

Modeling Multi-tenant Scheduling in ML Clusters using Astra-sim

Shawn Shuoshuo Chen¹

Collaborators:

Daiyaan Arfeen^{1,4}, Eric Ding^{2,4}, Peter Steenkiste¹, Srinu Seshan¹, Minlan Yu³

June 27, 2026

1



2



3



4



Today's talk

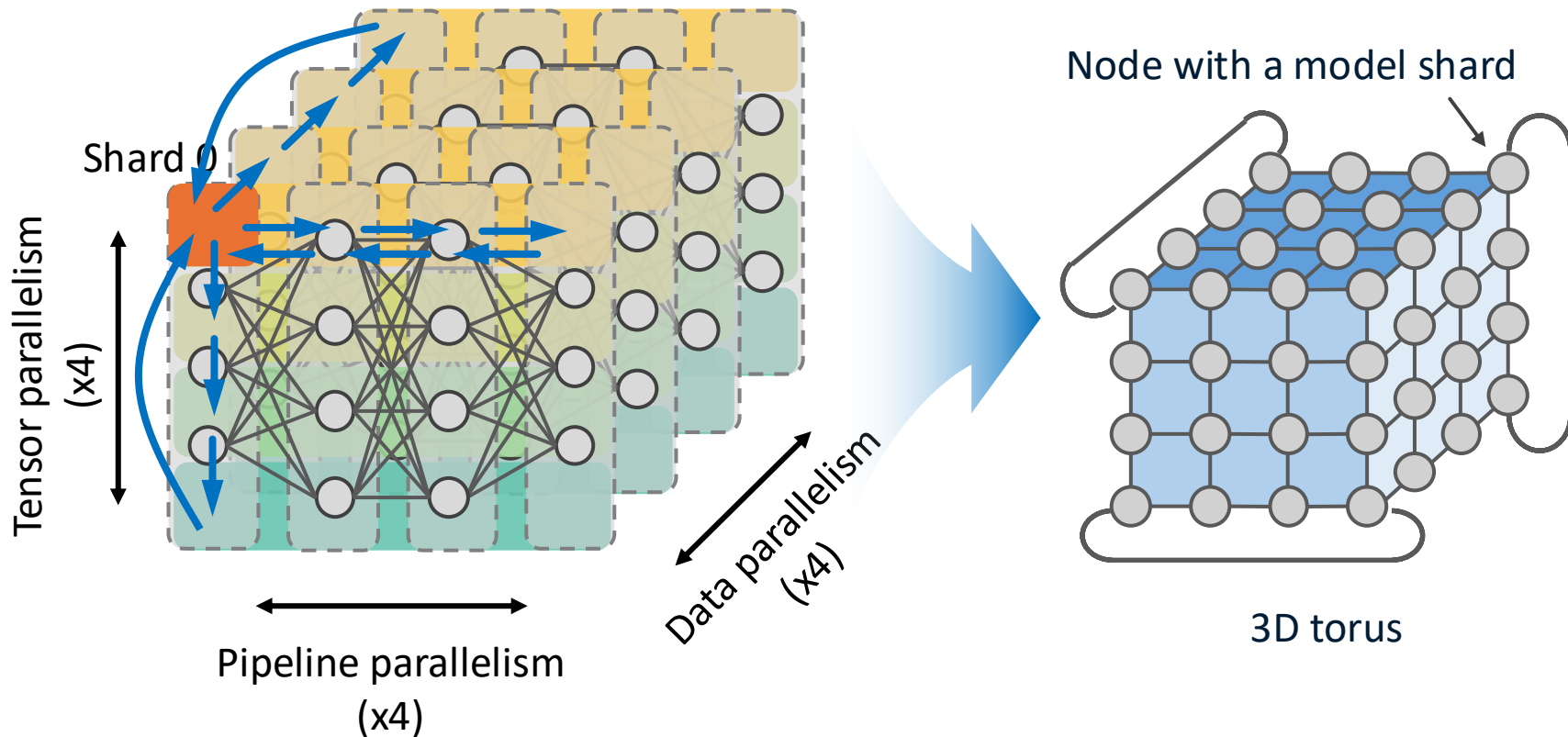
- 01 Introducing multitenancy and scheduling
- 02 Supporting reconfigurable networks in astra-sim
- 03 Supporting multitenant job traces in astra-sim
- 04 Evaluation

Why scheduling matters in **multitenant** ML clusters?

