging, providing customized task-specific acceleration and optimization. In addition, supercomputing continues to play a critical role in some areas such as meteorological forecasting, biomedicine, and aerospace. In the future, multiple forms of computing power will coexist and collaborate, gradually forming a multi-layered and heterogeneous computing infrastructure.

—With the rapid growth of computing demand, the traditionally siloed DC construction mode is not sustainable. Issues such as fragmented resources, low utilization, and regional imbalance greatly limit the effective utilization of computing power. To address these issues, China is constructing national computing power service hubs, aiming to achieve a coordinated layout and schedule of computing resources nationwide. At the same time, ISPs(Internet Service Providers) and CSPs(Cloud Service Providers) develop some new service modes such as "computing private networks", which support the visualization, scheduling, and metering of computing resources. With the preceding measures, computing power service is transforming from purely providing infrastructure resources to comprehensive services, further accelerating the intensification of computing infrastructure.

—As computing power infrastructure continues to scale out, traditional O&M, which mainly relies on the expertise and experience, is hard to cope with increasingly complex resource scheduling and troubleshooting. Intelligence has become a key enabler for improving the operational efficiency of computing systems. AI technologies are being gradually applied across the lifecycle of computing resource management. In terms of scheduling, reinforcement learning and predictive models enable intelligent task allocation and load balancing, preventing local overload or resource idleness. In terms of O&M, big data analytics and anomaly detection algorithms support fault prediction, root cause analysis, and automated recovery, significantly improving system availability. In terms of energy efficiency management, AI is used to dynamically adjusts CPU frequency, cooling system power, and power supply policies, effectively reducing PUE and promoting green, low-carbon, and sustainable development of computing infrastructure.

– As the CaaS model is becoming more and more mature, computing power service mode is accelerating its shift from "self-built and self-used" to "service-oriented, platform-based, and market-driven". For most enterprises and research institutions, building and maintaining AI facilities requires substantial investment. Against this backdrop, computing power leasing, with its plug-and-play and flexible supply model, significantly lowers the barriers to access and use computing resources, as well as costs. Enterprises can obtain required computing resources more quickly by leasing AI facilities from third-party AIDCs, accelerating the process of technological research and innovation. Computing power leasing effectively promotes the inclusive development of computing technologies, and provides strong support for the widespread application and continuous innovation of computing services[1].

5.2 Key Scenarios of Computing Services

With the rapid growth of computing power demand, computing power leasing is gradually becoming the mainstream supply model, multiple kinds of new computing services have emerged, such as transmitting sample data to AI cluster, transmitting sample data to AI servers, coordinated model training across AIDCs, and coordinated model inference between the customer and AIDCs. Metropolitan Area Networks(MANs), as the critical infrastructure for connecting customers and computing resources, provide essential network support for various computing services. MANs ensures those customers utilizing remote computing resources to obtain a user experience comparable to that of locally deployed computing resources.

5.2.1 Large-scale Sample Data Transmission to AI cluster

With the rapid development of AI/HPC, data volumes are growing at an unprecedented speed, and a single data transfer from an enterprise to AIDCs can reach hundreds of terabytes. International Data Corporation (IDC) estimates that global data output will reach 213.56 ZB by 2025 and 527.47

---

[1] computing services refer to services that meet the high-performance computing needs of individuals, enterprises, or institutions by providing, scheduling, and managing computing power.

ZB by 2029. Currently, many customers still rely on shipping physical hard drives to transfer large-scale data. Model training ususally needs tens of petabytes of sample data. If transfer such large scale data to AIDCs in this way., it will be very inefficient and costly. Since customers usually need the data transmission service intermittently in the period of the AI training, the traditional private lines service with fixed bandwidth such as 100 Mbps, 10 Gbps, is also inadequate for this. Therefore, there is an urgent need to upgrade the capalibilty of MANs to provide more efficient and cost-effective data transmission services.

5.2.2 Large-scale Sample Data Transmission to AI server

Data security requirements are common in intelligent computing scenarios across many sectors. For example, crash test and accident data in the automotive industry, as well as citizen identity and legal entity information involved in the government, should be strictly controlled and used in a secure way. These entities have strict security standards on sample data management and require core data to be stored on-premises. When leasing third-party computing resources, these entities adhere to the principle of localized data storage so as to avoid data leakage during model training. In these scenarios, compute nodes and data storage nodes are deployed across MANs(which means they rely on MANs to achieve high-frequency, real-time interactions). Only the required sample data are uploaded to AI servers during model training, and the sample data shall not be kept in disks in AIDCs during the AI training process.

