Congestion Control for AI Workloads with Message-Level Signaling

Yuxuan Li, Zhenghang Ren, Wenxue Li, Xiangzhou Liu, Kai Chen

iSING Lab @ HKSUT

August 2025

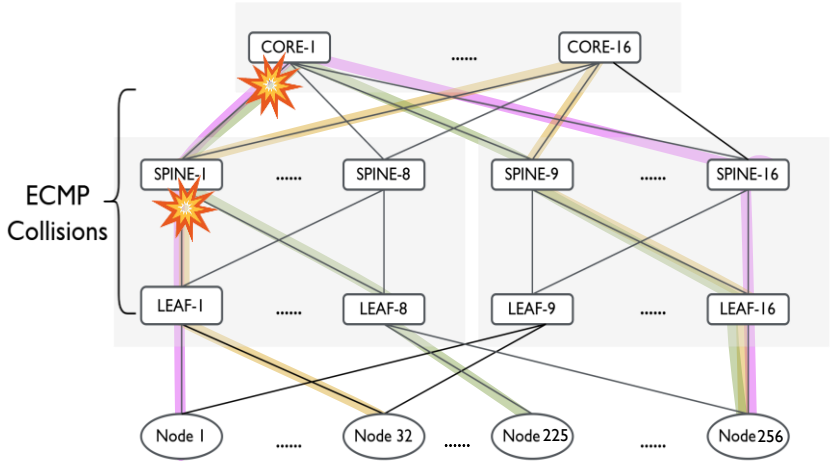
Challenges of Per-flow Load Balancing in AI Training Workloads

Collectives	Number of Flows per GPU [1]	Flow Size [2] 175B GPT3 model
AlltoAll	15	6MB
AllReduce	4	5.5GB
AllGather	4	560MB
Reduce Scatter	4	560MB

Low entropy

Large flow

AI flow characteristics in production



❶ **Low entropy large flows,** and **highly bursty** traffic patterns increase the likelihood of hash collisions.

❷ **Co-flow dependencies** exacerbate consequences when collisions occur.

[1] RDMA over Ethernet for Distributed AI Training at Meta Scale, SIGCOMM 2024
[2] Alibaba HPN: A Data Center Network for Large Language Model Training, SIGCOMM 2024

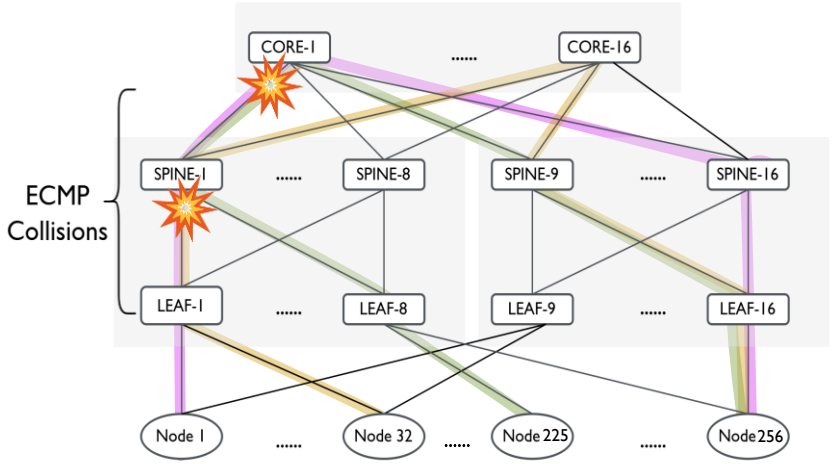
Challenges of Per-flow Load Balancing in AI Training Workloads

Collectives	Number of Flows per GPU [1]	Flow Size [2] 175B GPT3 model
AlltoAll	15	6MB
AllReduce	4	5.5GB
AllGather	4	560MB
Reduce Scatter	4	560MB

Low entropy

Large flow

AI flow characteristics in production



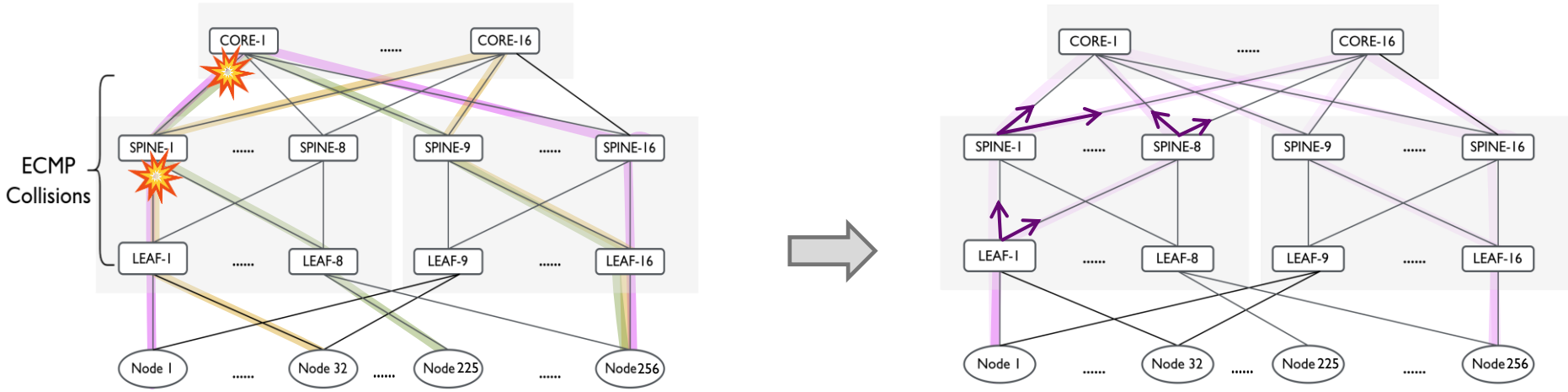
❶ **Low entropy large flows,** and **highly bursty** traffic patterns increase the likelihood of hash collisions.

❷ **Co-flow dependencies** exacerbate consequences when collisions occur.

Per-flow load balancing (i.e., ECMP) can significantly degrade AI communication throughput!

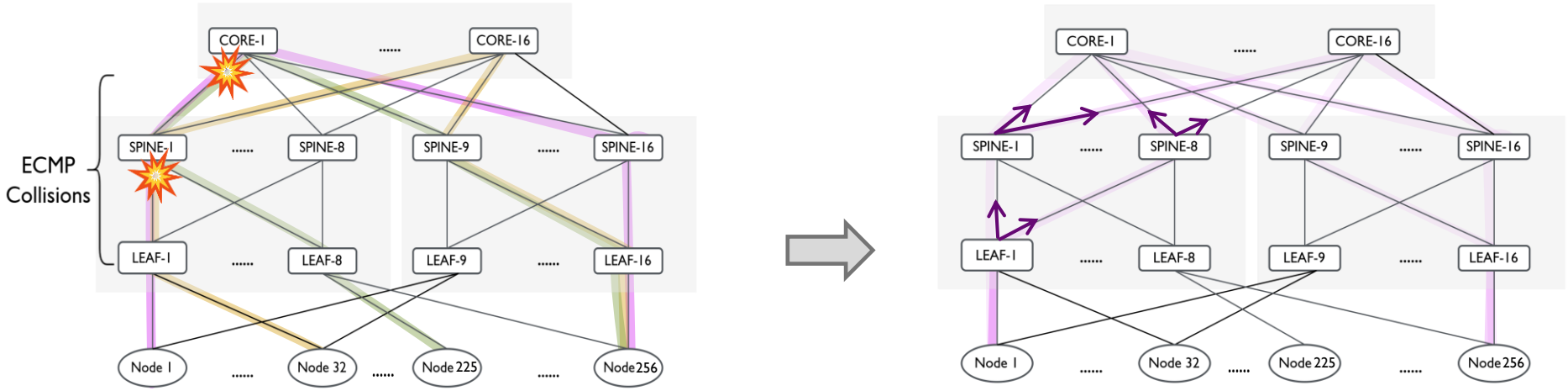
[1] RDMA over Ethernet for Distributed AI Training at Meta Scale, SIGCOMM 2024
[2] Alibaba HPN: A Data Center Network for Large Language Model Training, SIGCOMM 2024

AI Training Workloads Necessitate Packet Spray



1. Packet spray performs load balancing at packet granularity.
2. Adaptive routing (AR) can sense local congestion and dynamically re-route packets to less congested paths.

