

ccamp
Internet-Draft
Intended status: Informational
Expires: 24 April 2026

Y. Tan, Ed.
Y. Zheng
China Unicom
Q. Hu, Ed.
W. Wang
J. Zhang
Y. Zhao

Beijing University of Posts and Telecommunications
21 October 2025

Unified Optical Networks and AI Computing Orchestration (UONACO) Problem
Statement, Use Cases and Requirements
draft-tan-ccamp-uonaco-problem-statement-00

Abstract

Distributed artificial intelligence (AI) computing is increasingly deployed across geographically dispersed AI data centers (AIDCs) to meet the scale and performance demands of modern AI workloads. In such environments, the efficiency of distributed training, inference, and remote service access depends critically on tight coordination between optical transport networks and compute orchestration systems. However, today's infrastructure operates with isolated control planes: optical networks lack awareness of dynamic compute requirements, while compute schedulers have no visibility into real-time network conditions such as latency, bandwidth, or congestion. This decoupling leads to suboptimal resource utilization, degraded job performance, and inefficient scaling.

This document presents the problem statement, outlines three representative use cases—distributed AI training, distributed AI inference, and remote AI service access—and specifies the requirements for Unified Optical Networks and AI Computing Orchestration (UONACO). The goal is to enable bidirectional awareness, joint resource abstraction, and synchronized control across the compute-optical boundary, thereby supporting intent-driven, end-to-end provisioning of AI services over wide-area optical infrastructures.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <https://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 24 April 2026.

Copyright Notice

Copyright (c) 2025 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<https://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction	3
1.1. Requirements Language	3
2. Problem Statement	3
2.1. Isolated Control and Management	4
2.2. Independent Resource Efficiency Evaluation	5
3. Use Cases	5
3.1. Distributed AI Training	6
3.2. Distributed AI Inference	6
3.3. Accessing Remote AI Service	6
4. Requirements	7
4.1. Integrated Control and Management Architecture	7
4.2. Unified Abstraction of Computing and Network Resources	7
4.3. Joint Orchestration of Computing and Network Resources	7
5. IANA Considerations	8
6. Security Considerations	8
7. References	8
7.1. Normative References	8
Authors' Addresses	8

1. Introduction

The rapid proliferation of large-scale AI applications, particularly those involving distributed training and inference across geographically dispersed AIDCs, has exposed a critical gap in today's infrastructure: the lack of coordination between optical transport networks and compute orchestration systems. While optical networks provide the high-bandwidth, low-latency, and deterministic connectivity required for wide-area AI computing collaboration, their control planes remain largely agnostic to the dynamic, heterogeneous demands of AI workloads. Conversely, compute schedulers operate without visibility into the underlying network's real-time state, such as path latency, available bandwidth, or congestion levels.

This decoupling leads to suboptimal resource utilization, degraded job performance, and inefficient scaling of distributed AI jobs. For instance, a training job may be scheduled across distant AIDCs with abundant GPU capacity but poor optical connectivity, resulting in prolonged synchronization phases and significant compute efficiency loss. Similarly, inference services with strict latency requirements may be routed through paths that meet compute criteria but violate network service-level objectives.

To address these challenges, UONACO enables bidirectional awareness, joint resource abstraction, and synchronized control across the compute-optical boundary. This document describes sample usage scenarios that drive UONACO requirements and will help to identify candidate solution architectures and solutions.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Problem Statement