5.2.3 Coordinated model training across AIDC

The scaling law for foundation models remains in force. Over the past decade, the computing power required by foundation models has increased by approximately one million times and is expected to continue growing at a rate of more than four times per year. The scale of computing power within a single DC is constrained by many factors such as power supply and physical space. To meet the rapidly growing computing demand of foundation models, it is imperative to promote collaborative training across multiple AIDCs over MANs, so as to integrate geographically dispersed computing resources for the AI training task. In addition, most AIDCs in China are relatively small (over 70% of them have a scale of 100—300 PFLOPS) and are mainly distributed across research institutions, local governments, and cloud service providers. Integrating dispersed computing resources helps overcome limitations imposed by physical infrastructure and service providers, thereby enabling the creation of a unified computing service platform.

5.2.4 Coordinated model inference between Customer and AIDCs

The significant reduction in LLM inference costs has driven enterprises to deploy AI facilities on-premises, enabling rapid adoption and application of AI. However, customer on-premises computing resource face challenges such as capacity bottleneck and maintenance cost, making it hard to meet the computing power demand of continuously growing model inference. Therefore, efficient collaboration between the customer and AIDCs provides a more effective, convenient, and cost-efficient solution for enterprises to flexibly get more computing resources. This solution is based on the split learning, in which the model is partitioned and processed in parallel across the customer and AIDCs, with gradient parameters and other data efficiently synchronized over MANs. In addition, this solution supports the on-premises deployment of the input and output layers of LLMs, ensuring sample data remains within customers' own campuses or facilities.

5.3 Challenges Faced by Metropolitan Area Network

MAN connect heterogeneous computing resources and diverse terminals within a region, serving as a critical infrastructure for the sustainable development of the AI industry. To deliver computing service with one-click access and instant service—much like water and electricity services in the modern city—MAN faces the following three major challenges:

—**Challenge 1: bottlenecks in data transmission efficiency and stability:** LLMs training and knowledge base building typically requires TB/PB-scale data, placing a much higher demand on MAN capacity. At the same time, parameter synchronization has extremely stringent requirements on network performance. Low network quality may result in problems such as insufficient data transmission rate, excessive latency, and frequent packet loss, affecting the availability of computing resources. In addition, the development of some new computing services based on MCP and A2A requires MAN to efficiently and reliably steer the traffic among AI agents.

—**Challenge 2: difficulty in adapting O&M systems to the dynamic of AI services:** AI traffic exhibit dynamic traffic characteristics marked by the mix of elephant flows and mouse flows, and the coexistence of bursty and steady traffic. Moreover, AI services have extremely low tolerance for faults, even a minor fault may trigger task rollback and directly degrade user experience. Conventional O&M that relies on manual troubleshooting typically has minute-level MTTR, failing to meet the stringent requirements of AI services. In addition, existing O&M systems show clear deficiencies in computing power scheduling. They are incapable of unified sensing or elastic scheduling of distributed and heterogeneous computing resources, resulting in low computing resource utilization.

—**Challenge 3: complexity in building data security and trust mechanisms:** AI services involve large volumes of sensitive data, and traditional filtering-based AAA systems and encryption technologies are insufficient to defend against targeted attacks or ensure full-lifecycle data security. Furthermore, MAN carries ToH, ToB and ToC services simultaneously, requiring a balance between strict service isolation and flexible performance guarantees. Traditional VLAN isolation or VPN technologies lack the ability to dynamically adjust bandwidth, paths, and other network parameters, making it difficult to balance resource sharing and isolation. Moreover, emerging security technologies such as blockchain and quantum encryption have not been effectively adapted to or integrated with MAN, leading to a high costs and long period for building a trusted network.

To address these challenges, COMAN construction must consider communication efficiency, maintainability, reliability, and security, to ensure key computing metrics such as the effective computational power rate, computing power availability rate, and collective communication efficiency.

6 Architecture of COMAN

6.1 Overall Architecture

COMAN should adopt a modular "Building-block" architecture to enable flexible network adjustment and compatibility with heterogeneous components, thereby ensuring smooth upgrades and sustainable evolution of the network. From a functional perspective, COMAN consists of three major components: computing service oriented POD, computing service oriented POP, and computing service oriented interconnection zone (hereinafter referred to as “POD, POP, and interconnection zone” in this document), along with a compute-network brain. Figure 1 shows the overall architecture of the COMAN.

