

Requirements and Gap Analysis of Multicast in AI Data Centers

draft-zhang-rtgwg-multicast-requirements-gaps-aidc

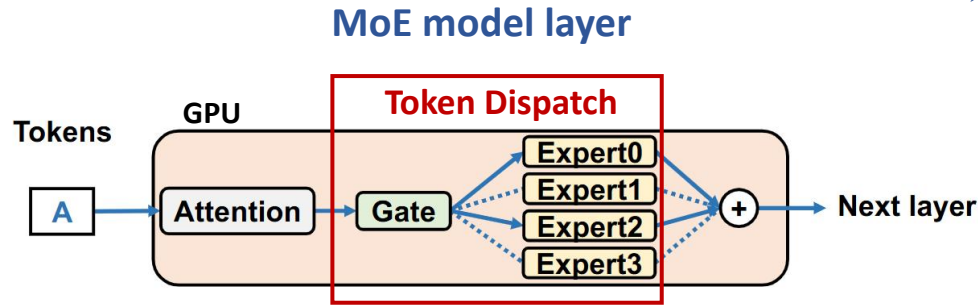
Junye Zhang, China Mobile

Weiqiang Cheng, China Mobile

Kefei Liu, China Mobile

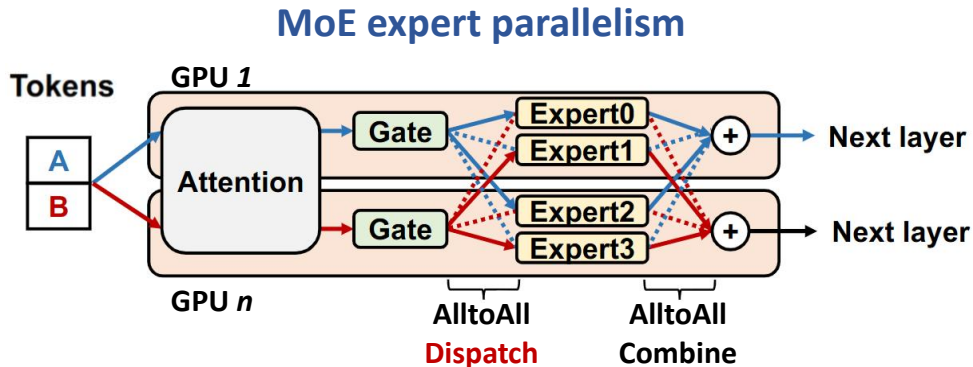
IETF 125, Shenzhen

Multicast Use Cases in AIDCs: MoE Token Dispatch



➤ Mixture-of-Experts (MoE) model

- **Popularity:** A popular model architecture for LLM training and inference, adopted by many LLMs such as Llama4, Mixtral and DeepSeekV3.
- **Model structure:** Traditional dense model layers are replaced by MoE layers with **sparse, dynamically-activated experts** to enhance computing efficiency.
- **Process:** A token is dispatched to multiple selected experts decided by a gating network. Each token may be routed to different experts.



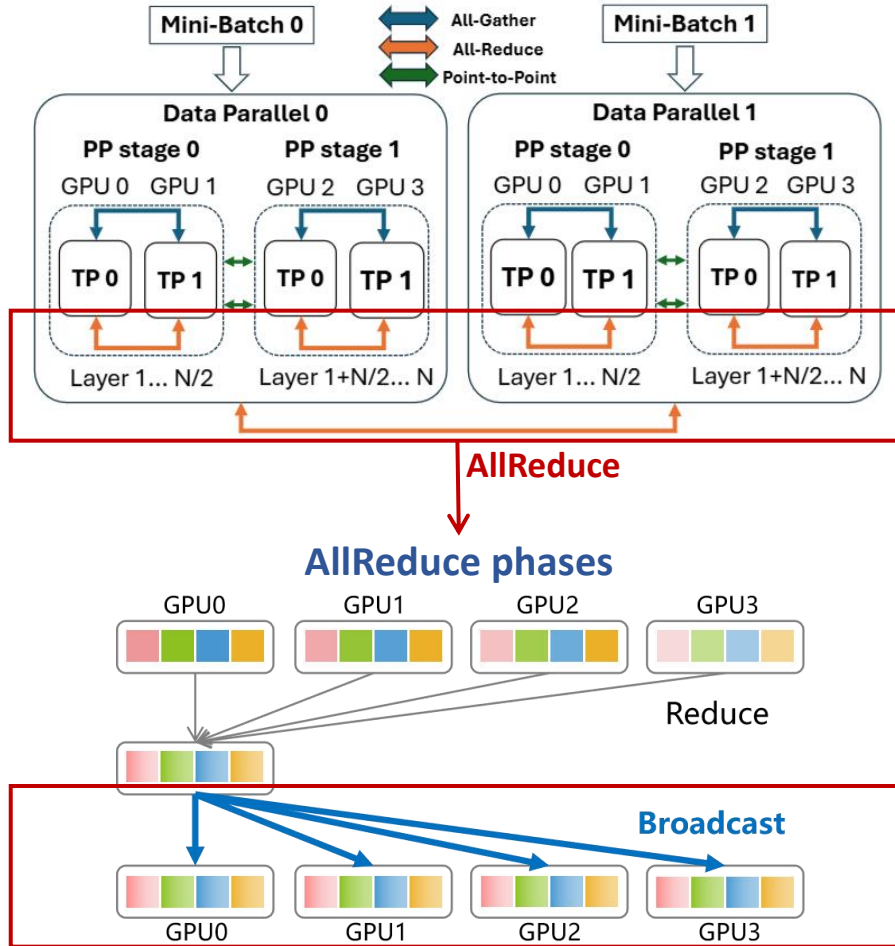
➤ Expert Parallelism (EP)