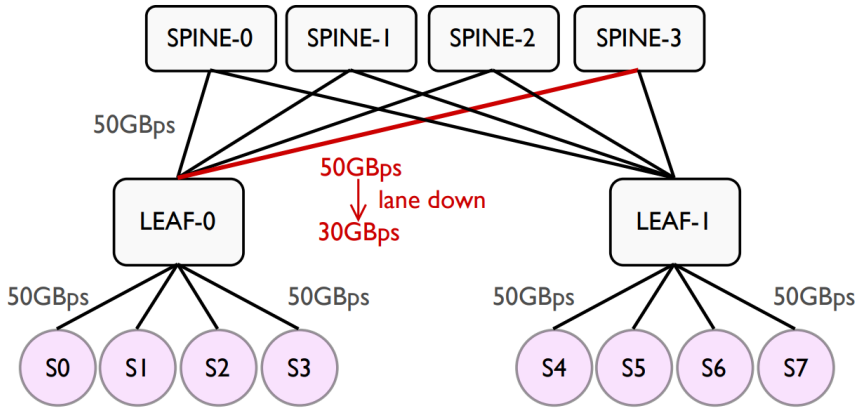
AI Training Workloads Necessitate Packet Spray



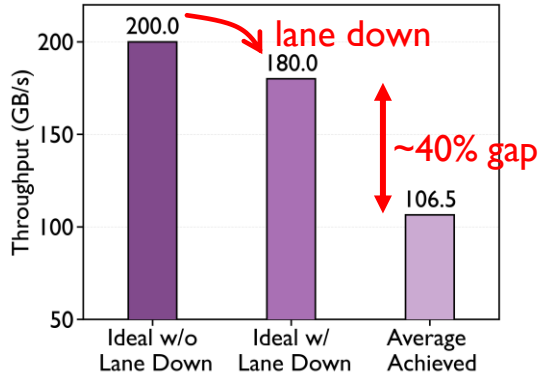
1. Packet spray performs load balancing at packet granularity.
2. Adaptive routing (AR) can sense local congestion and dynamically re-route packets to less congested paths.

Packet spray avoids hash collisions and can help absorb transient load imbalances.

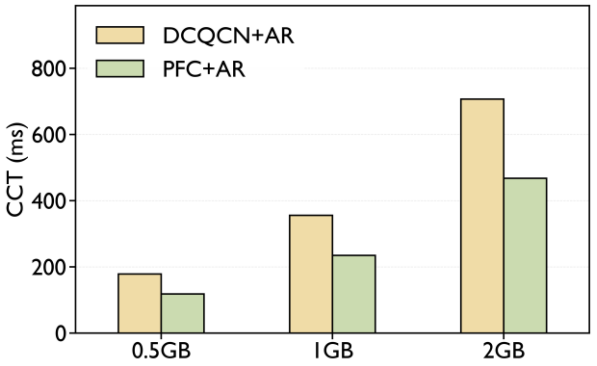
However, Packet Spray does not warrant high throughput (a toy example) !



- Simulation Results:



With packet spray (AR), throughput remains below 60%.



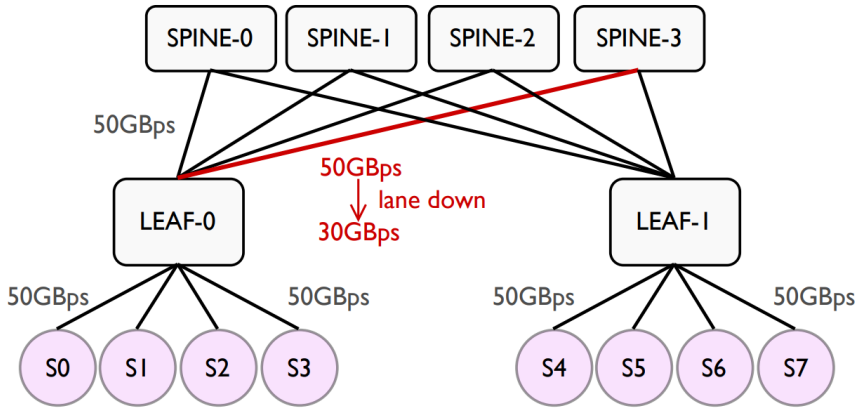
Turning off CC (i.e., AR with PFC) performs even better.

- Settings:

- 2-layer spine-leaf topology, 8 servers into 4 groups (2 servers each).
- Each group simultaneously executes an AllReduce operation with sizes of 0.5 GB, 1 GB, and 2 GB.
- DCQCN with recommended parameters in HPCC [1].

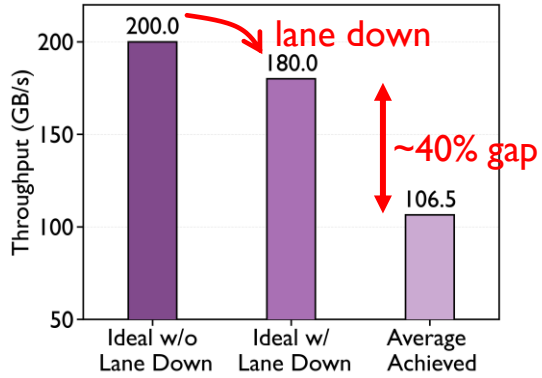
[1] HPCC: High Precision Congestion Control, SIGCOMM 2019

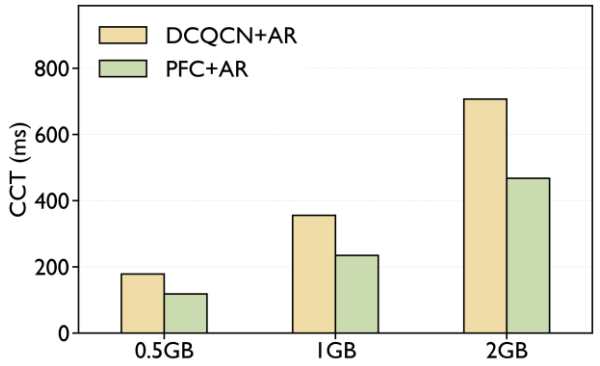
However, Packet Spray does not warrant high throughput (a toy example) !



- Settings:
  - 2-layer spine-leaf topology, 8 servers into 4 groups (2 servers each).
  - Each group simultaneously executes an AllReduce operation with sizes of 0.5 GB, 1 GB, and 2 GB.
  - DCQCN with recommended parameters in HPCC [1].

