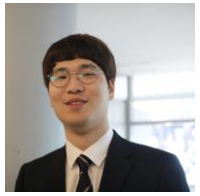
Georgia Tech School of Electrical and Computer Engineering
College of Engineering

http://synergy.ece.gatech.edu

Demo 3: Advanced ASTRA-sim Topics



William Won
Ph.D. Student, School of Computer Science
Georgia Institute of Technology
william.won@gatech.edu

Acknowledgments: Srinivas Sridharan (Meta), Sudarshan Srinivasan (Intel)

Objective

- Enabling intra/inter-dimensional scheduling
  - Using Themis scheduler
- Compiling and using different network backends
  - Congestion-aware network backend
  - (Garnet network backend)

Themis Scheduler

- Cite ISCA Paper

- Change system layer: themis and SCF
- Use 4D topology with ½ BW each level

- Use with or without themis, compare the result

Compiling ASTRA-sim with Congestion Backend

Compile ASTRA-sim with congestion-aware analytical backend

```
$ ./build_congestion.sh
```

Exercise 3-1: Running topology-unware collective

Objective:

(1) We will 1 GB All-Reduce on a **64-NPU Switch**
(2) Using **Direct All-Reduce** algorithm
(3) And compare the result

Configurations: Network

inputs/network/switch_analytical.json

```
{
  "dimensions-count": 1,
  "topologies-per-dim": ["Switch"],
  "units-count": [64],
  "links-count": [2],
  "link-latency": [500],
  "link-bandwidth": [50]
}
```

1D network

Switch topology

64 NPUs

2 links per NPU

500ns link latency

50GB/s link bandwidth

Configurations: Network

`inputs/network/switch_congestion.yml`

```
topology: Switch,
npus_count: 64,
bandwidth": [50],
latency": [500],
```

Switch topology

64 NPUs

50GB/s link bandwidth

500ns link latency

Configurations: System

inputs/system/direct.txt

```
scheduling-policy: LIFO
endpoint-delay: 10
active-chunks-per-dimension: 1
preferred-dataset-splits: 4
boost-mode: 0
all-reduce-implementation: direct
all-gather-implementation: direct
reduce-scatter-implementation: direct
all-to-all-implementation: direct
collective-optimization: localBWAware
```

ring All-Reduce Algorithm

Running Experiment

- Objective: Running
  - 1 GB All-Reduce
  - On 64-NPU Switch
  - Using Direct All-Reduce Algorithm

```
$ ./exercise_3-2.sh
```

Understanding Results

`result_3-2/tutorial_result.csv`

Name	Total Time (us)	Compute Time (us)	Exposed Communication Time (us)
analytical	**633.217**	0	**633.217**
congestion	**40008.217**	0	**40008.217**

- Add