

Network Working Group
Internet-Draft
Intended status: Informational
Expires: November 30, 2026

M. Tayyebi
ViYouna
May 29, 2026

A Fabric Coordination Layer (FCL) for High-Scale AI Training:
Distributed Credit-Orbit Pacing (DCOP) and Global Token Ledger (GTL)
draft-tayyebi-fcl-dcop-gtl-00

Abstract

This document specifies a Fabric Coordination Layer (FCL) designed to stabilize data transmission in frontier-scale distributed computing fabrics. As AI training clusters scale to hundreds of thousands of accelerators and link speeds exceed 800 Gbps, traditional reactive congestion control mechanisms suffer from severe feedback-loop latency. By allocating transmission authority via a Global Token Ledger (GTL) and enforcing it through Distributed Credit-Orbit Pacing (DCOP), the FCL mitigates incast-driven buffer overflows and significantly reduces tail-latency variance, thereby maximizing Model Flops Utilization (MFU).

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 November 30, 2026.

Copyright Notice

Copyright (c) 2026 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 New BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the New BSD License.

1. Introduction

Modern Artificial Intelligence (AI) training workloads rely on collective communication patterns (e.g., All-Reduce, All-to-All) that generate massive, synchronized data bursts. Traditional congestion control mechanisms (e.g., ECN, PFC) are fundamentally reactive; they respond only after congestion is detected.

As network speeds scale to 1.6 Tbps and beyond, the physics of data transport dictate that the "Time to Overflow" (T_{overflow}) for a

switch buffer is often shorter than the "Time to Control" (T_{control}) required for a pause frame or congestion notification to traverse the fabric. If $T_{\text{overflow}} \leq T_{\text{control}}$, the fabric must regulate injection prior to transmission.

This document introduces the Fabric Coordination Layer (FCL), comprising a Global Token Ledger (GTL) and Distributed Credit-Orbit Pacing (DCOP), which shifts the paradigm from reactive backpressure to predictive, authority-based coordination.

2. Terminology

- o Fabric Coordination Layer (FCL): A control plane that converts workload-level communication intent into enforceable transmission authority.
- o Global Token Ledger (GTL): A logically shared representation of transmission authority across nodes, enabling dark capacity harvesting.
- o Distributed Credit-Orbit Pacing (DCOP): A hardware-native control mechanism that regulates transmission timing within the NIC/DPU using authority, queue state, and predicted demand.
- o Deterministic Burst Volume (DBV): The predicted data volume for an upcoming communication epoch or collective operation.
- o Incast Collapse: A condition where highly correlated, synchronized traffic from multiple senders overwhelms a single receiver's port buffer.

3. The Fabric Coordination Layer Architecture

The FCL bridges upper-layer application intent with lower-layer physical constraints. It enforces a "Layered Authority Envelope" to ensure global conservation of bandwidth.

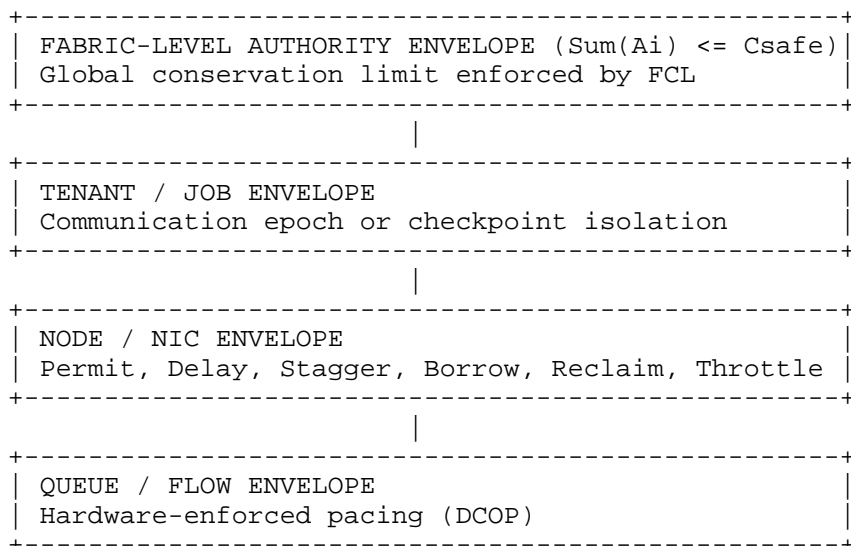


Figure 1: Layered Authority Envelope

The FCL mathematically guarantees that the sum of all active transmission authorities (A_i) never exceeds the safe physical absorption capacity (C_{safe}) of the fabric spine.

4. Global Token Ledger (GTL) Specification

The GTL prevents incast by globally managing transmission tokens.

Unlike traditional fair-share schedulers, GTL enables asymmetric reallocation. During a collective All-Reduce, a subset of nodes may be compute-bound while others are communication-bound. The GTL identifies idle nodes and temporarily aggregates their unused transmission capacity, lending it to active nodes to accelerate epoch completion without exceeding Csafe.

5. Distributed Credit-Orbit Pacing (DCOP) Specification

DCOP serves as the enforcement arm of the FCL, operating within the physical hardware (SmartNIC or DPU) to achieve nanosecond-scale precision.

When a node's local transmission demand exceeds its local capacity, DCOP queries the GTL. If capacity is available, DCOP triggers a single-clock cycle Read-Modify-Write (RMW) operation within the NIC's SRAM. This executes an atomic swap of credits from the GTL Peer Map to the local Transmit Pipeline, allowing burst rates that safely exceed nominal link rates without relying on reactive drop signals.

6. Security Considerations

Because transmission authority represents physical fabric bandwidth, the GTL must be secured against token spoofing and unauthorized harvesting. Implementations **MUST** utilize hardware-rooted cryptographic signatures for inter-node authority transfers to prevent malicious tenants from executing denial-of-service (DoS) attacks via bandwidth starvation.

7. IANA Considerations

This memo includes no request to IANA.

8. Informative References

[ViYouna-WP] ViYouna, "The End of Reactive Networking: Why multi-terabit AI clusters require coordinated transmission", 2026.