- Simulation Results:





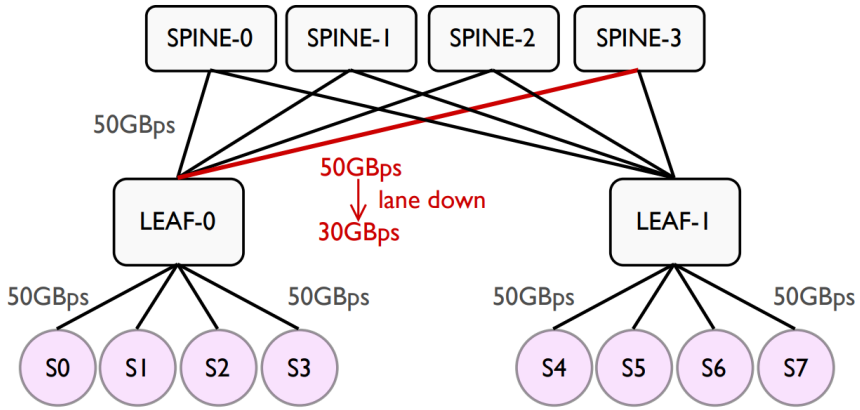
With packet spray (AR), throughput remains below 60%.

Turning off CC (i.e., AR with PFC) performs even better.

Traditional CCs are incompatible with packet spray!

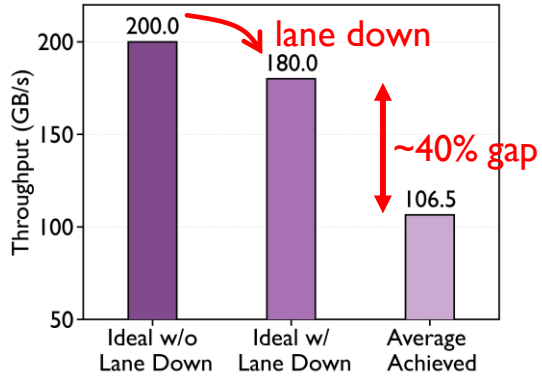
[1] HPCC: High Precision Congestion Control, SIGCOMM 2019

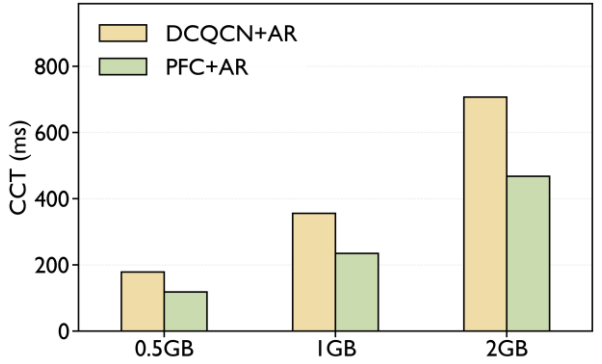
However, Packet Spray does not warrant high throughput (a toy example) !



- Settings:
  - 2-layer spine-leaf topology, 8 servers into 4 groups (2 servers each).
  - Each group simultaneously executes an AllReduce operation with sizes of 0.5 GB, 1 GB, and 2 GB.
  - DCQCN with recommended parameters in HPCC [1].

- Simulation Results:

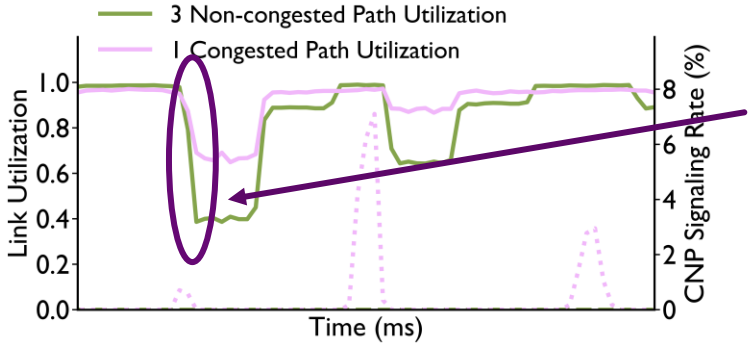




With packet spray (AR), throughput remains below 60%.

Turning off CC (i.e., AR with PFC) performs even better.

Traditional CCs are incompatible with packet spray!

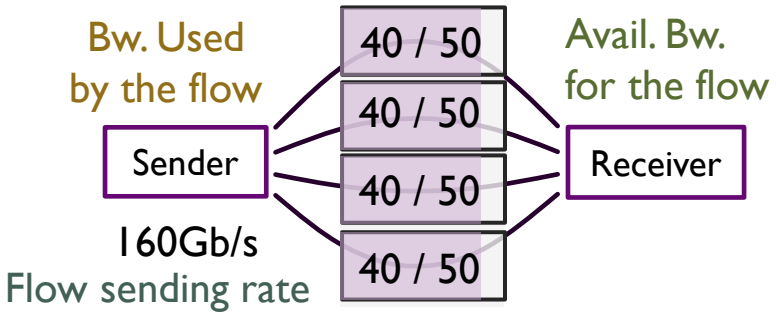


We observed that 1% CNP rate from 1 path causes up to 60% utilization drop on all other 3 paths!

