Deep Learning at 15 PF

Supervised and Semi-supervised Pattern Classification for Scientific Data

---

Thorsten Kurth, Jian Zhang, Nadathur Satish, Evan Racah, Ioannis Mitliagkas, Md. Mostofa Ali Patwary, Tareq Malas, Narayanan Sundaram, Wahid Bhimji, Mikhail Smorkalov, Jack Deslippe, Mikhail Shiryaev, Srinivas Sridharan, Prabhat, Pradeep Dubey

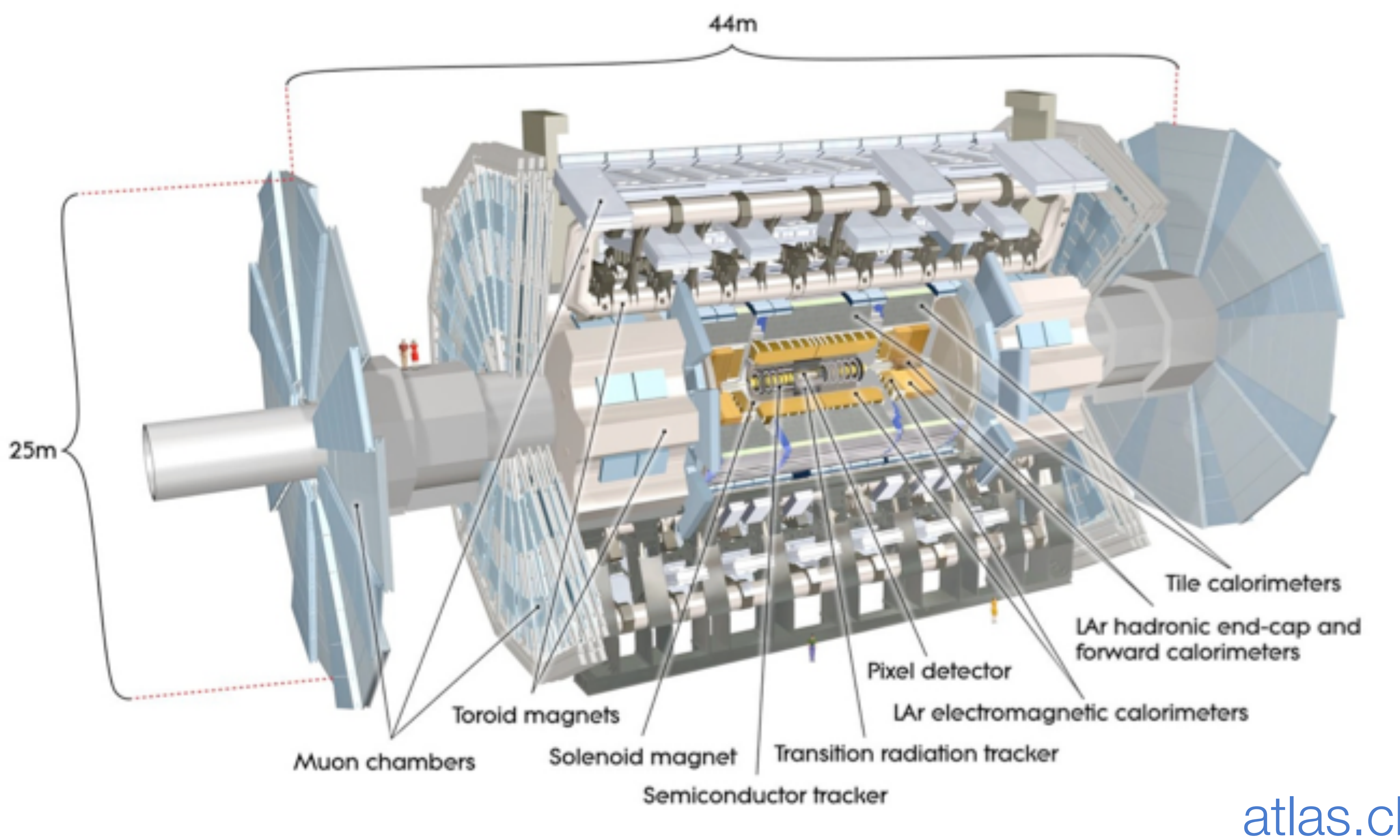
NIPS17
Long Beach, CA, USA

DOI: 10.1145/3126908.3126916

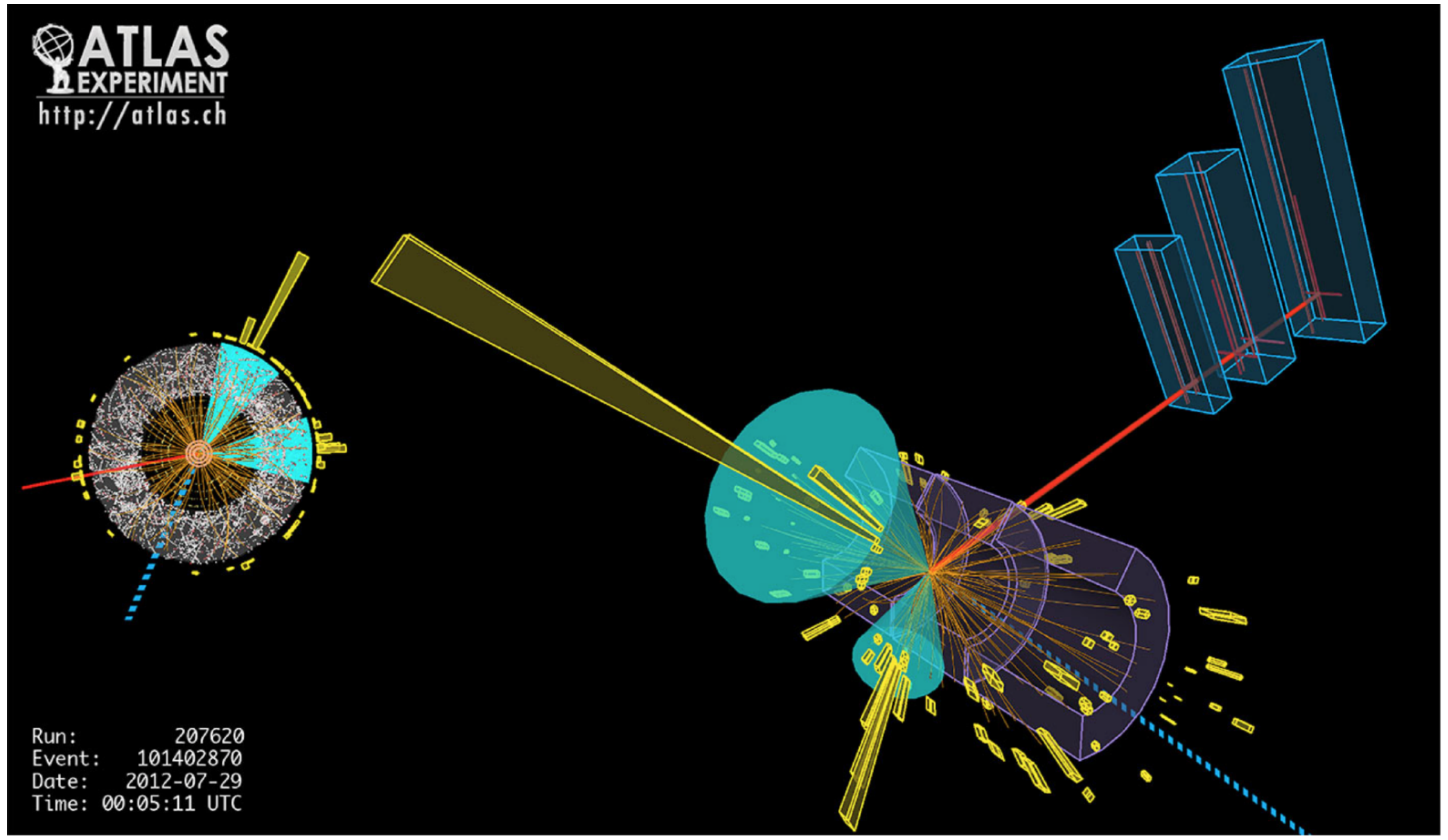
High-Energy Physics (HEP)

