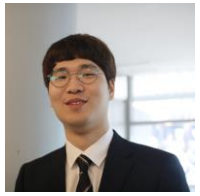
Georgia Tech School of Electrical and Computer Engineering
College of Engineering

http://synergy.ece.gatech.edu

Exercise 1: Getting Started with ASTRA-sim



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

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

Agenda

Time (CET)	Time (ET)	Topic	Presenter
15:00 – 16:00	9:00 – 10:00	**Introduction to Distributed Deep Learning Training Platforms**	Tushar Krishna
16:00 – 17:00	10:00 – 11:00	**ASTRA-sim**	Saeed Rashidi
17:00 – 17:10	11:00 – 11:10	**Break**	
17:10 – 17:50	11:10 – 11:50	**Demo and Exercises**	William Won and Taekyung Heo
17:50 – 18:00	11:50 – 12:00	**Extensions and Future Development**	Tushar Krishna and Saeed Rashidi

Tutorial Website

includes agenda, slides, ASTRA-sim installation instructions (via source + docker image)
https://astra-sim.github.io/tutorials/asplos-2022

Attention: Tutorial is being recorded

Objective

- Installing ASTRA-sim
  - Download
  - Compilation
- Writing Input Files
  - Network
  - System
  - Workload
- Running ASTRA-sim
  - Running ASTRA-sim
  - Understanding Results

Downloading ASTRA-sim

(1) Clone ASTRA-sim tutorials GitHub repository

```
$ git clone https://github.com/astra-sim/tutorials.git
$ cd tutorials/asplos2022/
```