[1] HPCC: High Precision Congestion Control, SIGCOMM 2019

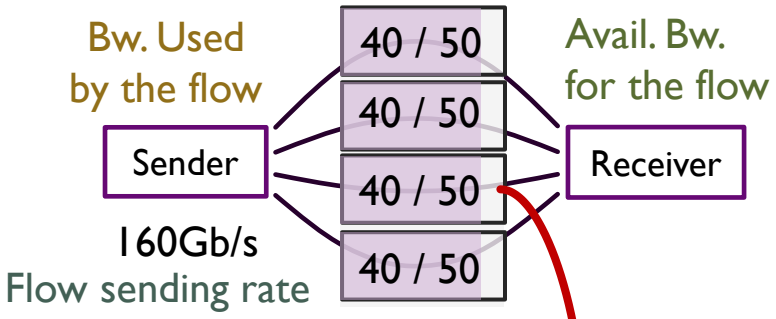
Root Cause: Overreaction of Existing CC

I Non-Congestion

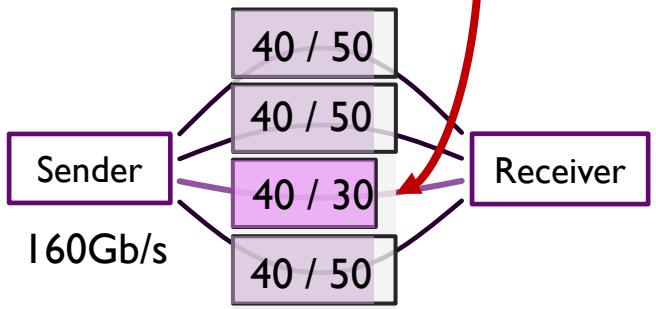


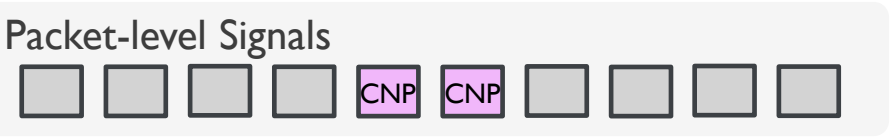
Root Cause: Overreaction of Existing CC

1 Non-Congestion



2 Partial Congestion

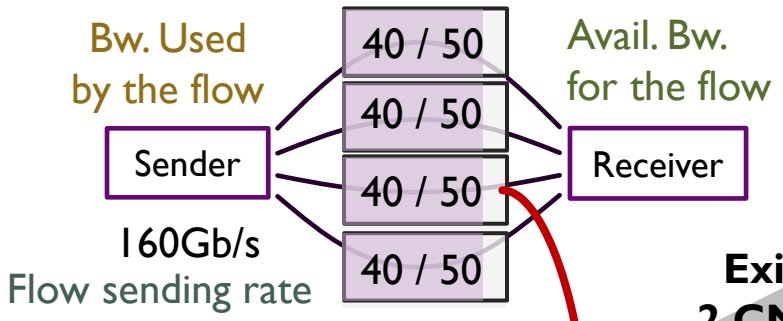




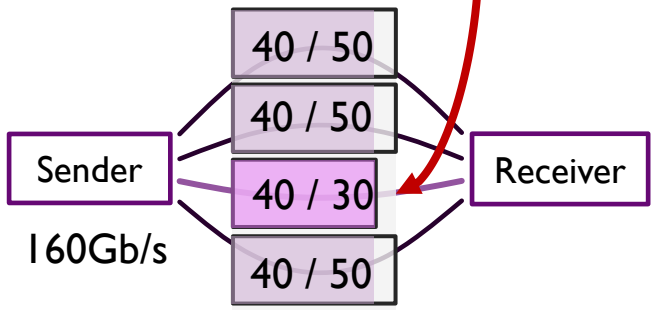
2 CNPs from 1 congested path

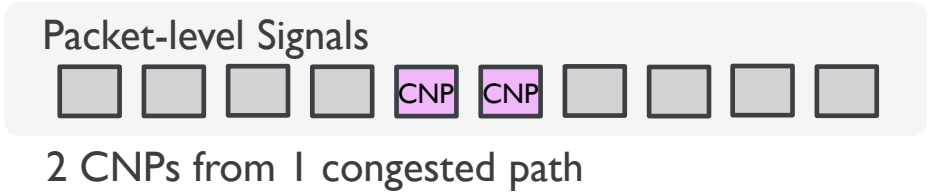
Root Cause: Overreaction of Existing CC

1 Non-Congestion



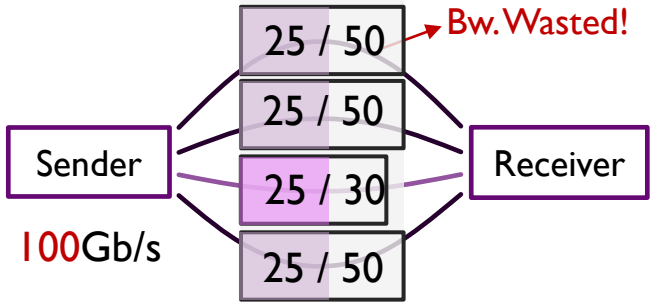
2 Partial Congestion





Existing CC:
2 CNPs trigger
2 slowdown adjustments

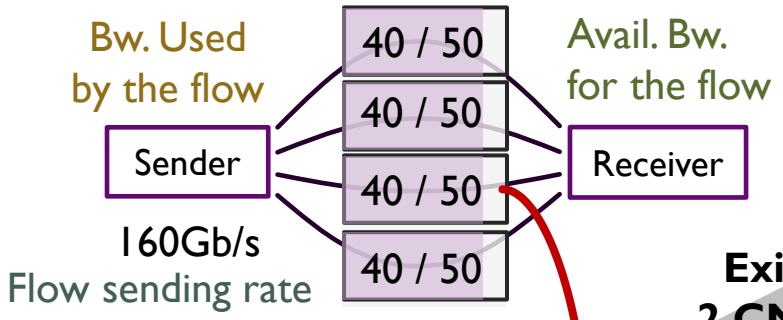
3 Overreaction



- Traditional packet-level CCs are incompatible with packet spray because they overreact to partial congestion without considering adaptive packet spray capability over the pathset!

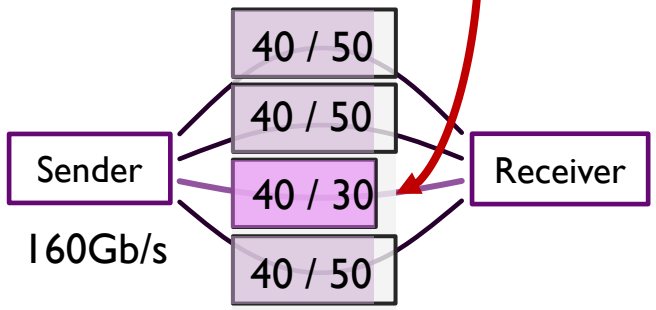
Root Cause: Overreaction of Existing CC

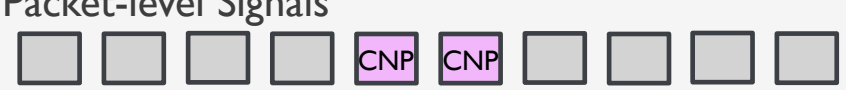
1 Non-Congestion



Existing CC:
2 CNPs trigger
2 slowdown adjustments

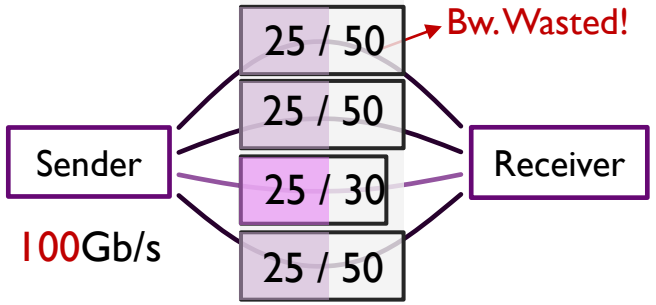
2 Partial Congestion





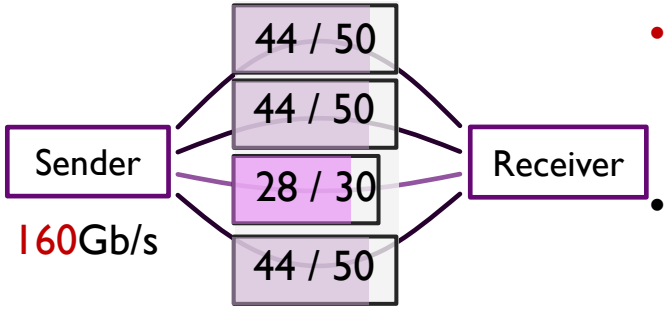
2 CNPs from 1 congested path

3 Overreaction



- Traditional packet-level CCs are incompatible with packet spray because they overreact to partial congestion without considering adaptive packet spray capability over the pathset!

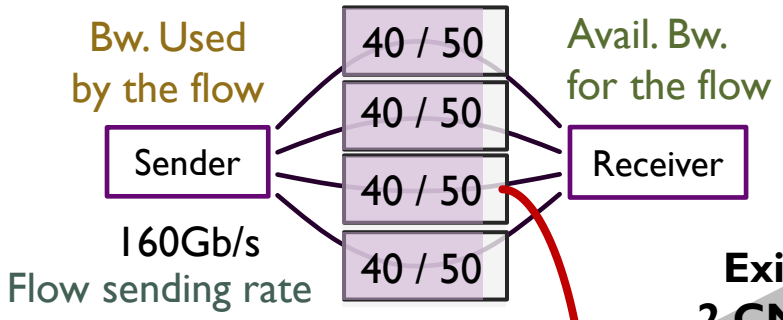
4 Right Reaction: Keep Rate



- Inherent re-pathing ability in AR can absorb partial congestion, well adapting to internal imbalances.
- The sender maintains full sending rate without causing congestion.