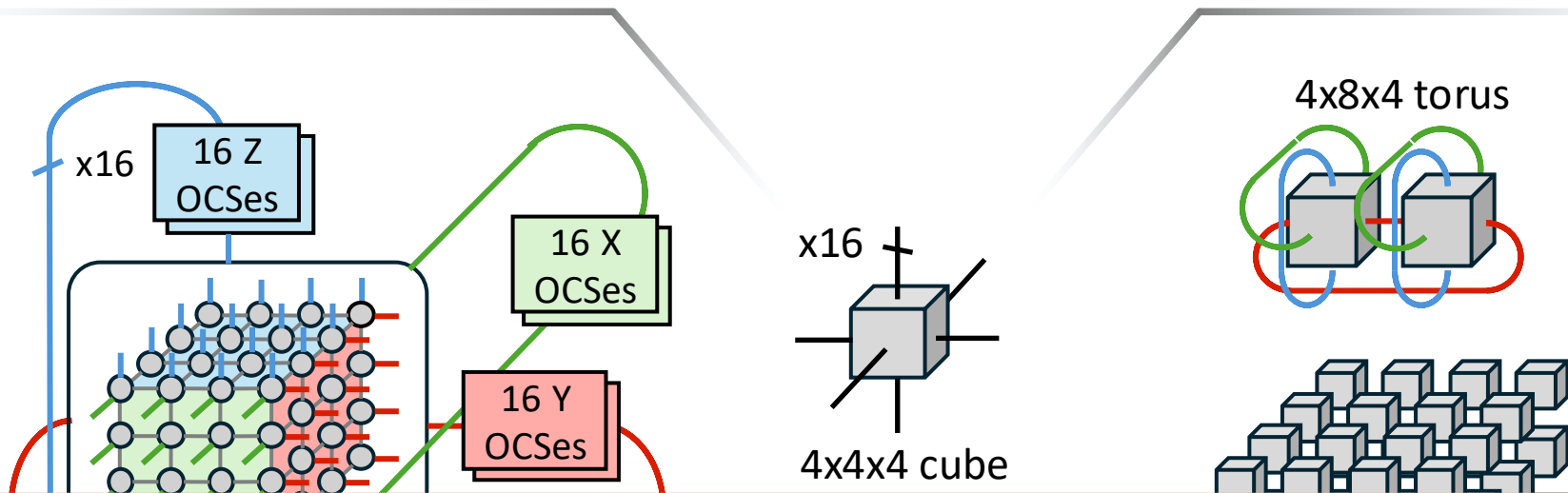
ML clusters experience performance degradation under multitenant workloads due to:

- **Topology Constraints:** ML parallelization strategies require specific GPU allocation to maintain near-neighbor routing efficiency.
- **Network contention:** Co-located jobs contend for physical link bandwidth.
- **Fragmentation:** Contiguous placement of jobs with various sizes creates “holes”.

ML workloads rely on near-neighbor communication.

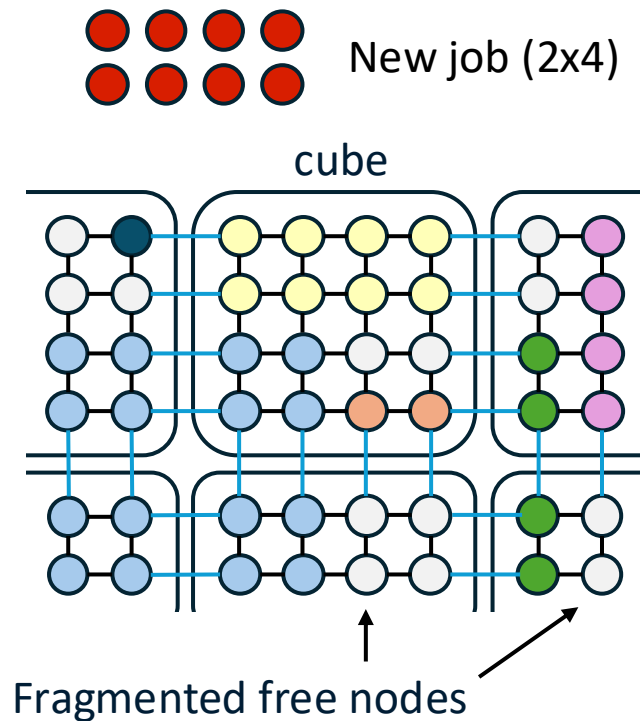


Architecture of Google's 3D torus TPU cluster



Cluster is reconfigurable at the cube level.

