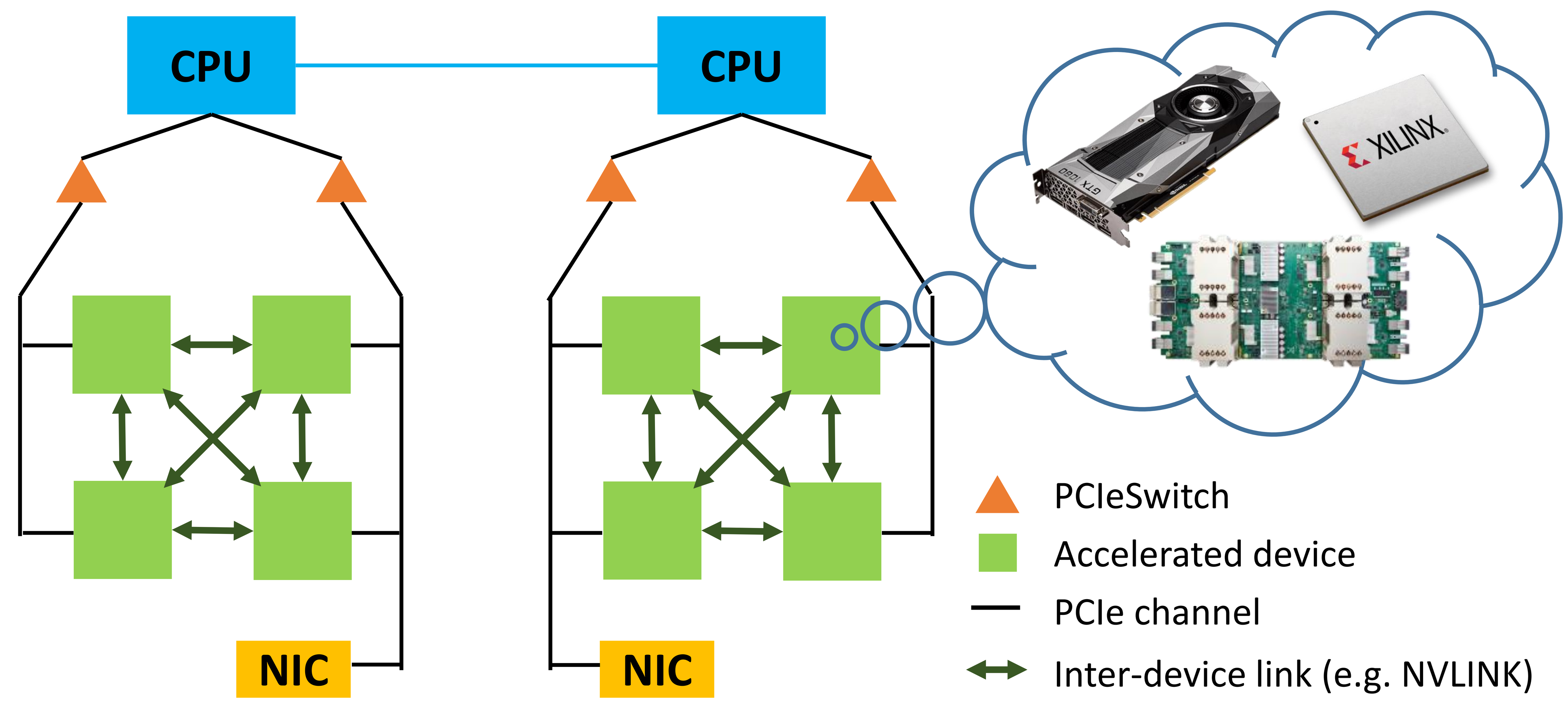
All You Need to Know about Scheduling Deep Learning Jobs

Student Research Competition #35
Wencong Xiao
Beihang University and Microsoft Research

Motivating Scenario

Deep learning shows great power in many areas → # DeepLearningJobs increase dramatically → Manage deep learning in AI infrastructure!
- improve resource utilization
- hide system issues for users