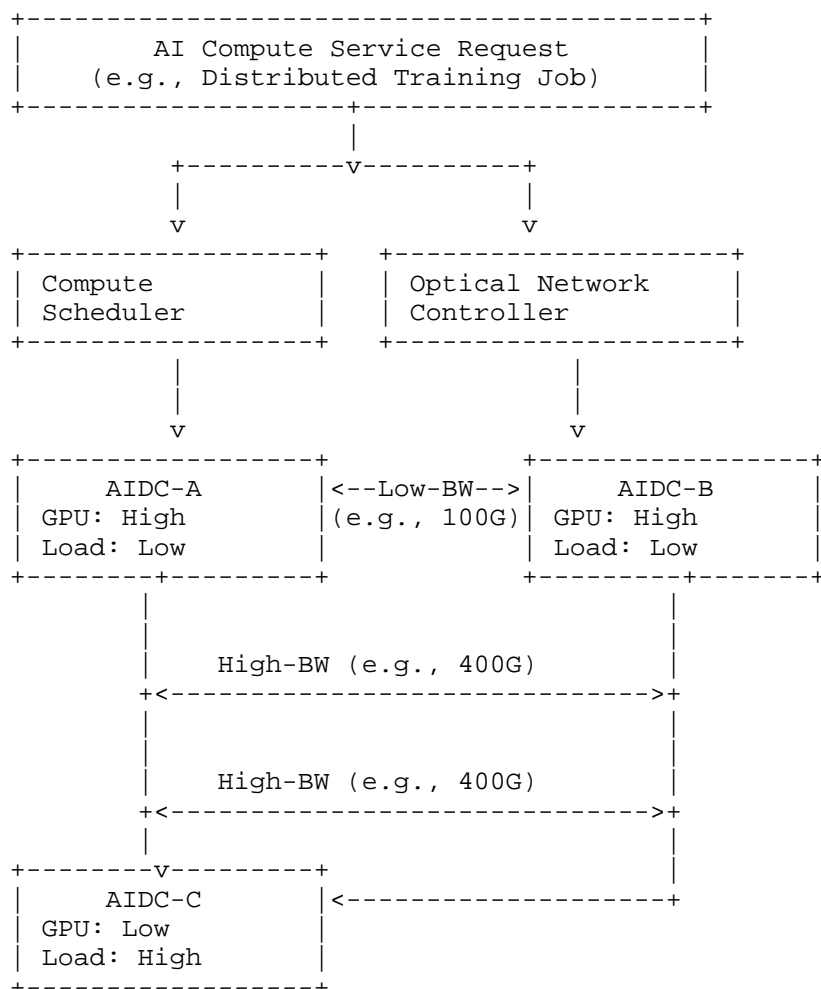


Figure 1: Suboptimal Resource Allocation under Decoupled Control

2.1. Isolated Control and Management

The primary challenge lies in the management and control isolation between the computing domain (e.g., AI training clusters, cloud pools) and the optical transport network. This separation creates a "chasm" that prevents holistic resource optimization.

The optical network control plane operates without awareness of the real-time characteristics and requirements of the compute jobs it carries. It cannot perceive critical parameters such as the

bandwidth intensity of a distributed training job, the strict latency budget of an inference request, or the fluctuating resource demands of a compute job. Consequently, it is unable to proactively establish, adjust, or tear down optical paths (e.g., OTN circuits, OXC switched paths) to optimally serve the compute workload, leading to suboptimal network configurations and underutilized bandwidth.

Conversely, the compute orchestration layer (e.g., Kubernetes schedulers, AI job managers) makes resource allocation and job placement decisions based primarily on local compute and storage metrics (e.g., GPU availability, memory). It lacks visibility into the underlying network's state, including path latency, available bandwidth, or congestion levels between candidate data centers. This results in compute jobs being scheduled across locations with poor network connectivity, causing significant communication bottlenecks and degraded overall job performance (i.e., "compute efficiency loss").

This bidirectional lack of awareness creates a fundamental mismatch, where neither domain can adapt to the needs of the other, severely limiting the potential of wide-area collaborative computing.

2.2. Independent Resource Efficiency Evaluation

Compounding the control isolation is the absence of a unified framework for evaluating the joint efficiency of compute and network resources. Today, network performance is evaluated using traditional metrics like bandwidth utilization, latency, and packet loss. Compute performance is assessed through metrics like FLOPS (Floating Point Operations Per Second), job completion time, and resource utilization (CPU/GPU/memory). These evaluation systems operate in silos. There is no standardized method to quantify the combined cost and benefit of a joint compute-and-network resource allocation decision. For instance, as shown in Figure 1, it is difficult to answer whether allocating a more powerful but distant GPU (with higher network latency) is more efficient than a less powerful but local one for a specific AI job. This lack of a common evaluation language prevents the development of truly optimal co-scheduling algorithms that balance compute power against network quality.

3. Use Cases

The growing scale and distribution of artificial intelligence workloads have created new demands on wide-area optical infrastructure, particularly in scenarios that span multiple artificial intelligence data centers (AIDCs). Three representative use cases illustrate the need for tighter integration between optical transport networks and compute orchestration systems.

3.1. Distributed AI Training

In distributed AI training, large models are increasingly trained across geographically separated AIDCs due to physical and operational constraints within any single site. This requires frequent synchronization of model parameters over long-haul optical links, where performance is highly sensitive to network latency, bandwidth availability, and packet loss. Without awareness of the underlying compute job characteristics, optical networks cannot provision connections that match the dynamic communication patterns of training iterations, leading to significant inefficiencies in GPU utilization and extended training times.