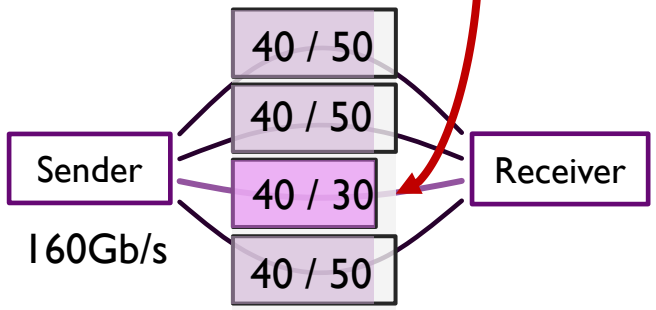
Root Cause: Overreaction of Existing CC

1 Non-Congestion



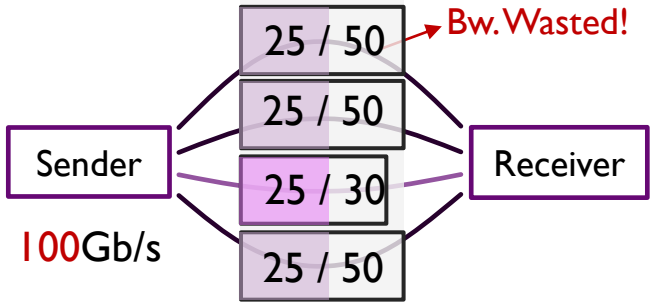
Existing CC:
2 CNPs trigger
2 slowdown adjustments

2 Partial Congestion



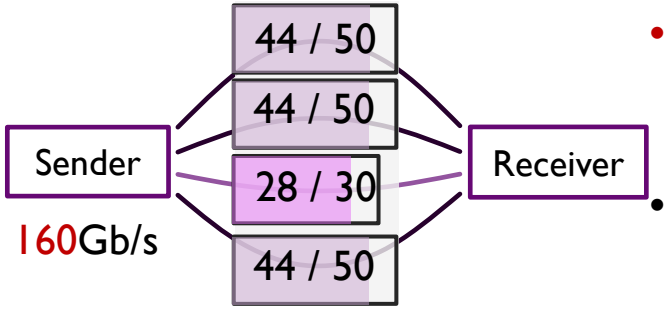
2 CNPs from 1 congested path

3 Overreaction



- Traditional packet-level CCs are incompatible with packet spray because they overreact to partial congestion without considering adaptive packet spray capability over the pathset!

4 Right Reaction: Keep Rate

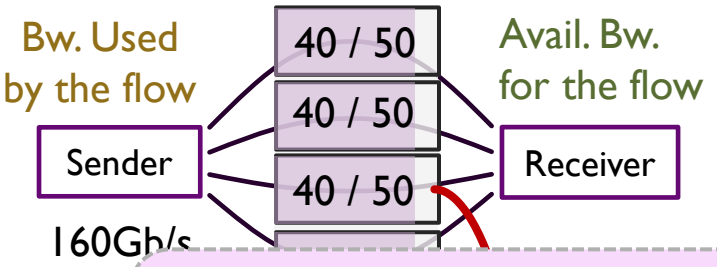


- Inherent re-pathing ability in AR can absorb partial congestion, well adapting to internal imbalances.
- The sender maintains full sending rate without causing congestion.

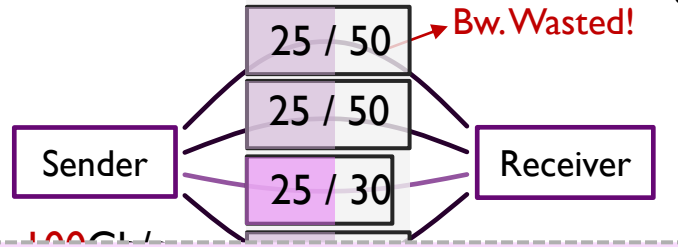
Existing CCs cannot distinguish localized partial congestion which AR can resolve on its own.

Root Cause: Overreaction of Existing CC

1 Non-Congestion



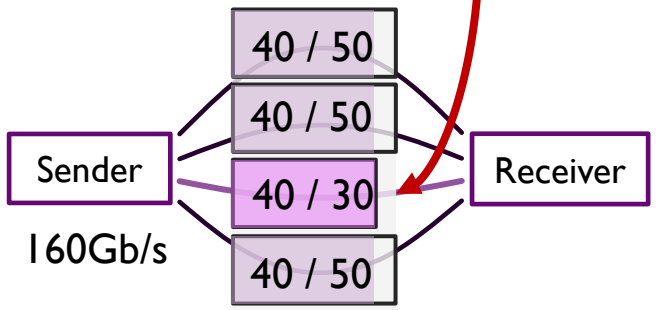
3 Overreaction

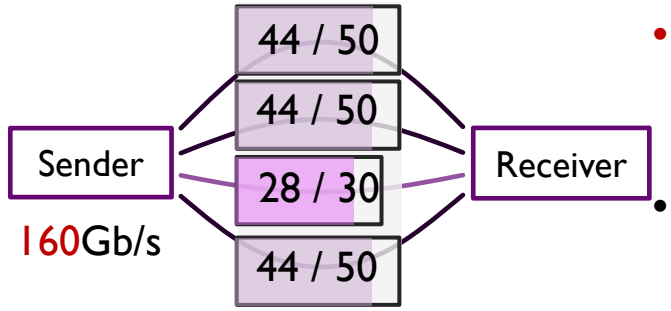


- Traditional packet-level CCs are incompatible with packet spray because they overreact to partial congestion without considering adaptive packet spray capability
  pathset!

How can we make a CC algorithm smart enough to distinguish between a local imbalance that AR can handle, and a true, global congestion that CC must handle?

2 Partial





- Inherent re-pathing ability in AR can absorb partial congestion, well adapting to internal imbalances.
- The sender maintains full sending rate without causing congestion.

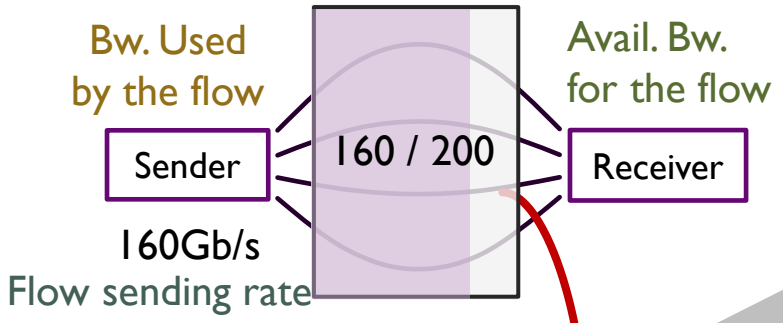
Packet-level Signals

2 CNPs from 1 congested path

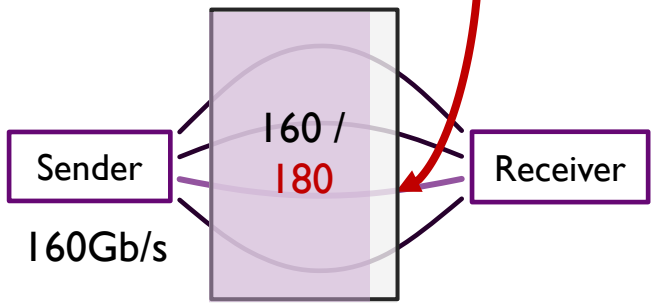
Existing CCs cannot distinguish localized partial congestion which AR can resolve on its own.