When there are few service requirements or the city scale is small, network components can be simplified. In such cases, the network can be deployed using spine-leaf or other architectures to quickly meet initial service transmission requirements.

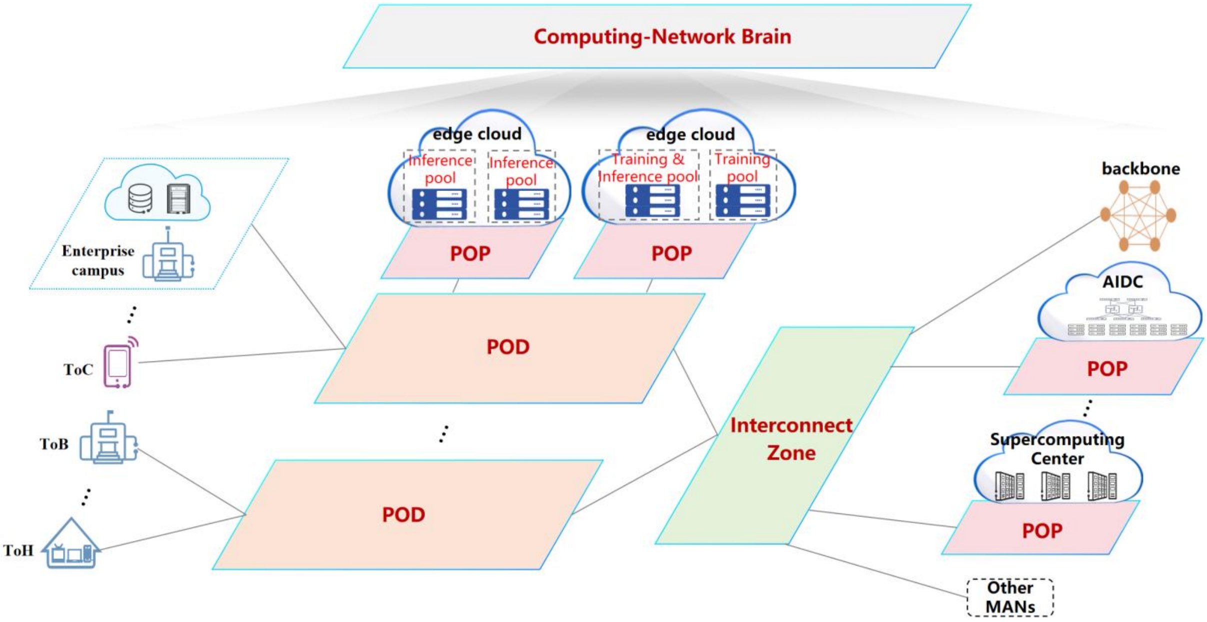

Figure 1 — Overall architecture of the COMAN

—**POD:** serves as a core unit for the converged carrying of all services. It provides wide coverage and unified access for various computing customers, and also enables fast access to edge computing resource pools.

—**POP:** connects COMAN and heterogeneous computing resource pools, enabling standard and rapid interconnection between them in parameter, sample, and service planes.

—**Interconnection zone:** connects COMAN and external networks, such as supercomputing centers and backbone networks. It is also responsible for unified traffic steering between PODs and POPs, as well as the inbound and outbound traffic of COMAN.

—**Compute–network brain:** acts as the "scheduling brain" and "management hub" of COMAN, enabling unified control, coordinated scheduling, and intelligent O&M of computing and network resources.

Efficient collaboration among these components ensures efficient transmission of computing services over COMAN. POD serves as the access entry and, through high–speed interconnection with POP, establishes efficient transmission channels between customers and computing resource pools. In collaboration with the interconnection zone, POD provides E2E lossless connections between customers and cross–domain computing resource pools. Meanwhile, these components can seamlessly collaborate with the compute–network brain through standard interfaces, enabling dynamic perception, intelligent orchestration, and full–lifecycle O&M and management of computing resources.