(2) Run setup script

```
$ ./clone_astra_sim.sh
```

Compiling ASTRA-sim

(1) Go to **Exercise 1** directory

```
$ cd exercise_1/
```

(2) Compile ASTRA-sim

```
$ ./build.sh
```

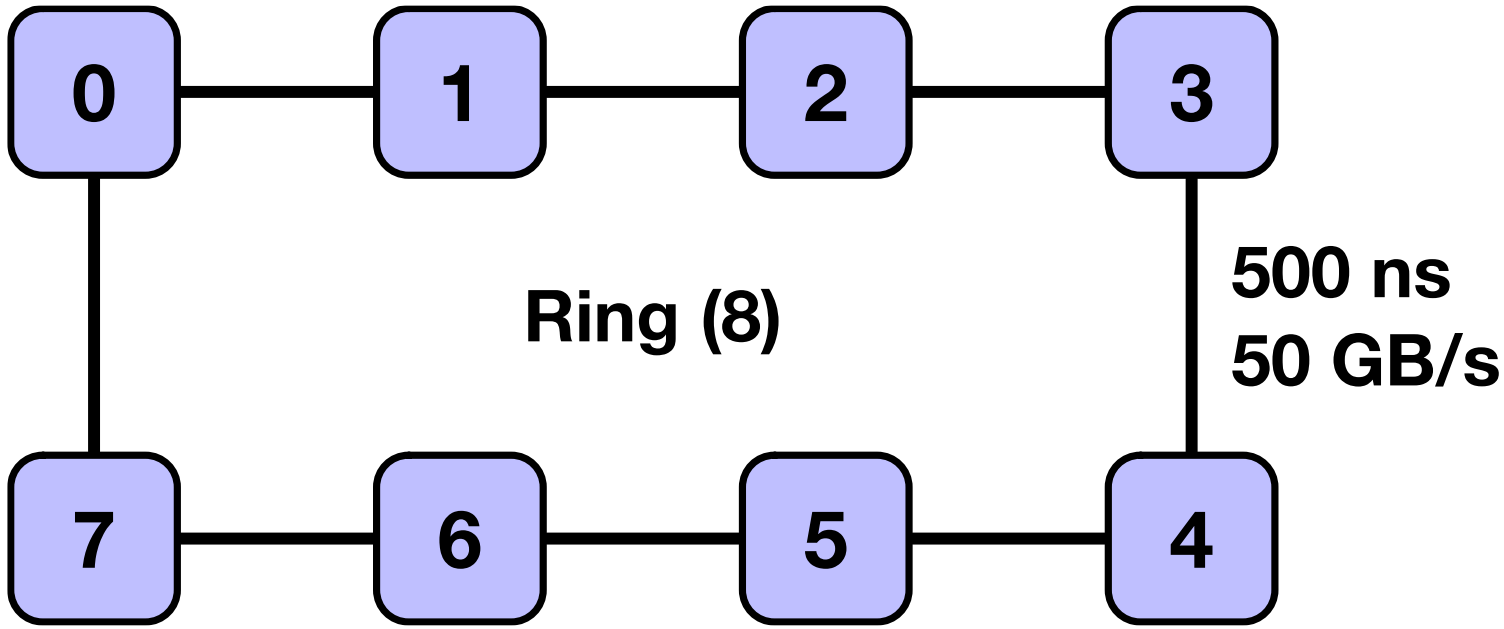
Exercise: Ring All-Reduce

Objective:

(1) We will configure an 8-NPU Ring
(2) And run **1 MB All-Reduce** on it

Configurations: Network

- **Ring** topology with **8 NPUs**
- **500 ns** (latency), **50 GB/s** (bandwidth)
- **2 links** per NPU



Configurations: Network

inputs/ring.json

```
{
  "dimensions-count": 1,
  "topologies-per-dim": ["Ring"],
  "units-count": [8],
  "links-count": [2],
  "link-latency": [500],
  "link-bandwidth": [50]
}
```