Heterogeneity in resource abstraction



Extensible resource abstraction

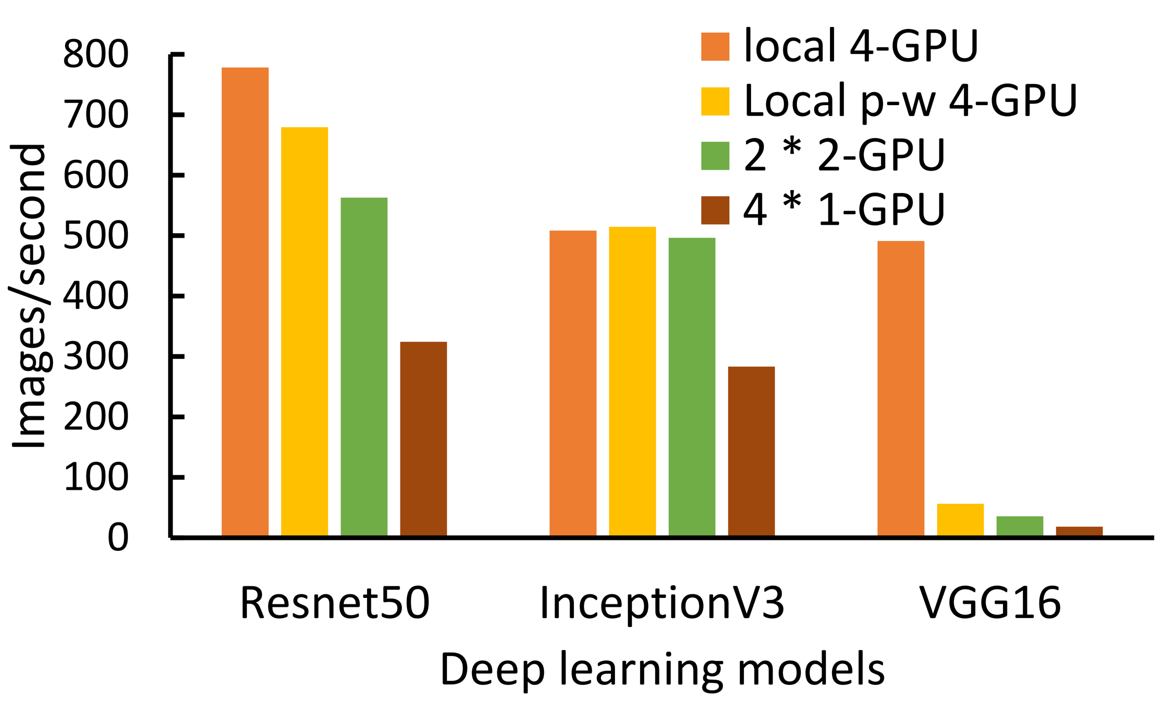
- *Diversified* & *fast evolving* hardware
  - Different generations of GPU
  - FPGA
  - ASIC (e.g. TPU, Cambricon)
- *Complicated* & *vendor-specific* topology
  - Motherboard
  - Inter-device link