Finding New Physics Candidates

- Find rare signals of new particles in collisions (events) at LHC
- Represent data from cylindrical detector as sparse 2D (224x224) image
- 3 channels for different instruments (hadron & EM calorimeter, track multiplicity)
- ~10M event (7.4 TB) simulated dataset
- Existing selections on derived high-level physics variables used as a benchmark
- **binary classification**



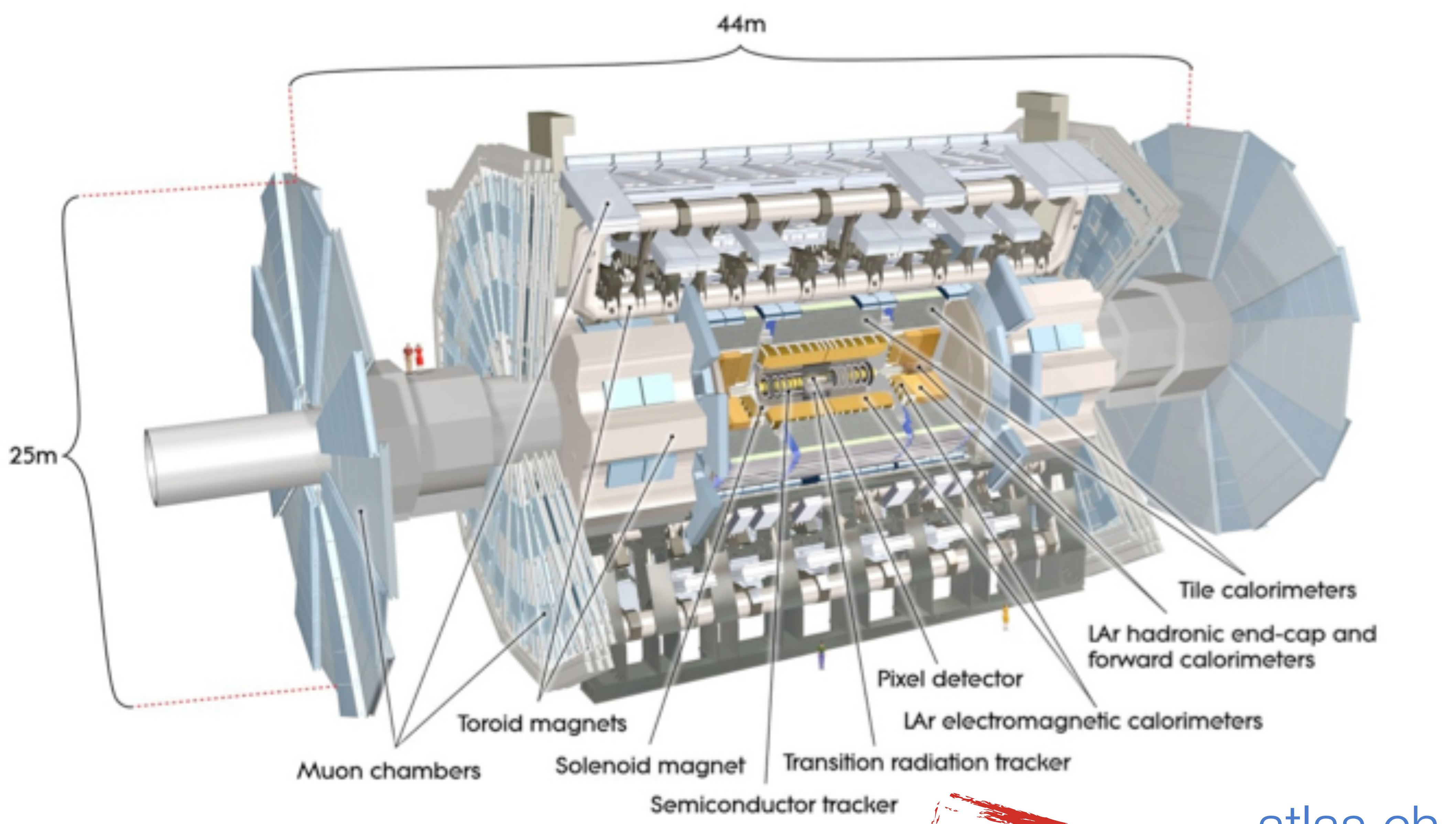
atlas.ch



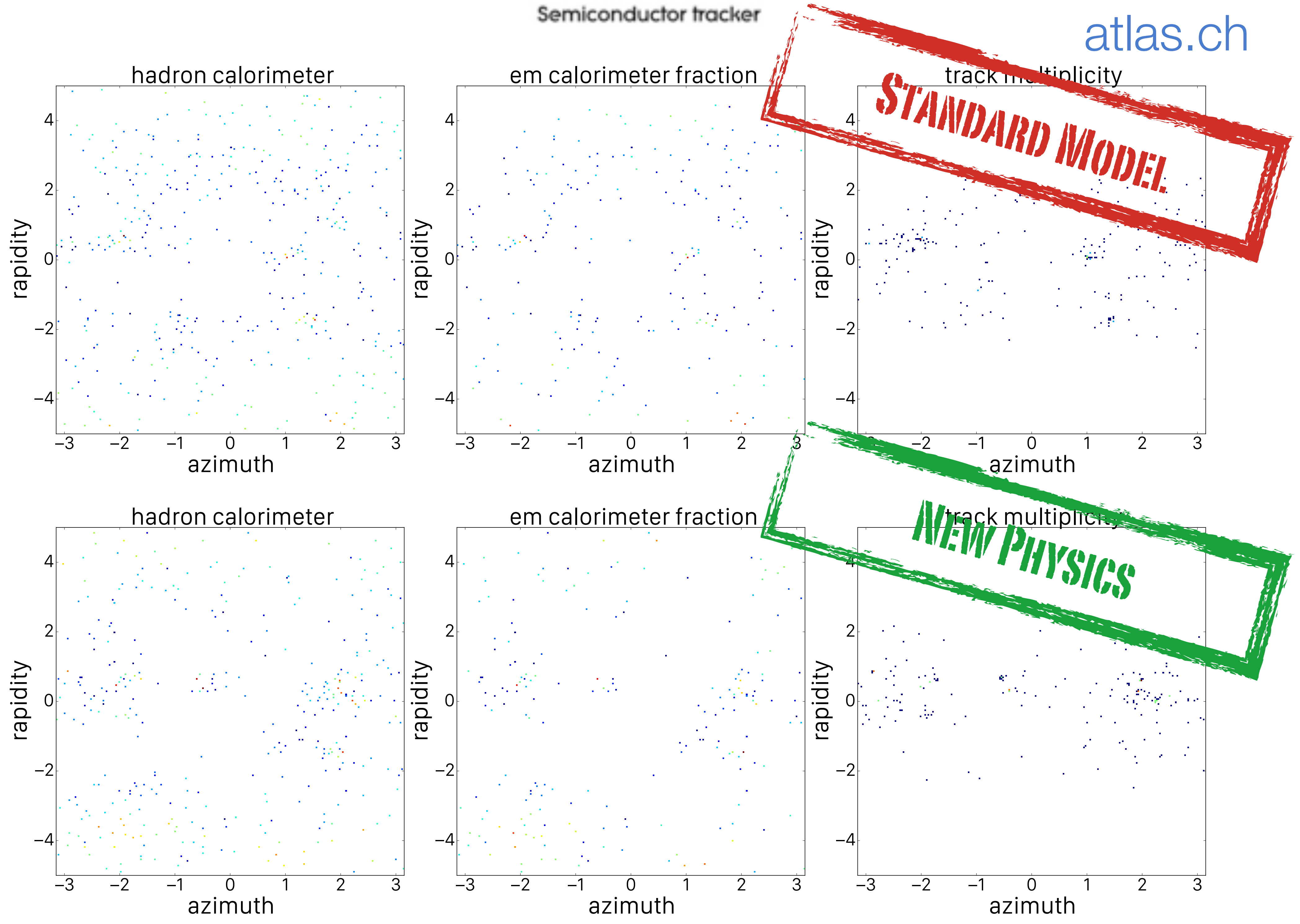
High-Energy Physics (HEP)