- **Motivation:** A parallelism strategy for MoE that distributes experts across GPUs connected by networks to handle large parameter scales of LLMs.
- **Communication process:** One forward pass requires **two AlltoAll communications via networks**, i.e., **Dispatch and Combine**. Tokens are sent to selected experts in the Dispatch phase, and the processed tokens are sent back in the Combine phase.
- **P2MP traffic:** The AlltoAll token dispatch follows the **Point-to-Multipoint (P2MP)** traffic pattern, which is a typical multicast use case.

Multicast Use Cases in AIDCs: AllReduce in LLM Training

LLM distributed training

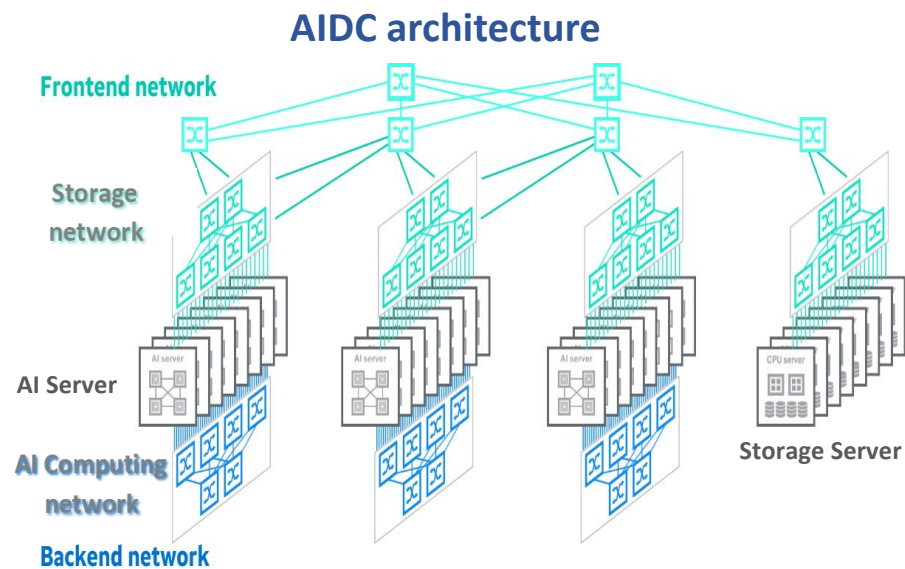
Tensor Parallel (TP) Pipeline Parallel (PP) Data Parallel (DP)



➤ AllReduce

- **Importanc:** A core collective communication primitive fundamental to distributed LLM training, especially for data parallel and tensor parallel. It synchronizes model gradients/parameters across computing devices (e.g., GPUs) to ensure training consistency.
- **Phases:** Can be decomposed into **Reduce and Broadcast phases**.
 - Reduce phase: aggregate gradient/parameter shards across devices.
 - Broadcast phase: distribute the consolidated results to participating devices.
- **P2MP traffic:** The Broadcast phase exhibits a P2MP traffic pattern, natively matching the multicast semantics.