6.2 Core Capabilities of Each Component

6.2.1 Computing service oriented POD

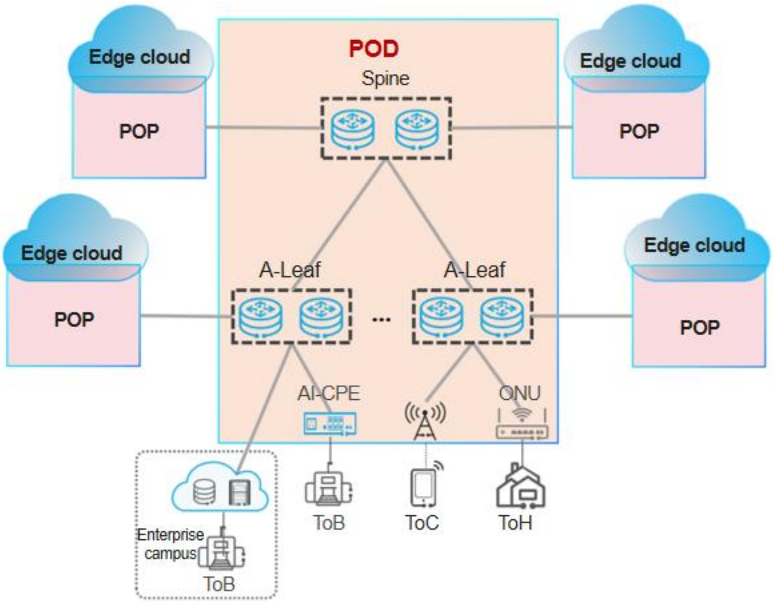

Figure 2 — Computing service oriented POD

Note: Figure 2 uses a Spine—Leaf architecture as an illustrative example only.

This component serves as the edge access layer of COMAN, enabling converged connectivity for customer terminals, enterprise branches, and home subscribers. By aggregating traffic through hierarchical device convergence, it forms a wide–coverage, elastic, and flexible computing service entry. Its core functions are as follows:

–Provides ubiquitous access for computing terminals and enterprise branch sites across multiple types of access media (including dark fibers, PONs, and 5G) and can be flexibly deployed based on factors such as access locations, administrative regions, and AIDC service coverage.

–Realizes unified access and converged transport for fixed, mobile, cloud, and computing services, ensuring non–blocking forwarding of various types of service traffic.

–Provides customers with computing channels featuring high throughput, low latency, low packet loss, and high reliability to ensure data transmission quality.

–Realizes fast access to edge computing resource pools in collaboration with POPs. This enables pooled utilization and flexible scheduling of edge computing resources, delivering low–latency, high–quality computing services to customers.

6.2.2 Computing service oriented POP

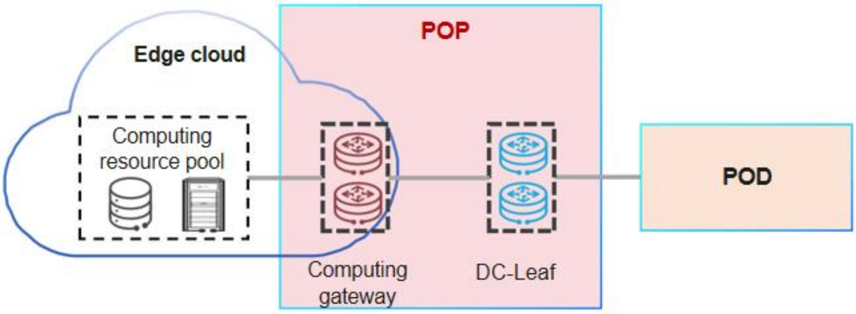

Figure 3 — Computing service oriented POP

This component implements rapid interconnection between COMAN and computing resource pools through standard networking between DC–Leaf and computing gateway devices. It supports converged multi–service carrying and flexible computing power scheduling. Its core functions are as follows:

—Provides standard, rapid interconnection between COMAN and self–owned/third–party heterogeneous computing resource pools in the sample, parameter, and service planes, ensuring "minute–level" computing service provisioning and "millisecond–level" service access latency.

—Leverages SRv6 to enable efficient forwarding of computing service traffic and E2E service programmability, delivering high–speed and stable computing services to customers.

—Interconnects with provincial or regional POPs to provide efficient connections between inter–domain computing resource pools, enabling collaborative computing across domains.