Fragmentation and the contention-utilization tradeoff

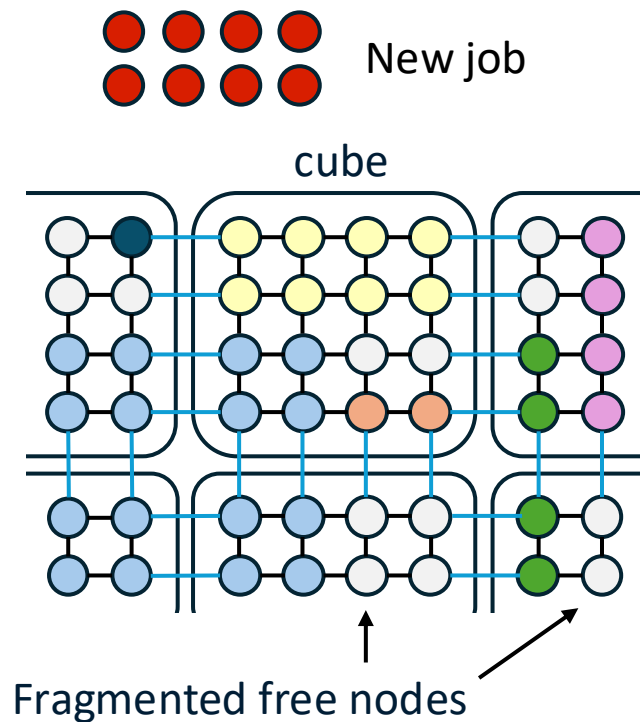


Options for the new job?

1. Enforce shape (contiguous)

- 👍 Optimal job runtime
- 👎 Extra queueing time
- 👎 Low utilization

Fragmentation and the contention-utilization tradeoff



Options for the new job?

1. Enforce shape (contiguous)



Optimal job runtime



Extra queueing time



Low utilization

2. Relax shape (non-contiguous)



Minimum queueing time



High utilization



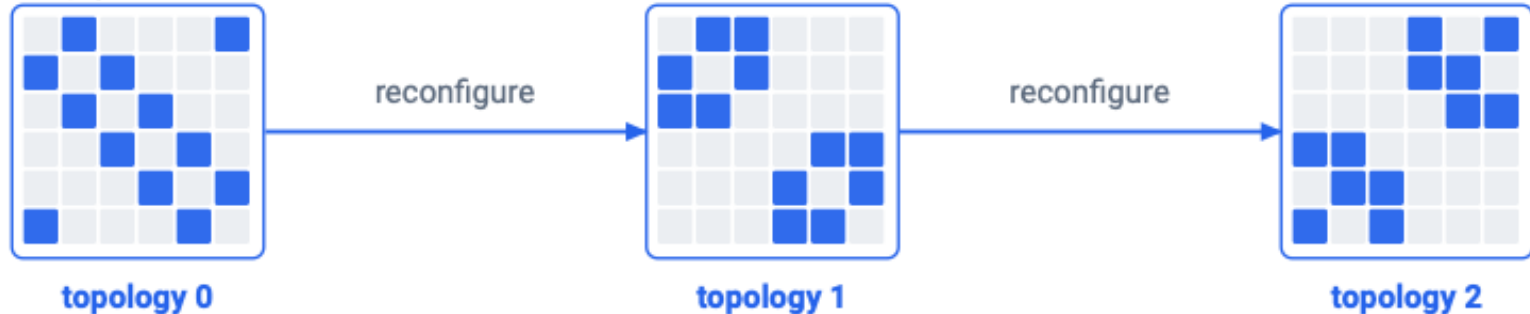
Longer runtime due to network contention

- Need to route through other jobs.
- 2.5x slower than option 1.

Supporting reconfigurable networks in astra-sim

- A new (runtime) reconfigurable network backend
 - Based on analytical model
- Arbitrary topology modeled as bandwidth/latency matrix

lit cell = a link
empty cell = no link

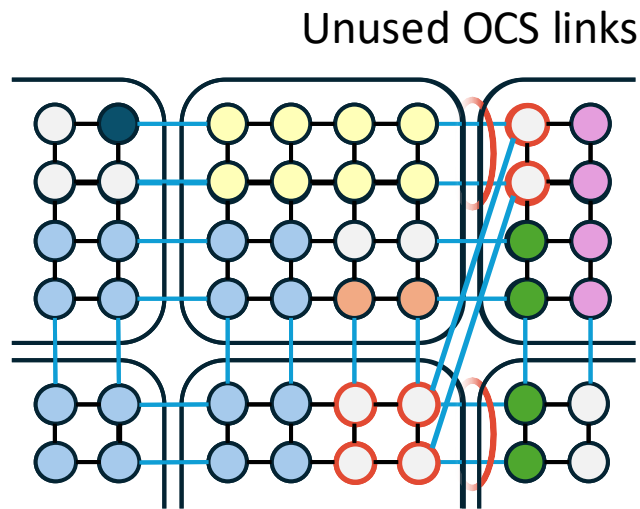


Supporting reconfigurable networks in astra-sim

- Reconfiguration allowed on idle nodes (between jobs)
- Change part of the bandwidth/latency matrix in runtime.
 - NB: TPU cluster has more specific constraints.