Multicast Use Cases in AIDCs: Model Distribution & Checkpoint Saving

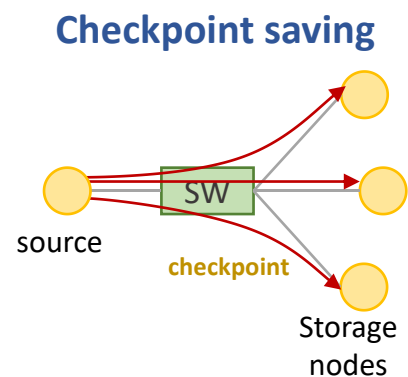
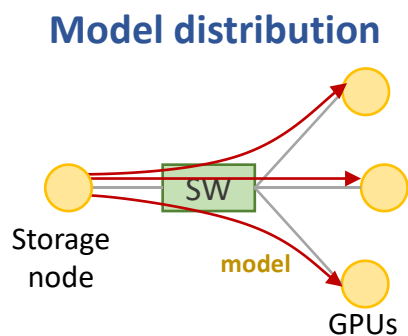


➤ Model distribution

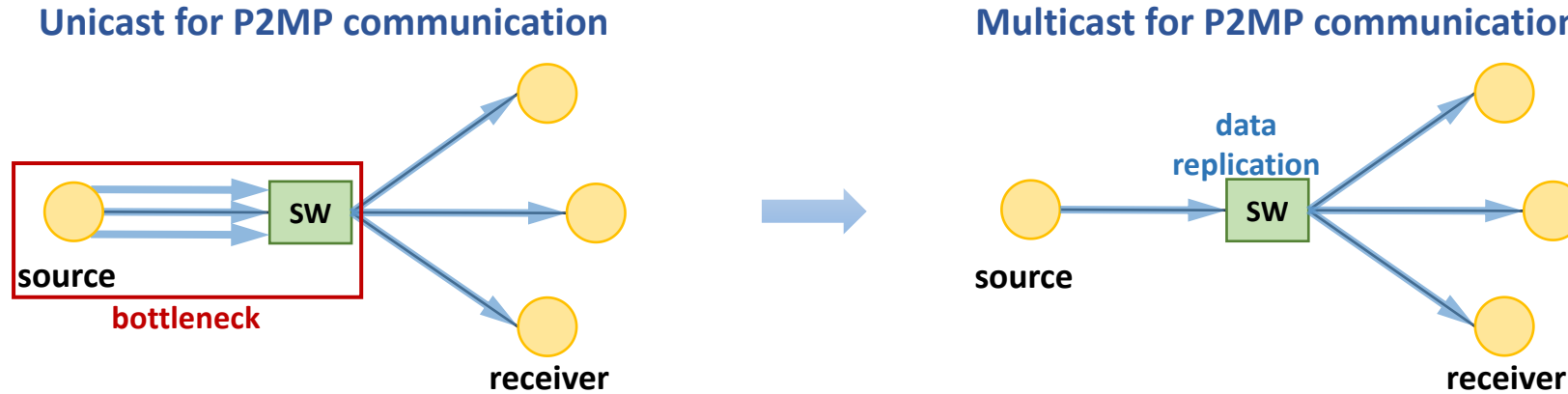
- **Scenario:** Initial distribution of model parameters or checkpoints from a storage node to computing nodes at job startup. Critical for distributed training before iterations begin where model data can reach terabytes.
- **P2MP Traffic:** The same massive data should be sent simultaneously from a storage node to GPUs, which is a typical P2MP multicast pattern.

➤ Checkpoint saving

- **Scenario:** During distributed training, checkpoints (including model parameters, model weights, optimizer states, ...) are periodically saved to persistent storage for fault tolerance. Checkpoints are typically written to multiple storage targets to ensure reliability.
- **P2MP traffic:** The multi-copy saving to multiple storage nodes will generate P2MP traffic, which is a typical multicast use case.