Finding New Physics Candidates

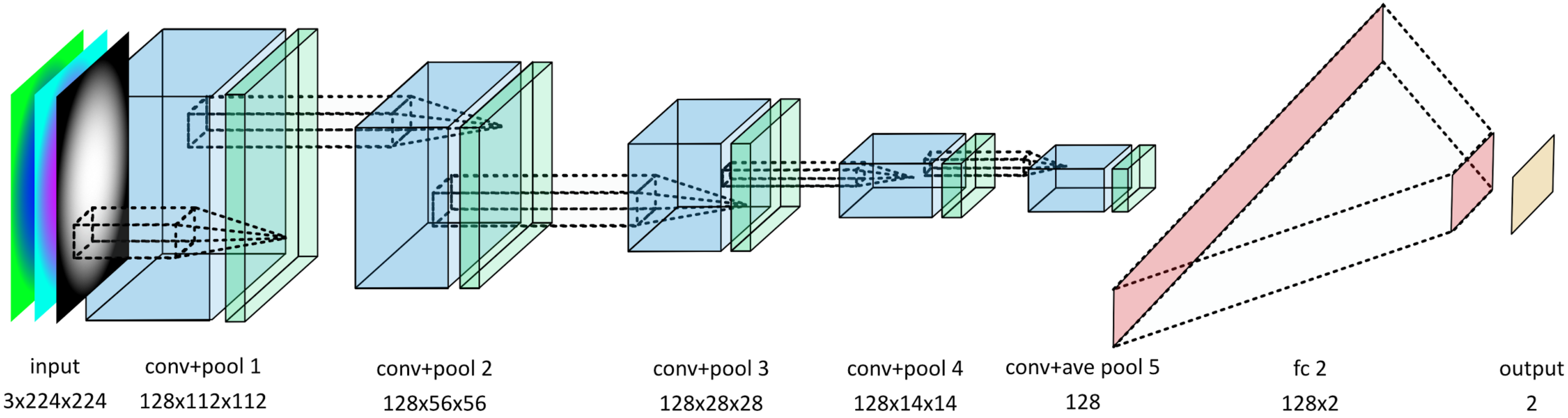
- Find rare signals of new particles in collisions (events) at LHC
- Represent data from cylindrical detector as sparse 2D (224x224) image
- 3 channels for different instruments (hadron & EM calorimeter, track multiplicity)
- ~10M event (7.4 TB) simulated dataset
- Existing selections on derived high-level physics variables used as a benchmark
- **binary classification**