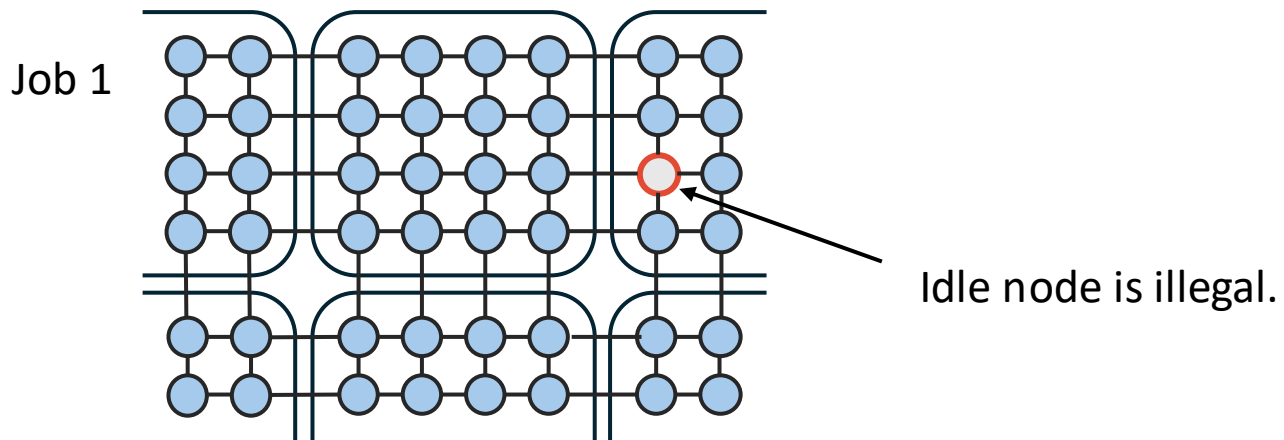
Reconfigure():

- Wait for in-flight comms to drain on affected links
- Update in-memory topology
- Recompute routing table

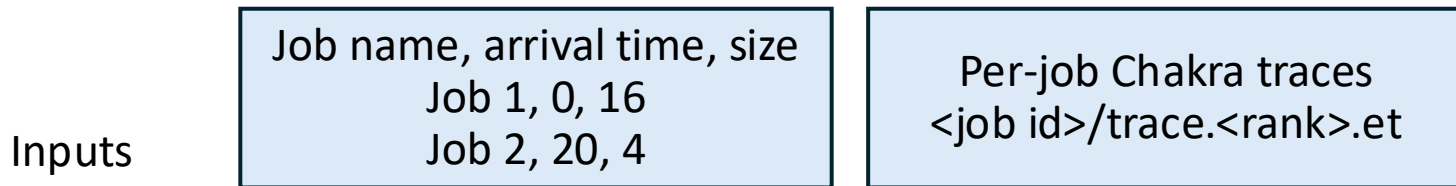


Supporting multitenant job traces in astra-sim

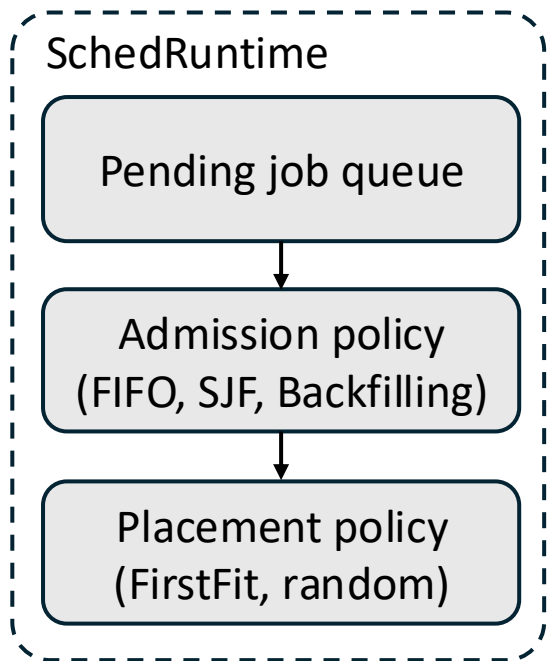
- Vanilla astra-sim:
 - one job per simulation
 - 1:1 rank-node mapping