Benefits of using network multicast in AIDCs



AIDCs exhibit numerous P2MP traffic. Benefits of **using network multicast to offload data replication from servers to network devices** to implement P2MP communication in AIDCs:

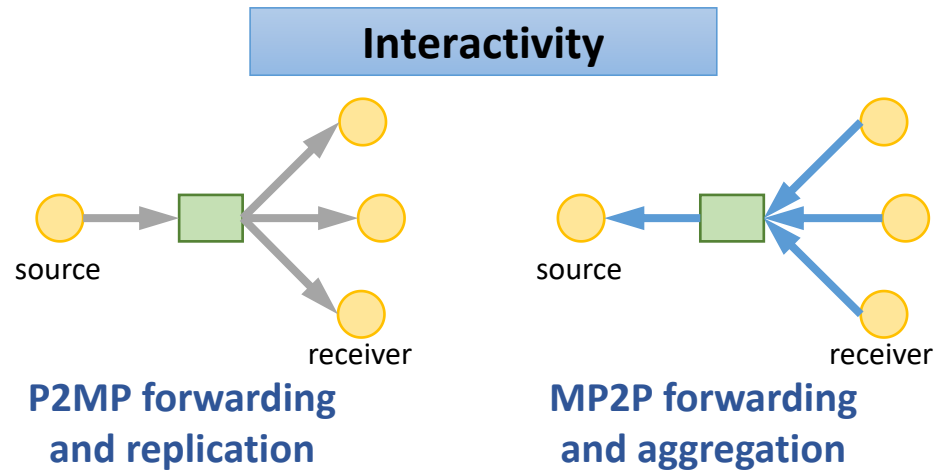
➤ Processing & Latency

- ✓ Reduce processing overhead and latency of source GPUs/NICs
- ✓ Enhance scalability by reducing source-side overhead

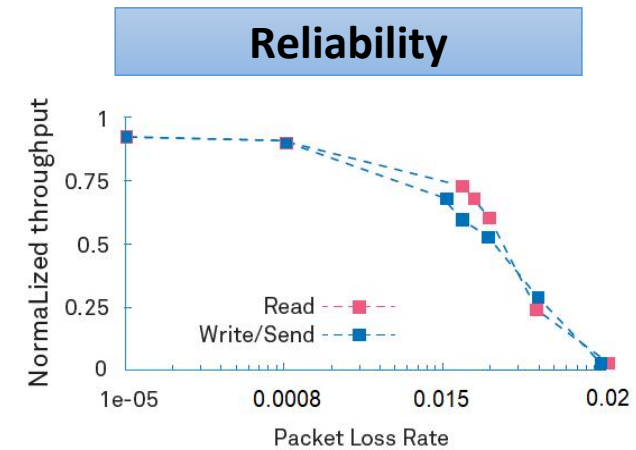
➤ Transmission & Bandwidth

- ✓ Eliminate redundant transmission from the source
- ✓ Improve receiver-side throughput by easing bandwidth pressure at the source
- ✓ Reduce network traffic, mitigating network congestion risks

Requirements: Interactivity & Reliability



- **P2MP forwarding and replication:** the fundamental functions of multicast
- **MP2P forwarding:** natively support efficient Multipoint-to-point (MP2P) forwarding, particularly for feedbacks such as ACKs and congestion signals
 - No ACKs → Unaware of packet loss, directly affecting parameter integrity and training accuracy
 - No congestion signals → Lead to network congestion and packet loss
- **MP2P packet aggregation:** support MP2P packet aggregation to avoid feedback storm with numerous receivers



RDMA Throughput degrades to 0 with only 2% loss rate

- **Network Lossless:** the high-reliability requirement of AIDC networks (also need support of interactivity)
- **Fast failure detection:** fast detection of link failures and node failures and efficient detection of gray failures
- **Fast failure recovery:** fast failover without service interruption
- **Minimized failure domain:** confine the impact of a failure to the smallest possible set of parts (the faulty segment instead of the entire multicast tree)

Requirements: Dynamics, Sparseness & Simplicity

Dynamics

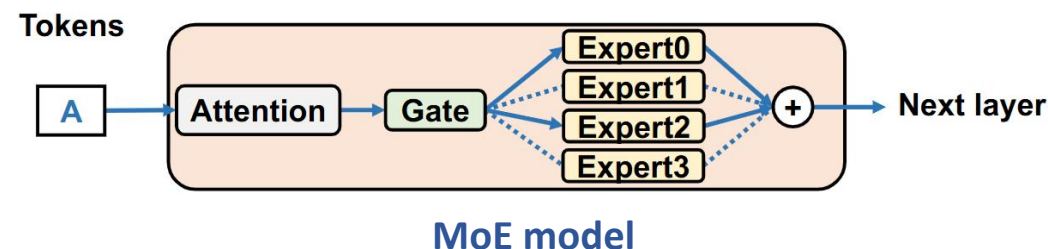
- **Microsecond-level adaptation:** adapt to dynamic member set changes in microseconds in scenarios like MoE token dispatch
- **Real-time path adaptation:** enable instant forwarding based on real-time multicast member selection without the need for pre-defined fixed groups
- **Ultra-low adjustment overhead:** minimize control and data plane overhead during dynamic changes to ensure performance and prevent packet loss and task failure