Lead to diversified resource requests

- Resource type
- Locality sensitivity

Deep learning jobs characteristic

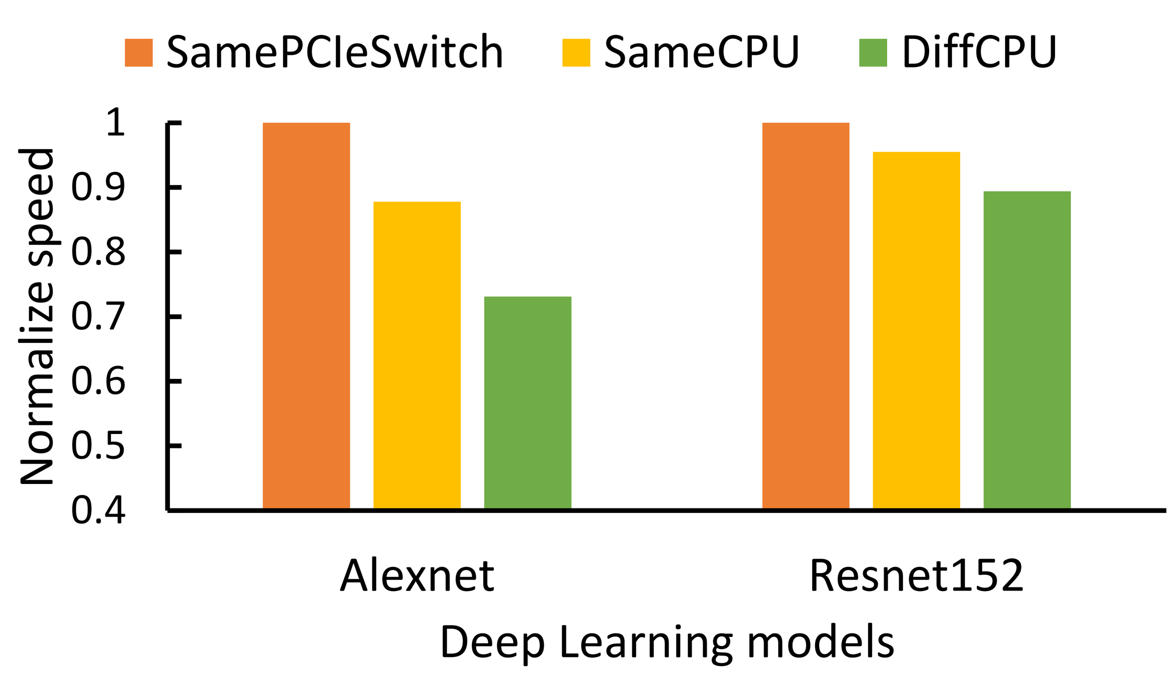
Inter-server locality



Experiment setting
- Local: multi-GPU in single machine
- 2*2-GPU: 2 machines with 2 GPUs each

Result
- Spreading resource among different servers slowdown the performance
- Different workloads can tolerate different level of resource spreading

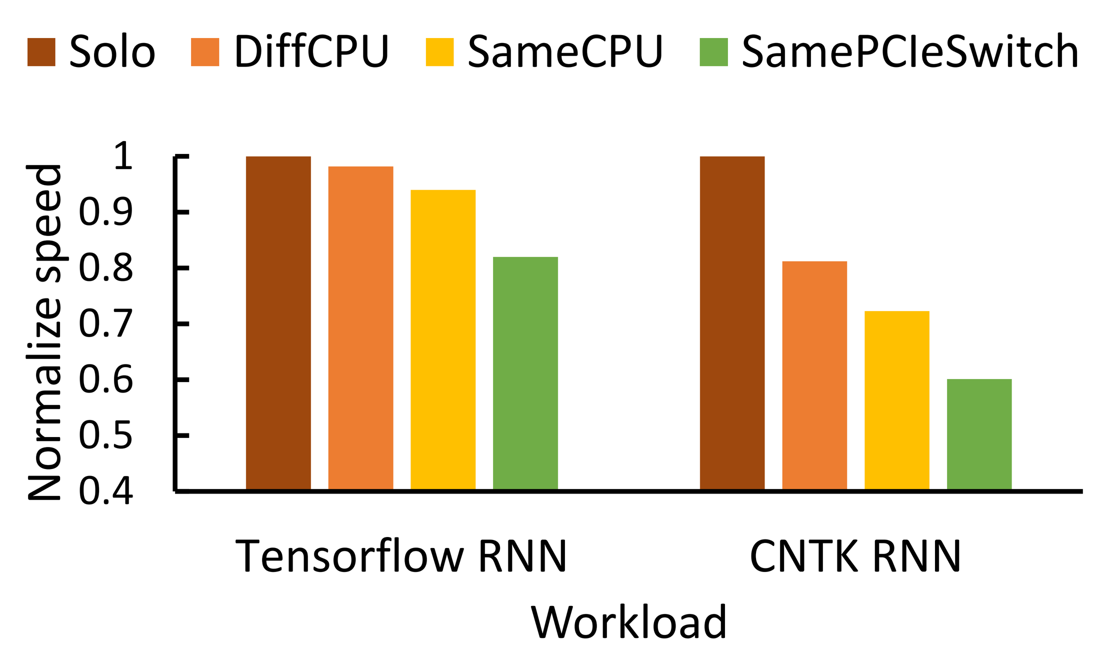
Intra-server locality



Experiment setting
- 2-GPU jobs run on different GPUs in the same machine

Result
- Up to 27% performance slowdown for sub-optimal GPU co-location

Job interference



Experiment setting
- Solo: a 1-GPU job runs solely as baseline
- Others: two 1-GPU jobs run on different GPUs in the same machine

Result
- Up to 40% performance slowdown for job interference

Conflict scheduling policy
- Inter-job policy: spread out among different machines to avoid interference -> *resource fragmentation*
- Intra-job policy: prefer closely co-location for better performance -> *require consecutive resource*

Multi-GPU jobs suffer
- 80% deep learning jobs in Microsoft cluster are 1-GPU jobs
- Large jobs
  - Long queuing delay for strict locality requirement
  - Sub-optimal performance for inconsecutive devices

Conclusion and future work: New Scheduling System for Heterogeneous Datacenter

- Flexible and compact resource abstraction with locality and topology awareness
- Decentralized design to decouple cluster-wide policy from individual job scheduling decision
- Static workload pattern analysis and dynamic job migration