6.2.3 Computing service oriented Interconnection Zone

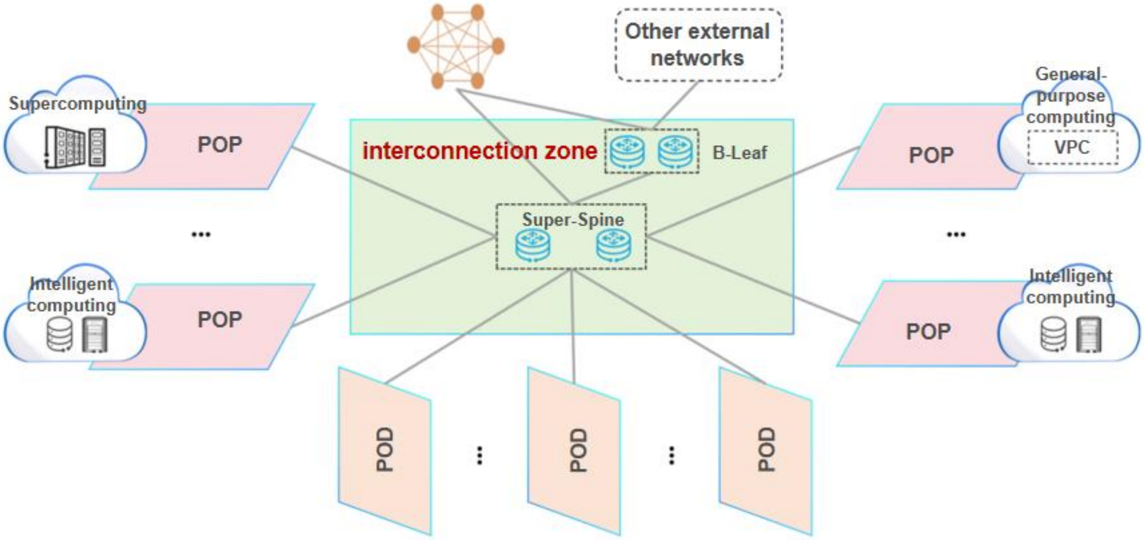

Figure 4 — Computing service oriented interconnection zone

Note: Figure 4 uses a Spine—Leaf architecture as an illustrative example only. In this example, the Super-Spine handles traffic exchanged between PODs, between heterogeneous computing resource pools, and between PODs and computing resource pools. It also interconnects with the backbone network. The B-Leaf handles cross-domain or cross-POD computing service traffic and provides interconnection to other external networks.

This component, as the traffic hub of COMAN, provides interconnection between PODs, POPs, computing resource pools and external networks. Its core functions are as follows:

—Flexibly connects components such as PODs and POPs in a scalable way even without changing existing connections to external networks.

—Provides unified and rapid steering for the inbound and outbound traffic of COMAN, as well as traffic between PODs, simplifying network hierarchy.

—Provides fast interconnections between heterogeneous computing resource pools, ensuring efficient integration and coordinated scheduling of heterogeneous computing resources.

6.2.4 Compute-network Brain

Figure 5 — Compute-network brain

This component is an integrated platform for unified scheduling, management, and optimization of network resources and service processes. Serving as the "control hub" of COMAN, it provides core functions such as resource scheduling, service orchestration, performance monitoring, fault management, and O&M. Its core functions are as follows:

—Resource management: integrates network information (e.g., topology and bandwidth status) with computing information (e.g., node types and load status) to enable unified modeling, dynamic sensing, and visualized management of computing and network resources.

—Resource scheduling: intelligently allocates optimal compute nodes and network resources based on service SLA requirements and real-time resource status and supports task migration, elastic scaling, and cross-domain collaboration, so as to improve resource utilization and service quality.

—Service orchestration: uses standard templates and automated workflows to enable rapid deployment and E2E integration of fixed, mobile, cloud, and computing services, enhancing service delivery efficiency and flexibility.

—Operational assurance: leverages digital twins, AI analytics, and other intelligent O&M methods to predict, locate, and automatically rectify faults, and proactively identifies and quickly responds to security risks, ensuring secure and stable system operation.

7 Key Technical Capability Requirements for COMAN

To clarify the capabilities required for each module, this chapter details the key technical requirements for the components of COMAN: the computing service oriented PODs, POPs and interconnection zone must support lossless and deterministic data transmission to ensure high-quality for computing services; the compute-network brain must ensure computing resource utilization efficiency and user experience comprehensively through capabilities in resource allocation, path optimization, and SLA assurance.

7.1 Key Technical Requirements for Computing service oriented POD

This component can be flexibly deployed according to service scale, AIDC service coverage, and administrative regions to accommodate the rapidly growing service demand. In addition, it can connect to various terminals and enterprise branch sites, providing unified access and converged carrying for fixed, mobile, cloud, and computing services while ensuring non-blocking forwarding of traffic within the region. To ensure a high computing efficiency, this component must also provide low-latency, elastic, and lossless data transmission channels.