Sparseness

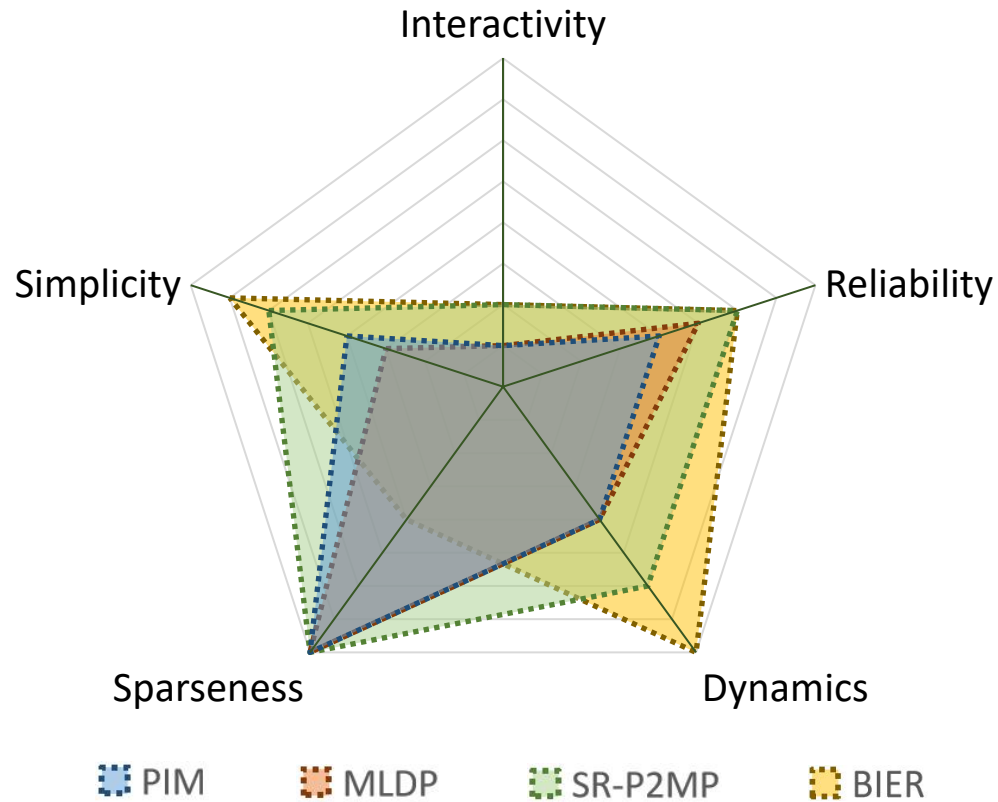
- **Efficient sparse member identification:** be efficient for forwarding, and avoid unnecessary scanning or signaling of non-member nodes in sparse-member scenarios (e.g., active 9/256 experts in DeepSeekV3)
- **Lightweight state maintenance:** avoid maintaining redundant states for non-member nodes to reduce the memory and processing burden

Simplicity

- **Simple control plane:** essential features, minimizing signaling and state synchronization to reduce operational complexity and latency
- **Efficient data plane:** efficient member identification and optimized packet to ensure low processing latency



Gap analysis



- **Interactivity:** No native support for efficient backward MP2P forwarding or packet aggregation, hindering closed-loop feedback (e.g., ACKs, congestion signals).
- **Reliability:** Fail to meet the lossless requirement of AIDC networks (partly limited by lack of interactivity). Tree-based techniques suffer from large failure domains.
- **Dynamics:** PIM/MLDP convergence is too slow for high-frequency member changes; SR-P2MP requires global recalculation; BIER is relatively good.
- **Sparseness:** BIER's BitString overhead degrades forwarding efficiency in sparse member scenarios.
- **Simplicity:** PIM/MLDP are highly complex; SR-P2MP is moderate; BIER is more simple but still needs further optimization for AIDC's ultra-high performance.

None of the existing multicast techniques can satisfy all the multicast requirements in AIDCs. A new architecture or deep enhancement to current multicast techniques are urgently needed for AIDCs.

Next Steps

- Seek feedback from the working group
- Welcome discussion and collaboration

Thanks