Active-Routing: Compute on the Way for Near-Data Processing

Jiayi Huang, Ramprakash Reddy Puli, Pritam Majumder
Sungkeun Kim, Rahul Boyapati, Ki Hwan Yum and EJ Kim

- Data is exploding

- Data is exploding
  - Requires more memory

- Data is exploding
  - Requires more memory

Die-stacked memory and memory network

- Data is exploding
  - Requires more memory

Die-stacked memory and memory network

- Large amount of data movement

- Data is exploding
  - Requires more memory

Die-stacked memory and memory network

- Large amount of data movement
  - Stall computation and consume energy

- Data is exploding
  - Requires more memory

Die-stacked memory and memory network

- Large amount of data movement
  - Stall computation and consume energy

Near-data processing in memory and network

- Data is exploding
  - Requires more memory

Die-stacked memory and memory network

- Large amount of data movement
  - Stall computation and consume energy

Near-data processing in memory and network

Active-Routing for dataflow execution in memory network

- Data is exploding
  - Requires more memory

Die-stacked memory and memory network

- Large amount of data movement
  - Stall computation and consume energy

Near-data processing in memory and network

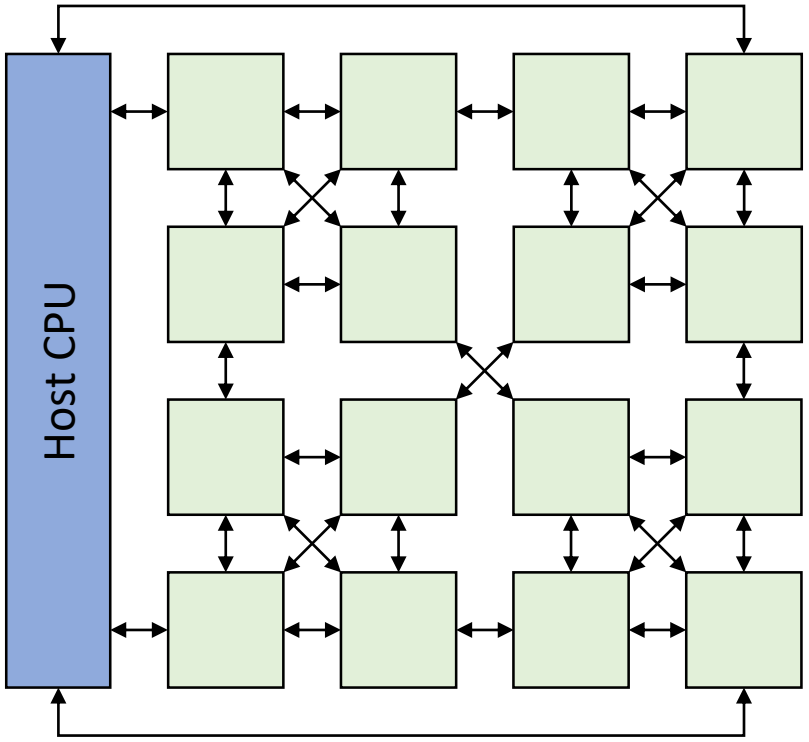
Active-Routing for dataflow execution in memory network