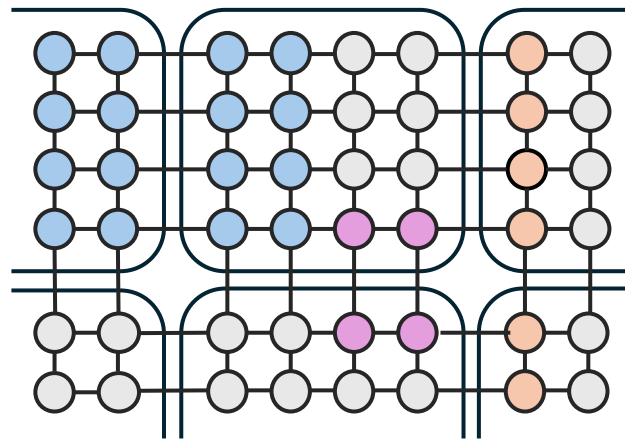
Supporting multitenant job traces in astra-sim



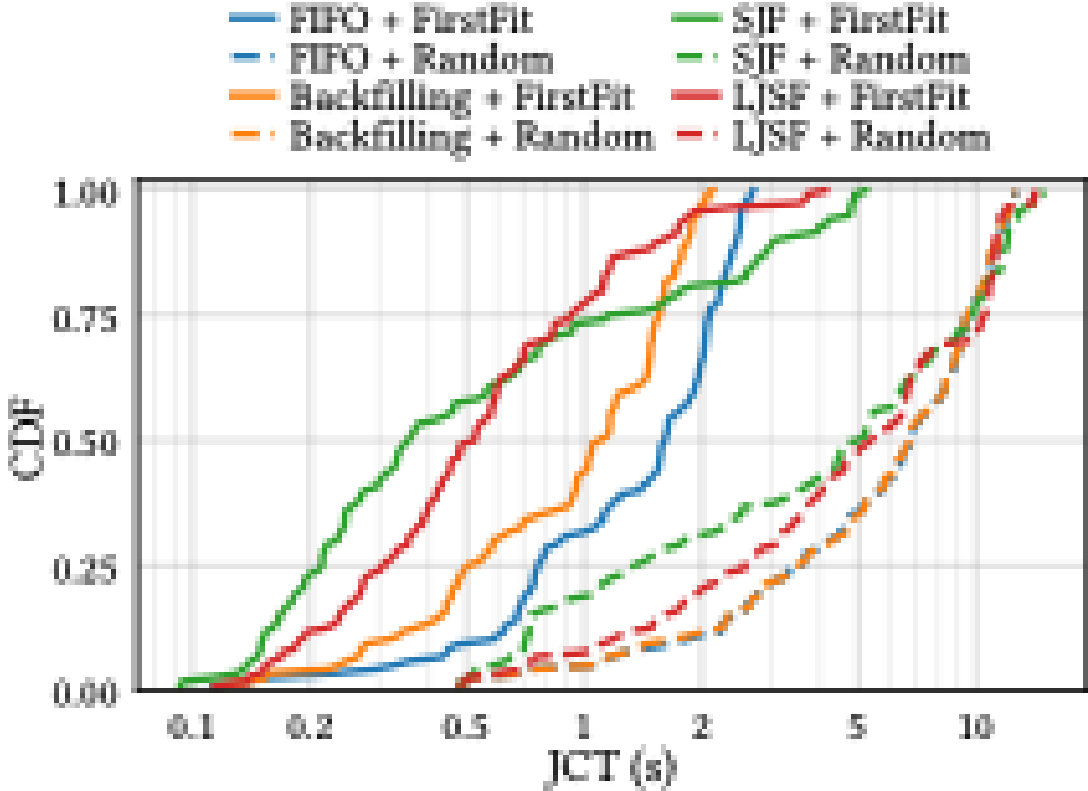
Astra-sim



Shared cluster



Evaluation: Comparing scheduling policies in astra-sim



Summary

Astra-sim extended to support

- Reconfigurable network backend
- Multitenant job traces

Development completed on modeling TPU clusters and scheduling policies.

More interesting work planned

- Storage networking
- Scale up/out/across modeling

Let's chat!

Backups