3.2. Distributed AI Inference

Distributed AI inference presents a complementary challenge, where workloads exhibit diverse service-level requirements ranging from strict latency bounds for interactive applications to high-throughput processing for batch analytics. To meet these varied objectives, inference jobs are often deployed across a hierarchical compute fabric that includes cloud, regional, and edge AIDCs. Effective placement of these jobs depends not only on local compute capacity but also on the quality of optical connectivity to the user or upstream service. However, current inference orchestrators lack visibility into optical path conditions such as available bandwidth or propagation delay, while optical networks remain oblivious to the latency and throughput expectations of the inference jobs they carry. This disconnect can result in suboptimal job placement and violations of service-level agreements, highlighting the need for bidirectional signaling between compute and network control planes.

3.3. Accessing Remote AI Service

The third use case, accessing remote AI inference services reflects the growing trend of enterprises and end users consuming AI capabilities on demand via APIs. In this scenario, users request access to remote inference or training resources through high-level APIs, expecting predictable performance and reliability. Traditional best-effort Internet transport is inadequate for such services, which instead require deterministic, isolated, and dynamically adjustable optical channels. Fulfilling these requests demands an integrated control framework capable of jointly evaluating compute availability and optical path feasibility, then orchestrating end-to-end resource allocation across both domains. This scenario underscores the need for standardized protocols that can carry both compute intent and network constraints, enabling automated, intent-driven provisioning of cross-domain AI services.

4. Requirements

4.1. Integrated Control and Management Architecture

A new, integrated control architecture is required to break down the management silos. This architecture must facilitate bidirectional fusion between the compute and network control planes.

It should allow the optical network to dynamically create, adjust, and tear down connections (e.g., leveraging OXC, ODUk, and fgOTN flexible scheduling mechanisms) based on explicit instructions or implicit signals derived from compute job requirements. Conversely, it should enable the compute orchestration system to perform resource planning and online tuning based on the real-time quality and state of the underlying optical network. The goal is to compress the response time for provisioning a compute-and-network service (e.g., "dynamic compute entry") from the current hour-level down to the minute-level.

4.2. Unified Abstraction of Computing and Network Resources

A common language is needed for both domains to understand each other's capabilities and state. This requires a unified resource abstraction model.

The optical network's heterogeneous resources (e.g., at the OTN, optical, and Ethernet layers; or physical ports, spectrum, time slots) must be abstracted into a model that can be understood by the compute orchestrator. This model should expose key attributes like available bandwidth, latency, and reliability.

Similarly, heterogeneous compute resources (CPU, GPU, memory, storage) must be abstracted into a model that conveys their real-time capabilities (e.g., FLOPS, available memory/VRAM) and state to the network controller. This unified abstraction is the foundation for any joint decision-making process.

4.3. Joint Orchestration of Computing and Network Resources

Building on the integrated architecture and unified abstraction, a joint orchestration mechanism is required to make intelligent, end-to-end resource allocation decisions. This mechanism should include compute-aware optical network dynamic adjustment algorithms that can perform application-driven optical path reconfiguration and adaptive bandwidth allocation. It should also include network-aware compute joint orchestration algorithms that can perform dynamic compute resource optimization and contention-aware job scheduling.

The orchestration should be driven by a unified evaluation system that can assess the joint efficiency of a scheduling decision by considering both compute performance metrics (e.g., job completion time) and network performance metrics (e.g., bandwidth cost, latency) in a single, cohesive framework.

5. IANA Considerations

TBD

6. Security Considerations

TBD

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

Authors' Addresses

Yanxia Tan (editor)
China Unicom
Email: tanyx11@chinaunicom.cn

Yanlei Zheng
China Unicom
Email: zhengyanlei@chinaunicom.cn

Qiaojun Hu (editor)
Beijing University of Posts and Telecommunications
Email: qiaoj475@bupt.edu.cn

Wei Wang
Beijing University of Posts and Telecommunications
Email: weiw@bupt.edu.cn

Jie Zhang
Beijing University of Posts and Telecommunications
Email: jie.zhang@bupt.edu.cn

Yongli Zhao
Beijing University of Posts and Telecommunications
Email: yonglizhao@bupt.edu.cn