Exploit **memory throughput** and **network concurrency**

Active-Routing Three-Phase Processing

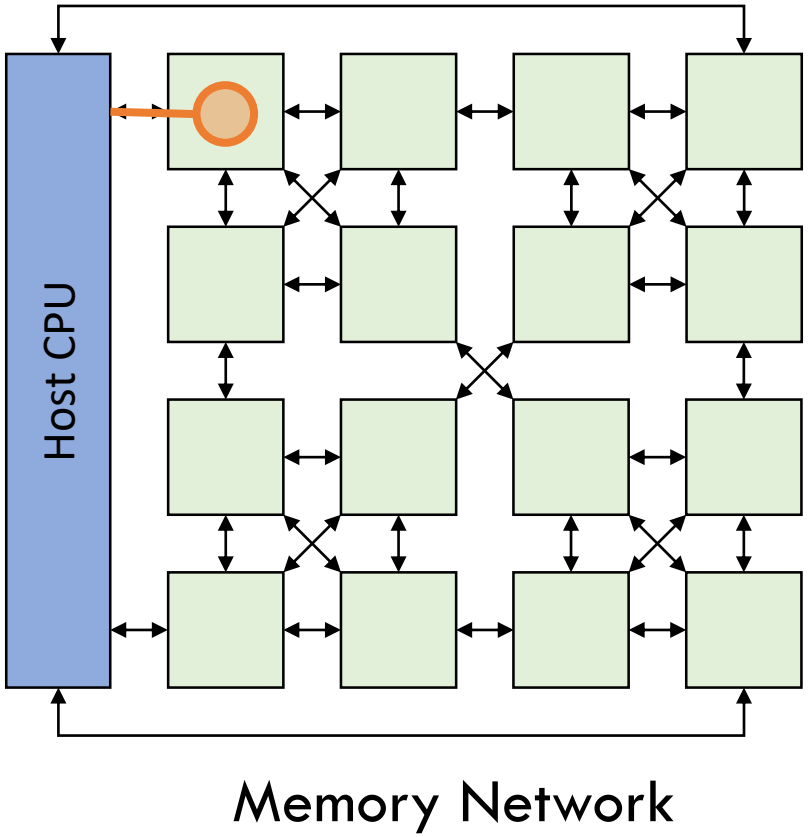
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction



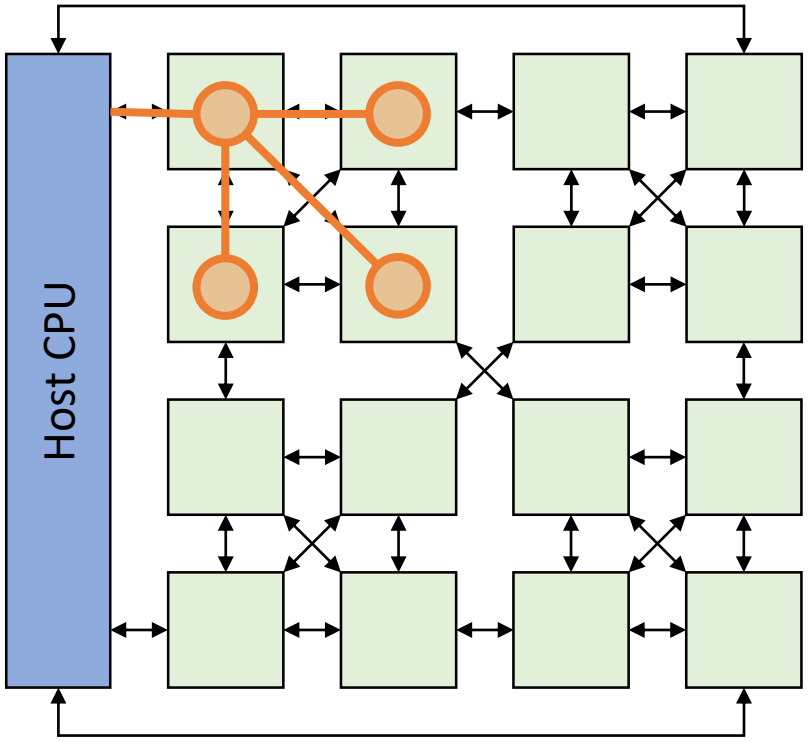
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction



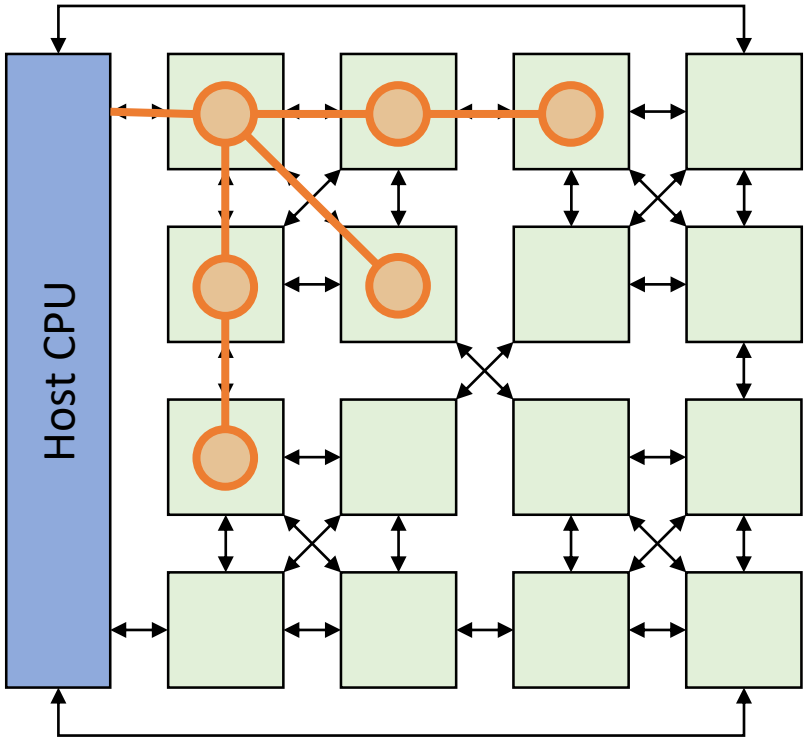
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction



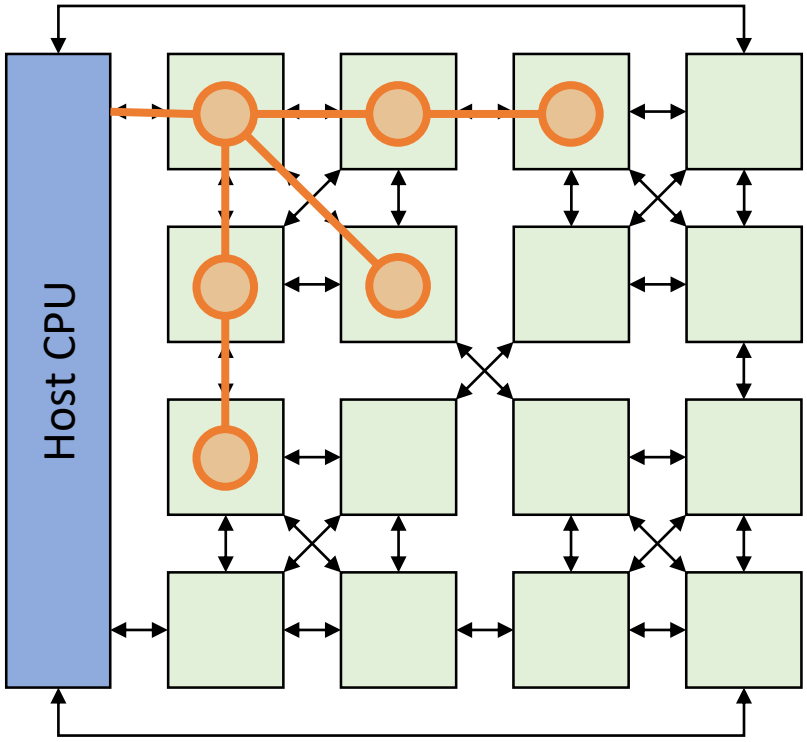
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction



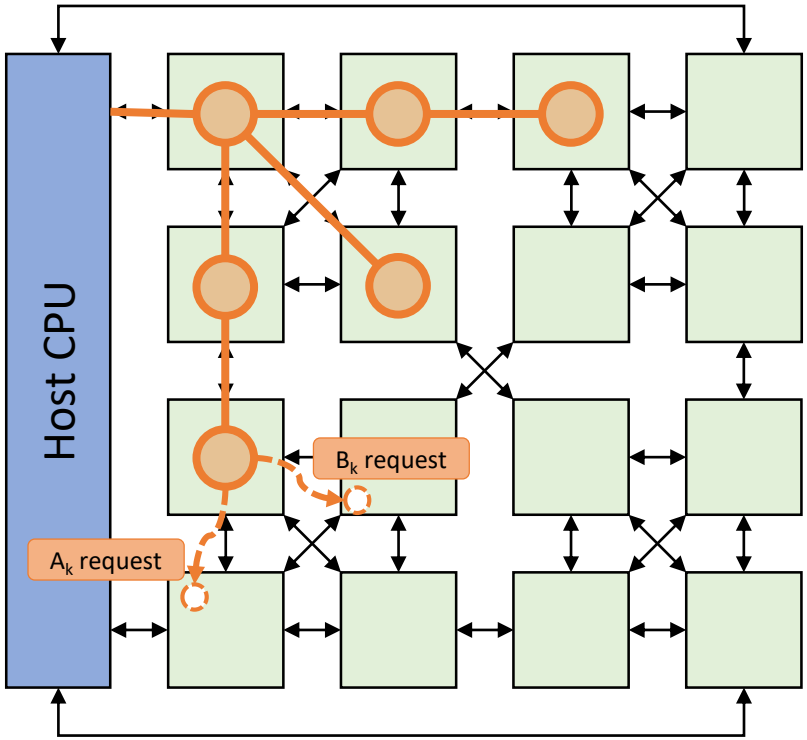
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing