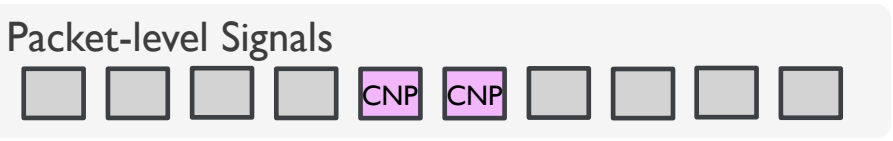
Network under AR Works like a Network Pipe (Pathset)

1 Non-Congestion



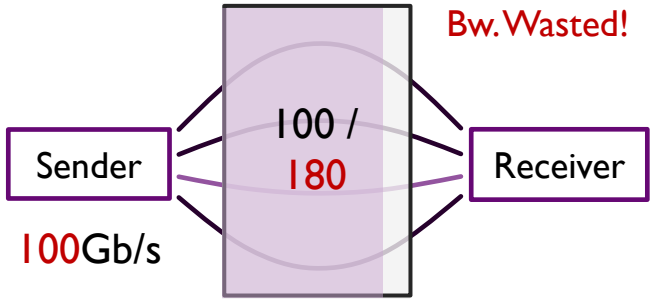
2 Partial Congestion





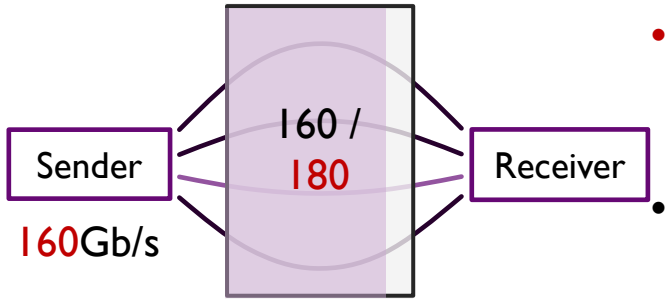
Existing CC:
2 CNPs trigger
2 slowdown adjustments

3 Overreaction



- Traditional packet-level CCs are incompatible with packet spray because they overreact to partial congestion without considering adaptive packet spray capability over the path set!

4 Right Reaction: Keep Rate

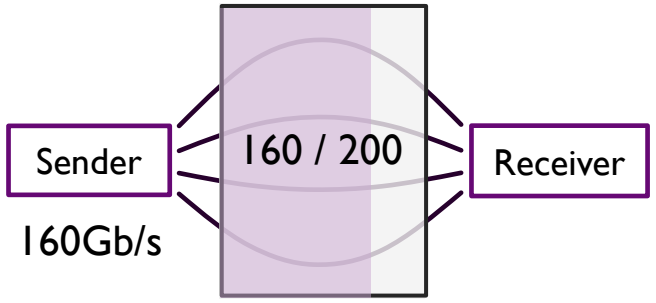


- Inherent re-pathing ability in AR can absorb partial congestion, well adapting to internal imbalances.
- The sender maintains full sending rate without causing congestion.

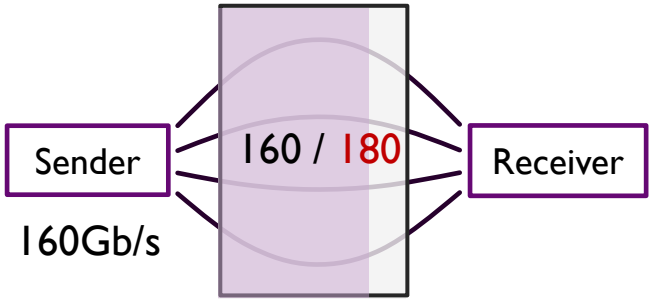
Network under AR Works like a Network Pipe (Pathset)

- As AR can best effort balancing all packets to all available paths, the network can be considered as a high-bandwidth **network pipe**.
- Traditional per-packet signal can only capture congestion in **individual path**.

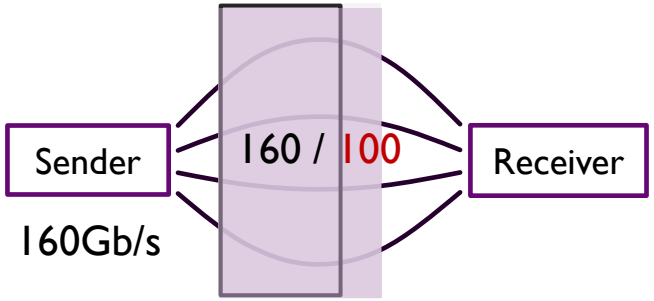
1 Non-Congestion



2 Partial Congestion



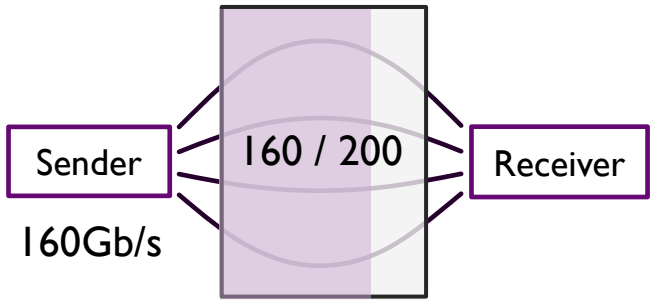
3 Global Congestion



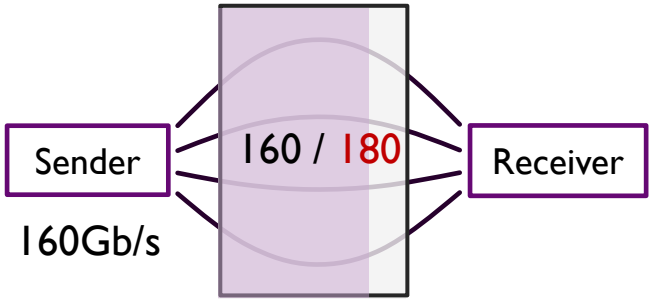
Network under AR Works like a Network Pipe (Pathset)

- As AR can best effort balancing all packets to all available paths, the network can be considered as a high-bandwidth **network pipe**.
- Traditional per-packet signal can only capture congestion in **individual path**.