This component must meet the following key capability requirements:

1. Integrated access and converged carrying: This component can provide integrated access and converged carrying for fixed, mobile, cloud, and computing services based on the unified SRv6/EVPN protocol stack. Customers can access network resources, cloud resources, and AIDCs through a single access point. This implementation ensures a consistent service experience, lowers network complexity and O&M costs, and improves deployment efficiency and user experience.
    (1) It can support multiple interconnection mechanisms (such as private lines and SD-WAN) over various access media (such as dark fibers, PONs, and 5G), providing seamless connections between enterprise/home LANs and third-party cloud/computing resources.

This enables the creation of logically unified and securely isolated VPNs for efficient and secure transmission of data and computing power.

(2) It can automatically identify and mark computing service traffic based on traffic characteristics as well as computing service type, model type, QoS, and security requirements. It can also work with the operations management system to dynamically allocate optimal forwarding paths and various network resources—including network slices and bandwidth—to computing services, enabling efficient transmission and SLA assurance for computing services.

2. Lossless transmission: This component can provide lossless transmission capabilities, effectively addressing the high sensitivity of RDMA transmission to packet loss(a packet loss ratio of 0.1% can degrade throughput by 50%). Besides the 400GE high-speed ports and large port buffers, other technologies such as fine-grained congestion control should be deployed to prevent the network from congestion and packet loss, ensuring efficient computing power transmission.

(1) COMAN devices can provide 400GE high-speed ports to meet the high-throughput and low-latency requirements of intelligent computing services. These devices can also provide GB-level port buffers and thousands of RDMA lossless queues and dynamically optimize buffers and queue allocation based on traffic changes to maintain network stability and transmission efficiency in scenarios involving high bursts, high concurrency, and heavy workloads.

(2) It can provide fine-grained congestion control capabilities. Specifically, it can dynamically allocate port buffers and priority queues per service or tenant, and send backpressure signals to upstream devices in an E2E or hop-by-hop manner for fine-grained traffic control. In addition, it can provide ECN, PFC, and other congestion control and optimization mechanisms to prevent and rapidly respond to congestion in a multi-dimensional and fine-grained manner in high-concurrency scenarios.

3. Ultra-high throughput: This component can provide ultra-high-throughput data transmission capabilities to ensure non-blocking transmission of east—west and north—south traffic. Moreover, it can use in-depth packet identification (capable of identifying IB transport layer information) to automatically identify elephant flows and split them into subflows, and further achieve the global traffic load balance traffic with SRv6. By leveraging these technologies, this component eliminates performance degradation caused by high-bandwidth, low-concurrency elephant flows, maximizes link utilization, and ensures a network throughput of over 90%.

4. High oversubscription: This component can support high oversubscription, efficiently transmitting cloud-edge collaborative training and inference services and dispersed computing integration services. Computing services feature high concurrency and bursty traffic, posing high requirements on network bandwidth. This component can use the "intelligent buffering & scheduling" mechanism to achieve high-convergence-ratio networking: the high-speed buffers instantly absorb traffic spikes to relieve burst pressure, while service priority-based dynamic scheduling ensures that critical traffic, such as control signaling and task data, is forwarded first, preventing computing resources from sitting idle. High convergence ratios such as 8:1, 16:1, and 32:1 can be introduced on demand to significantly lower the bandwidth requirements of enterprise private lines. In this way, the bandwidth costs can be optimized while the computing efficiency remains stable, enhancing both cost savings and computing efficiency.
5. Deterministic service: This component can provide deterministic service capabilities by combining technologies such as FlexE and deterministic IP slicing with SRv6 programmable paths and offer hierarchical network slicing services featuring both software and hardware guarantees for multiple tenants in an E2E, differentiated manner. Moreover, it can provide multi-layer deterministic capabilities—including resource reservation, path orchestration, and traffic scheduling—to ensure low-latency, low-jitter, and high-reliability transmission of computing services in scenarios involving complex service workloads and large-scale traffic concurrency.
6. Security encryption: This component can provide multi-layer encryption mechanisms such as IPsec and MACsec to ensure confidentiality and anti-tampering of data. Moreover, it can support high-speed encryption and decryption with mainstream international encryption standards, ensuring E2E encrypted transmission without compromising forwarding performance. Advanced security technologies such as quantum encryption and trusted execution environment can also be adopted to further enhance network security protection.