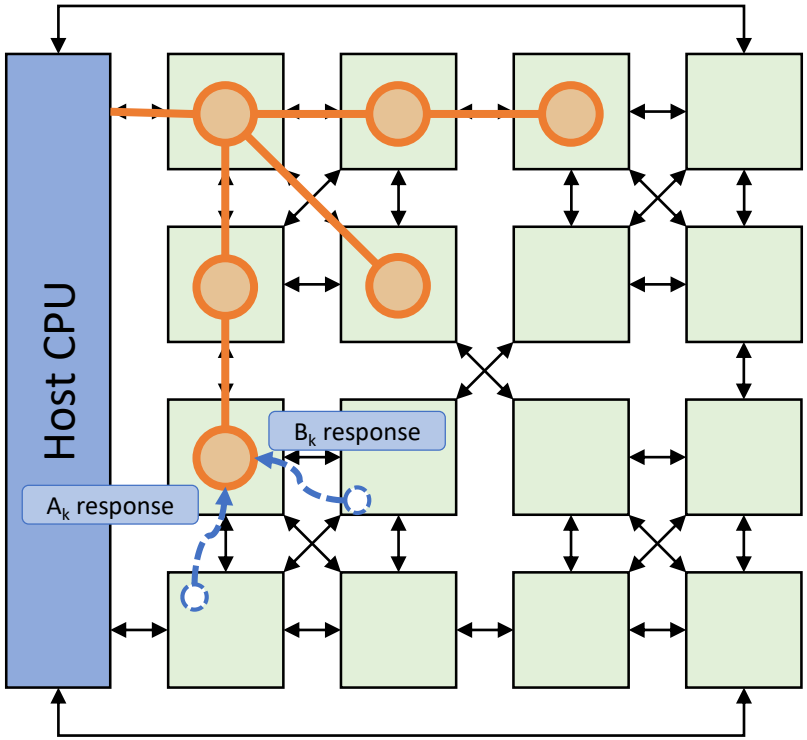
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing



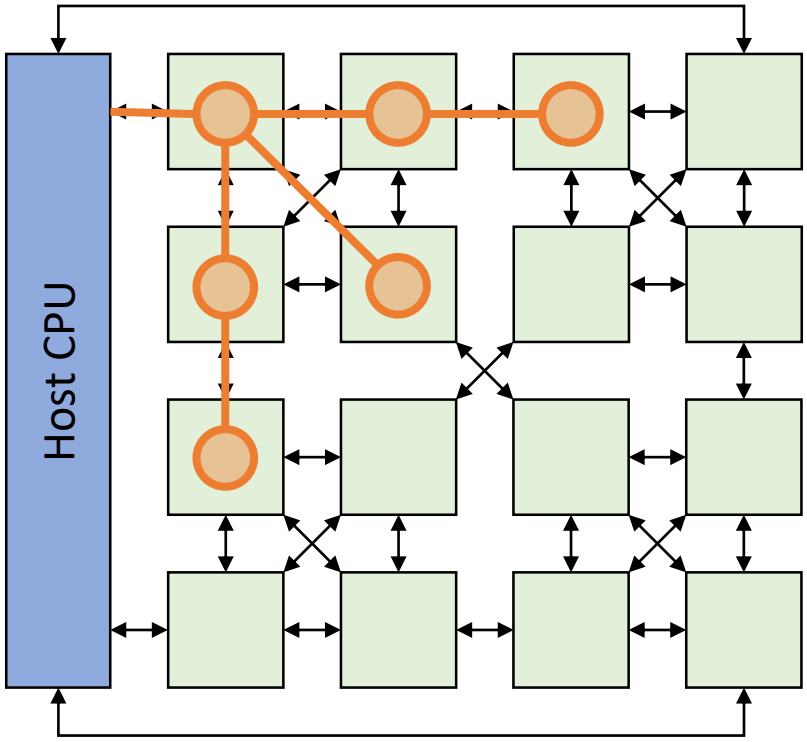
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing