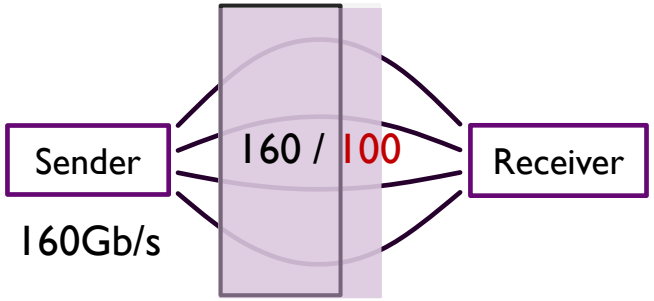
1 Non-Congestion



2 Partial Congestion



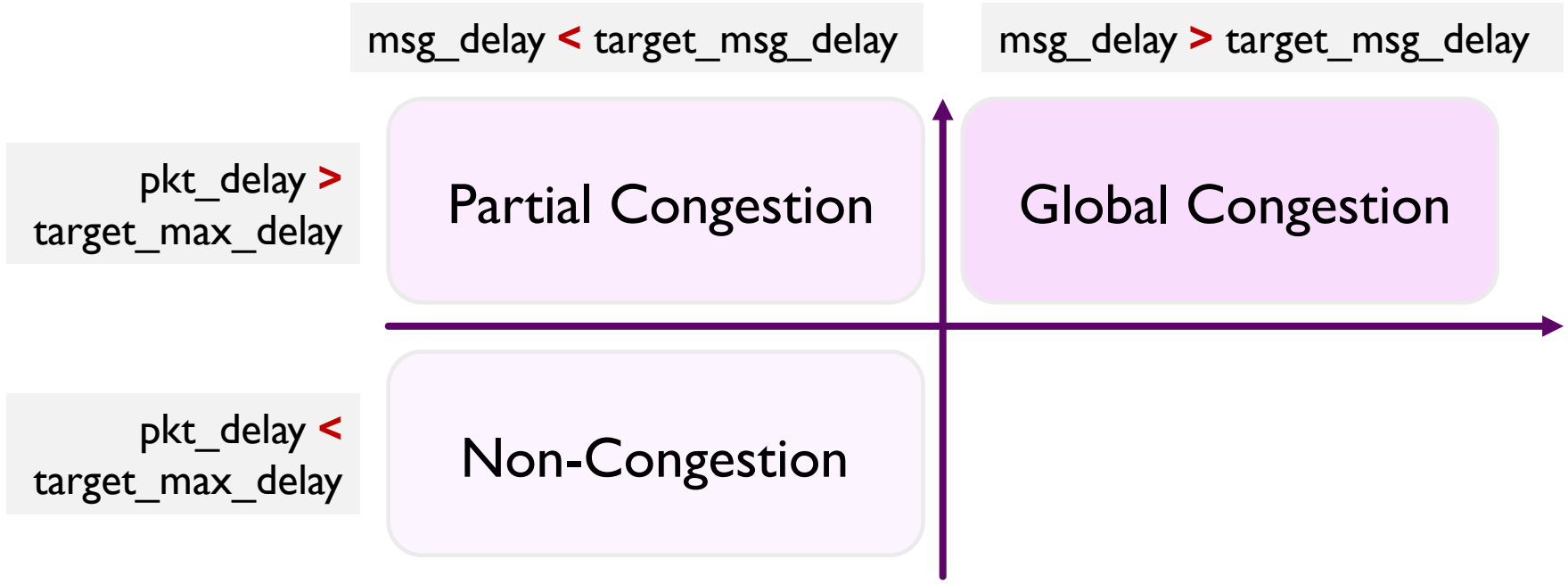
3 Global Congestion



How to appropriately interpret the congestion status in the pipe?

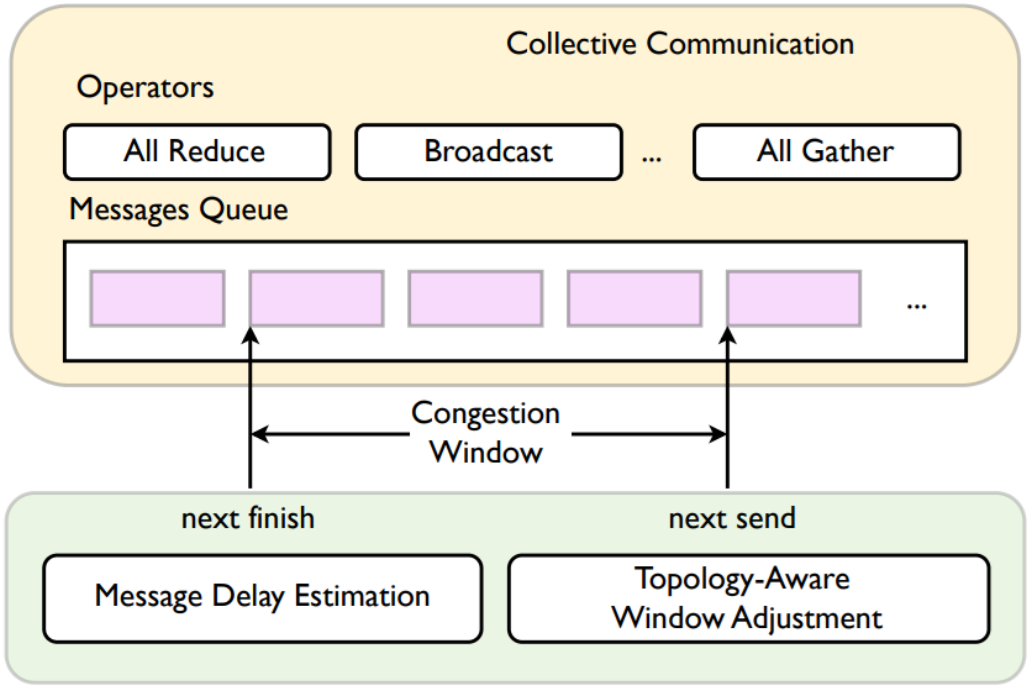
Message-level Congestion Signal

- Appropriately interpret Network Pipe congestion via Message delay and Packet delay
  - Message delay: message RTT that reveals network pipe congestion status.
  - Packet delay: sampled packet RTTs that convey multiple individual path congestion.

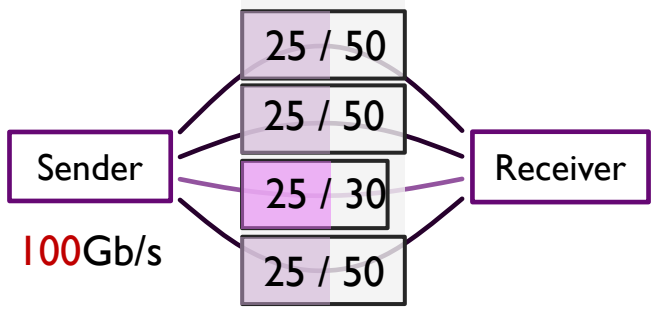


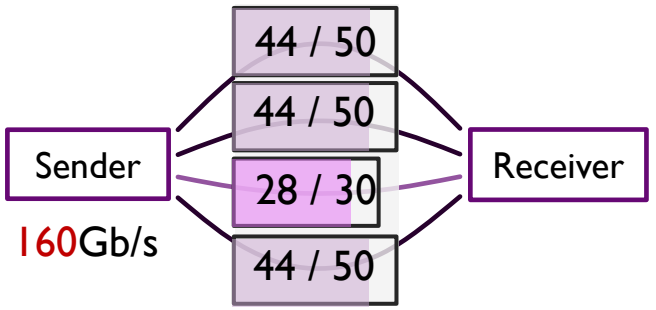
Congestion Control with Message-level Signals

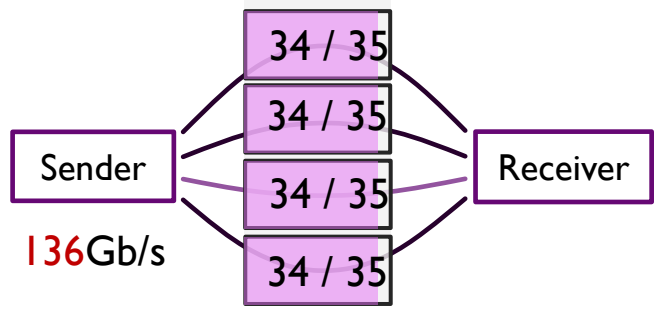
1. Window-based Control
   - Control at message granularity to regulate inflight bytes.
2. Integrate into CCLs
   - Chunk-based mechanism divides large messages into several msg chunks sized typically 64KB.