7.2 Key Technical Requirements for Computing service oriented POP

This component is constructed along with computing resource pools, enabling standard and rapid interconnection between the computing metro network and heterogeneous computing resource pools in the parameter, sample, and service planes of AI networks. It supports deep coordination between the service side and network side, achieving lossless transmission of RDMA traffic and deterministic multi-tenant SLA assurance, thus comprehensively ensuring stable, efficient operation of computing services such as training based on local storage and remote computing and cloud-edge collaborative training and inference.

This component must meet the following key capability requirements:

1. Standard interconnection: This component can implement seamless interconnection between the network and computing resource pools through standard networking between DC-Leaf devices and computing gateways (with physical links and logical channel resources provisioned in advance). On one hand, it can allow modular network components to be newly deployed or upgraded on demand according to the type and scale of connected computing resources and access locations. This ensures agile network expansion without the need for a complete architectural overhaul. On the other hand, it can flexibly adapt to the coverage scopes of self-owned/third-party AIDCs at different levels, enabling rapid integration and pooled management of heterogeneous computing resources. This supports smooth long-term network evolution, avoids large-scale network reconstruction caused by iterative computing services, and ultimately reduces O&M costs and architectural adjustment risks.
2. Terminal-network coordination: This component can implement deep coordination between AI servers and network, enhancing key capabilities such as congestion avoidance, traffic scheduling, and QoS parameter optimization. It can provide standard, structured interfaces and information exchange mechanisms (including unified data models, open southbound/northbound interfaces, and low-overhead real-time telemetry channels) between service status (e.g., service type, priority, and other information) and network status (e.g., topology, bandwidth utilization, and other information).
3. Lossless transmission: As a critical interconnection node, this component can provide lossless transmission capabilities aligned with those of PODs, ensuring E2E RDMA traffic transmission for computing services. Specifically, it can be equipped with DC-Leaf devices featuring 400GE ports and GB-level port buffers—along with fine-grained congestion control mechanisms—to prevent the network from packet loss and ensure high-performance transmission of RDMA traffic across domains.
4. Deterministic service: This component can provide deterministic service capabilities to ensure E2E deterministic service transmission in collaboration with other network components. On one hand, it can use technologies such as FlexE, deterministic IP slicing, and VPN to provide hierarchical network slicing services featuring both software and hardware guarantees for multiple tenants to meet service SLA requirements. On the other hand, it can provide multi-layer deterministic capabilities covering resources, paths, and services to ensure low-latency, low-jitter, and high-reliability data transmission in scenarios involving complex service workloads and large-scale traffic concurrency.

7.3 Key Technical Requirements for Computing service oriented Interconnection Zone

This component serves as a hub that connects COMAN to heterogeneous computing resource pools and external networks (such as the backbone network). Concurrently, it enables flexible expansion of PODs and POPs. By leveraging technologies such as 400G/800G high-bandwidth links and elephant-flow load balancing, this component can achieve a high throughput in the scope of the MAN. With mechanisms such as fine-grained congestion control and network slicing, this component can ensure lossless and deterministic data transmission for various services.

This component must meet the following key capability requirements:

1. High-bandwidth links: This component can have 400G/800G links deployed at scale to transmit massive data in sample plane and parameter plane to AIDCs. With the high-bandwidth links carrying aggregated traffic, it can improve bandwidth utilization and dynamic scheduling capabilities, thereby building high-capacity network infrastructure, reducing per-bit transmission costs, and paving the way for evolution to higher-speed technologies.
2. Ultra-high throughput: Many intelligent computing services involve a small number of elephant flows with extremely high data rate. This component can use in-depth packet identification (capable of identifying IB transport layer information) to automatically identify elephant flows and split them into subflows, and further achieve the fine-grained traffic engineering with SRv6. This maximizes network bandwidth utilization, ensuring a network throughput of over 90%.
3. High oversubscription: As the hub that interconnects various components, this component can support high oversubscription to cope with concurrent synchronization of large-scale data and sudden traffic bursts. High convergence ratios such as 8:1, 16:1, and 32:1 can be introduced on demand to ensure the transmission efficiency of computing power and achieve the optimal balance between network construction costs and computing efficiency.
4. Lossless transmission: For computing service scenarios where sample-plane and parameter-plane data must traverse long distances with minimal loss, this component can rely on 400GE/800GE ports, large port buffers, fine-grained congestion control, and thousands of RDMA lossless queues to prevent packet loss and ensure high-performance RDMA transmission across domains.
5. Deterministic service: This component can work with PODs and POPs to provide deterministic data transmission for various types of computing services. It can use technologies such as FlexE and deterministic IP slicing, combined with SRv6 programmable paths, to implement

low-latency path computation and congestion avoidance, ensuring differentiated E2E SLA assurance for various computing services. It can also build multi-layer deterministic capabilities covering resources, paths, and services to ensure low-latency, low-jitter, and high-reliability transmission of services in scenarios involving heavy workloads.