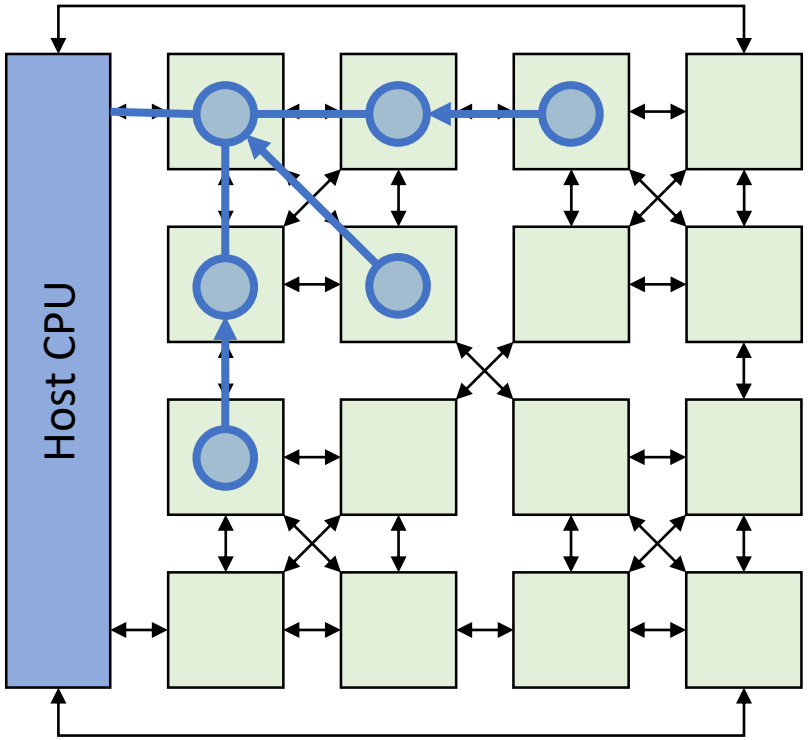
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing



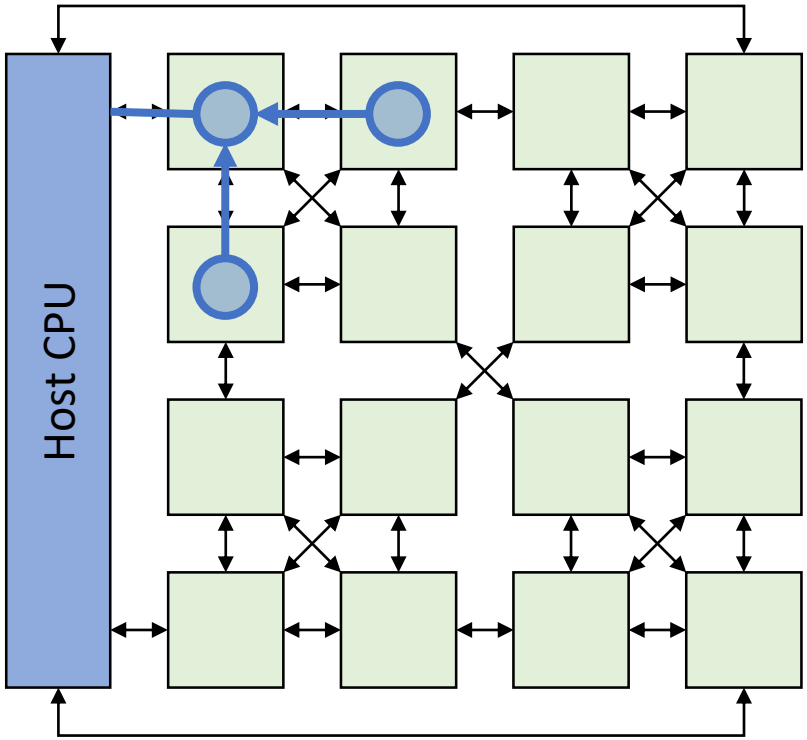
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing
- Gather Phase for tree reduction