MCC: Message-level Signaling (and CC) to Eliminate Overreaction!

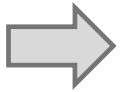






Packet-level Signals

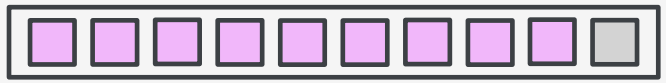
CNP indicates path congestion



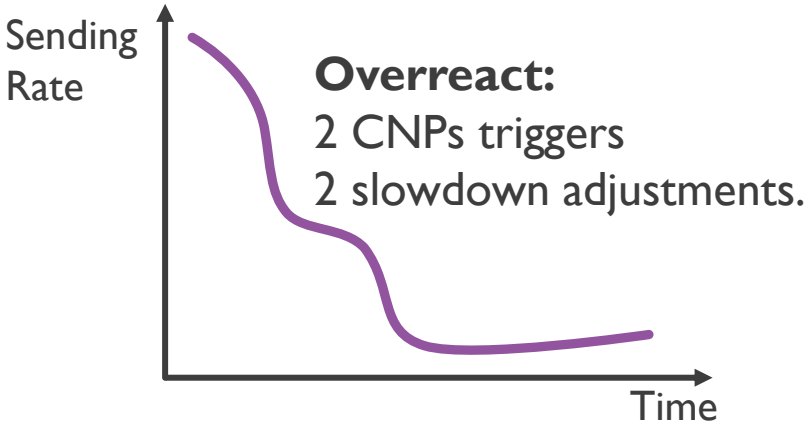
Message-level Signal

individual path is congested

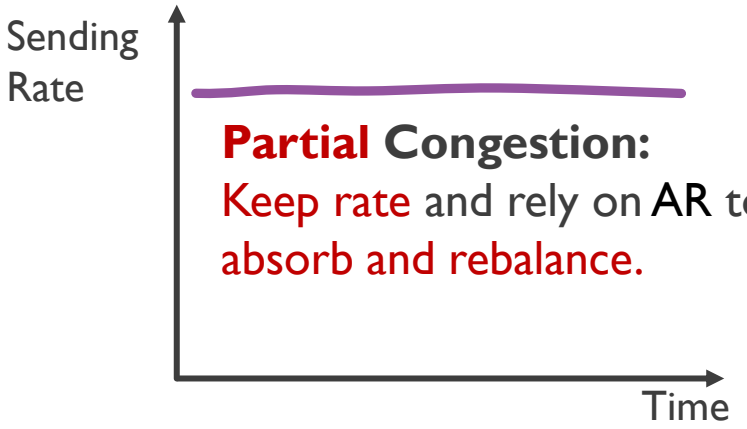
Message-level Signal



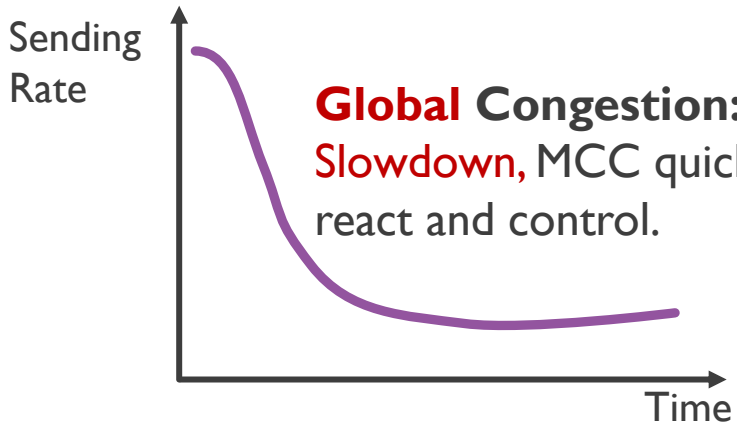
indicates all paths are congested



Overreact:
2 CNPs triggers
2 slowdown adjustments.



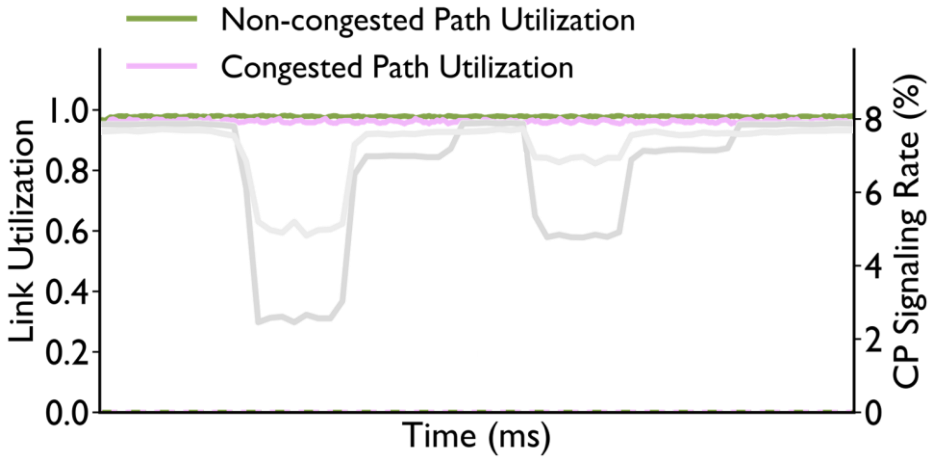
Partial Congestion:
Keep rate and rely on AR to absorb and rebalance.



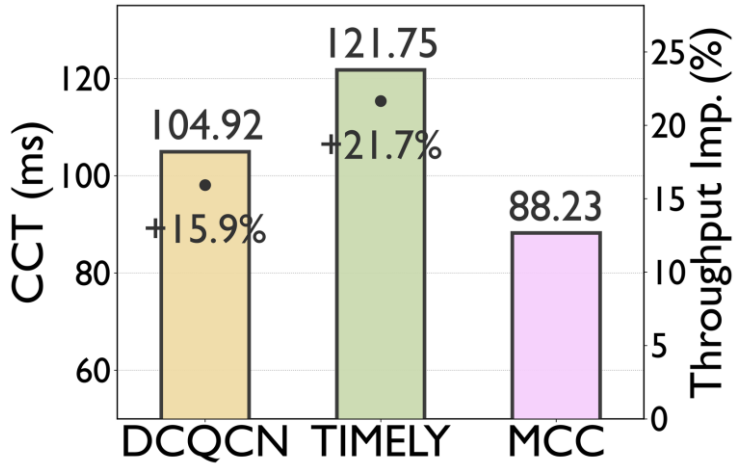
Global Congestion:
Slowdown, MCC quickly react and control.

MCC Preliminary Simulation Results

- Preliminary simulations results of MCC in NS-3.



Keep nearly 100% link utilization for all paths in partial congestion



Up to 22% higher throughput and at least 16% lower CCT

Settings:

- DCQCN and TIMELY with recommended parameters in HPCC.
- 8 servers into 4 groups (2 servers each).
- Each group executes an AllReduce operation, starting execution at the same time.

MCC ensures **compatibility** with packet spray and provides **high performance** for AI training workloads.

Congestion Control for AI Workloads with Message-Level Signaling

Yuxuan Li, Zhenghang Ren, Wenxue Li, Xiangzhou Liu, Kai Chen

iSING Lab @ HKSUT

August 2025