7.4 Key Technical Requirements for Compute-network Brain

This component is the intelligent control hub of COMAN. It centrally manages and intelligently schedules computing resources, network resources, and security policies within the domain through a central management and control platform, enabling collaboration and service association between computing and network resources. It provides automated services throughout resource allocation, path optimization, and SLA assurance, enabling self-healing, proactive security defense, and energy-efficiency optimization. It delivers committed, measurable, and operational integrated computing-network services, comprehensively improving computing resource utilization and service experience assurance capabilities.

This component must meet the following key capability requirements:

1. Elastic bandwidth: This component can elastically allocate network bandwidth based on service cycles, task characteristics, and QoS requirements. Specifically, it can dynamically allocate bandwidth at Mbps and Gbps levels on demand. This prevents resource idleness and network congestion while enabling agile response to traffic bursts, significantly improving resource utilization.
2. Task-based service: This component can implement on-demand bandwidth application and dynamic bandwidth assignment on a per-task basis and automatically configure bandwidth resources for E2E transmission paths according to customer requirements (e.g., minimum guaranteed bandwidth, peak bandwidth, priority). During traffic bursts, it can detect load changes through real-time load monitoring and prediction mechanisms and dynamically expand bandwidth to ensure stable transmission for high-priority services, maximizing overall network and computing resource utilization.
3. Network-level load balancing: Intelligent computing service traffic typically consists of a small number of flows with large traffic volumes. However, the current load balancing technology based on hashing field that requires a large number of flows to achieve statistical multiplexing, failing to meet the requirements of intelligent computing services. Therefore, this component needs to actively collect and analyze AI traffic patterns and perform unified planning and

scheduling from a global perspective to optimize network-wide load balancing, ensure high service throughput, and maximize overall network resource efficiency.

4. Network autonomy: This component can monitor network-wide information, including topology, load, device health, and traffic characteristics (elephant flows, congestion status, and so on) in real time to provide a reference for scheduling decisions. Moreover, it can support AI-driven self-healing by building fault diagnosis models using deep learning and knowledge graphs, enabling root cause inference and risk identification for rapid network recovery. It can implement an E2E autonomous workflow (perception — analysis — decision — execution) to advance network autonomy from L3 to L5.
5. Computing-network collaborative scheduling: This component can build a computing power measurement system based on geographical locations, resource types, and real-time load status to quantify and detect heterogeneous computing resources in all domains in a unified manner, providing basic support for flexible scheduling of computing power resources. On top of this, it can establish an intelligent computing power scheduling system to dynamically map services to the optimal computing nodes and service paths according to SLA requirements and allocate slices, bandwidth, and other computing and network resources on demand. It can monitor resource utilization and service quality in real time for task-oriented E2E scheduling, improving the overall utilization efficiency of computing network resources and response agility of services.
6. Refined management and control: This component can use digital twins and in-band flow measurement to build a refined management and control system for simulation rehearsal and exception handling. It can build a mirrored network based on digital twins to replicate configurations and traffic characteristics for accurate simulation of scenarios involving configuration changes, route adjustments, and more. It can detect E2E network performance in real time and accurately measure indicators such as the latency, jitter, and packet loss ratios of key service flows on a per-hop basis. This enables rapid anomaly detection and root cause identification, effectively improving network fault warning and troubleshooting capabilities.
7. Proactive security defense: This component can leverage multi-dimensional data collection and AI analysis to identify potential threats and automatically adjust security policies. Specifically, it can continuously monitor and analyze traffic behavior, access frequency, protocol compliance and other relevant data, use AI models to detect threats such as lateral movement, data theft, and unauthorized scanning, and automatically adjust security policies

accordingly to build proactive defense capabilities. For example, it can apply dynamic access control to high–risk source addresses and use encrypted tunnels or trigger path switching for sensitive service flows.