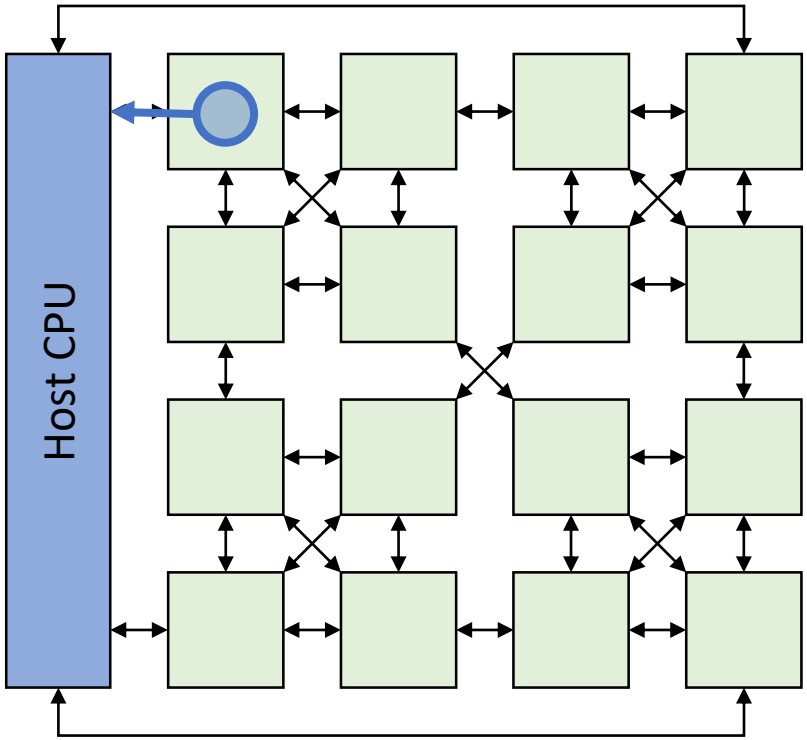
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing
- Gather Phase for tree reduction



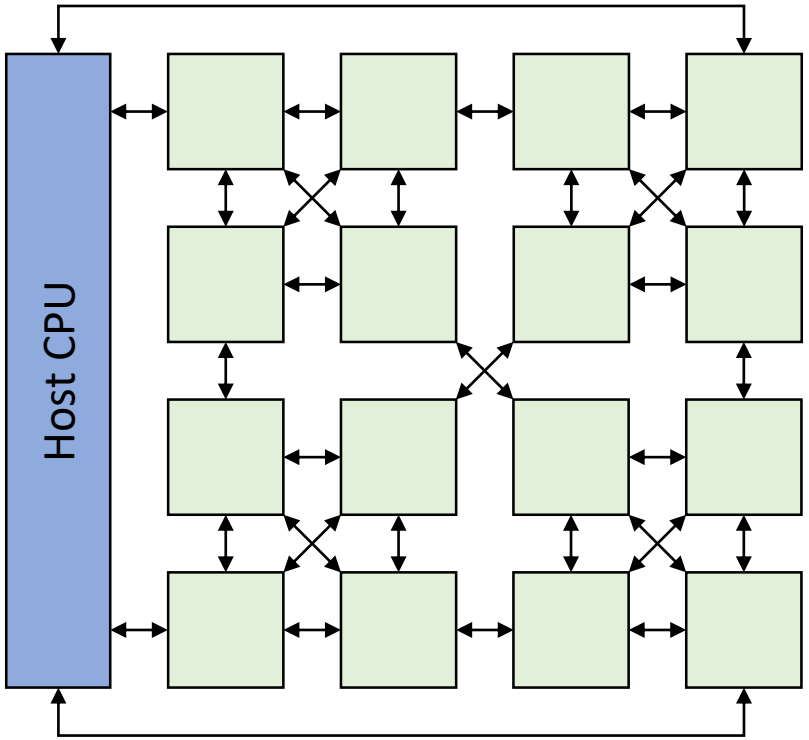
Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing
- Gather Phase for tree reduction



Active-Routing Three-Phase Processing

- Active-Routing Tree Construction
- Update Phase for data processing
- Gather Phase for tree reduction



Active-Routing: Compute on the Way for Near-Data Processing

Up to 7x speedup and improves 60% on average
reduce average energy-delay product by 80%

Wednesday, 12:10am-12:35pm, Session 9A-3