1D network

Ring topology

8 NPUs

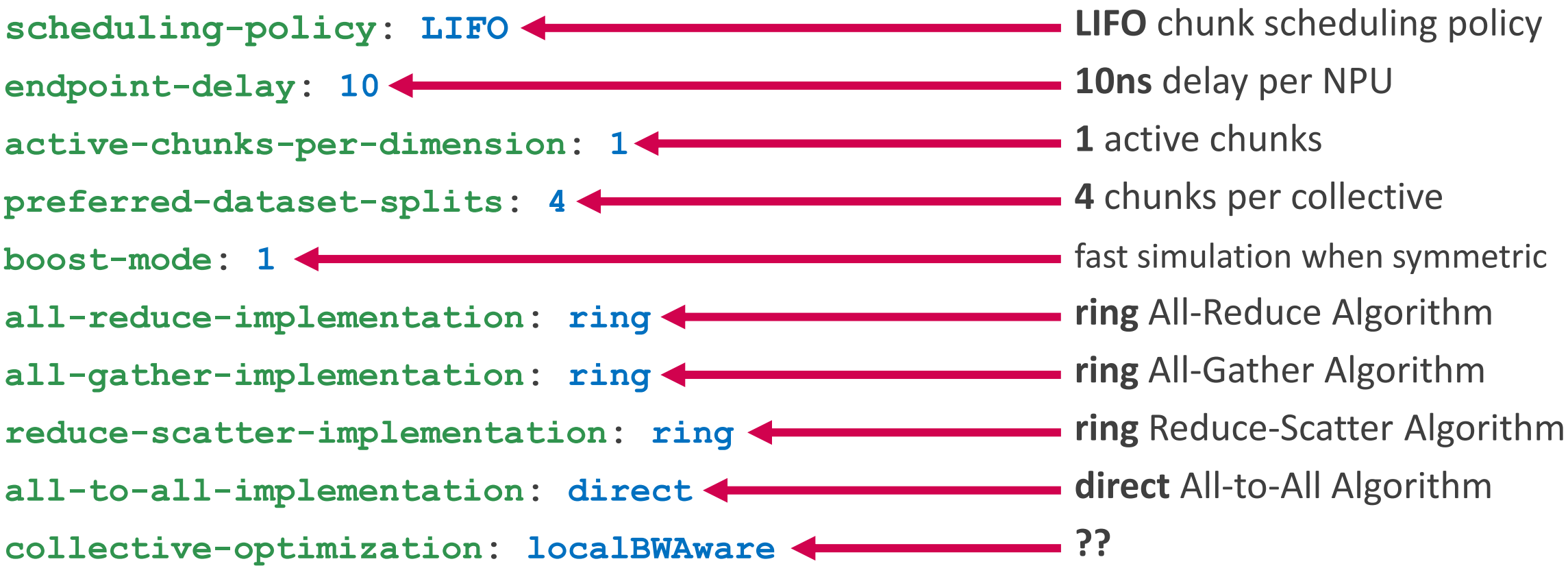
2 links per NPU

500ns link latency

50GB/s link bandwidth

Configurations: System

inputs/ring.txt



TODO: should check @Saeed about localBWAare

Configurations: System

inputs/ring.txt

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

4 chunks per collective

ring All-Reduce Algorithm

Configurations: Workload

`inputs/all_reduce.txt`

```
MICRO
1
allreduce -1 1 NONE 0 1 NONE 0 1 ALLREDUCE 1048576 1
```

training loop

#layers

layer data

Metadata		Forward			Input grad			Weight grad			Layer
Layer Name	(rsvd.)	Compute Time	Comm. Type	Comm. size	Compute Time	Comm. Type	Comm. Size	Compute Time	Comm. Type	Comm. Size	Delay
allreduce	-1	1	NONE	0	1	NONE	0	1	ALLREDUCE	1048576	1

1 MB

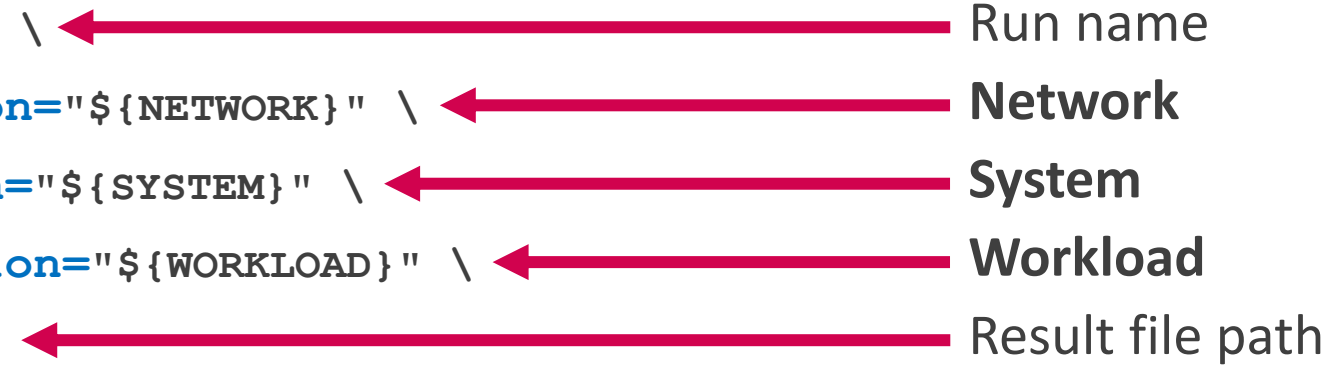
Running ASTRA-sim

Run ASTRA-sim

```
$ ./exercise_1.sh
```

exercise_1.sh

```
"${BINARY}" \
    --run-name="exercise_0" \
    --network-configuration="${NETWORK}" \
    --system-configuration="${SYSTEM}" \
    --workload-configuration="${WORKLOAD}" \
    --path="${RESULT_DIR}/"
```



Running ASTRA-sim

45,681 ns (45.681 µs)

```
all passes finished at time: 45681, id of first layer: allreduce
path to create csvs is: /usr/scratch/will/tutorials/asplos2022/exercise_1/result/
success in openning file
*****
Time to exit: Sun Feb 27 06:46:51 2022
all-reduce Collective implementation: ring
reduce-scatter Collective implementation: ring
all-gather Collective implementation: ring
all-to-all Collective implementation: direct
Collective optimization: localBWAware
Total sim duration: 0:0 hours
Total streams injected: 4
Total streams finished: 4
Percentage of finished streams: 100 %
*****
Exiting
```

Understanding Results

`result/tutorial_result.csv`

Name	Total Time (us)	Compute Time (us)	Exposed Communication Time (us)	Total Message Size (MB)
Exercise 1	45.681	0	45.681	1.75

45.681 µs

No compute

All communication exposed

1.75 MB/NPU