atlas.ch



HEP Network Architecture

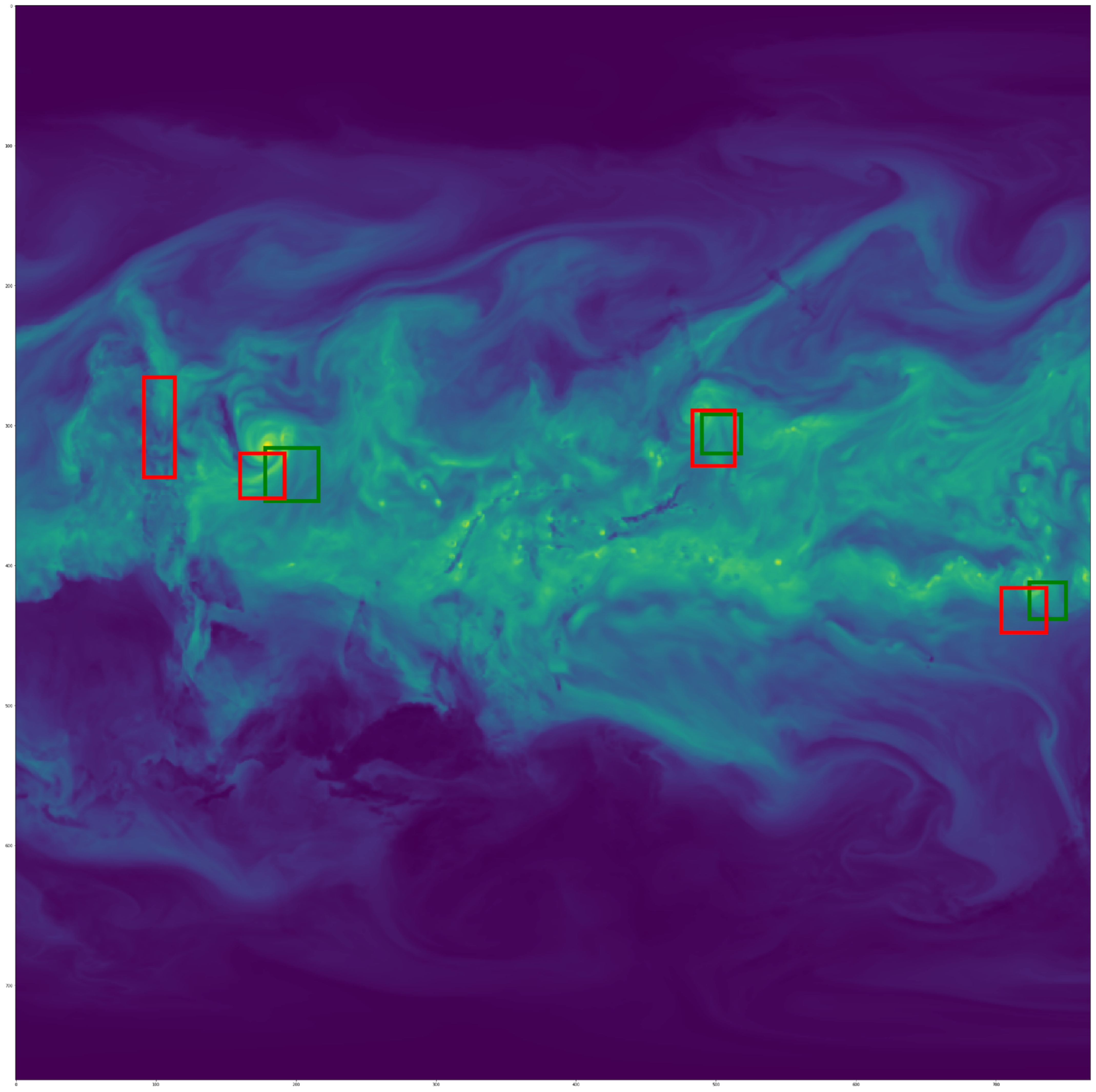


- sparse/lightweight layers
- 3 channels, suitable image dimensions
- total model size: ~2.3 MB
- training: SGD + momentum, per-layer LR, weight decay

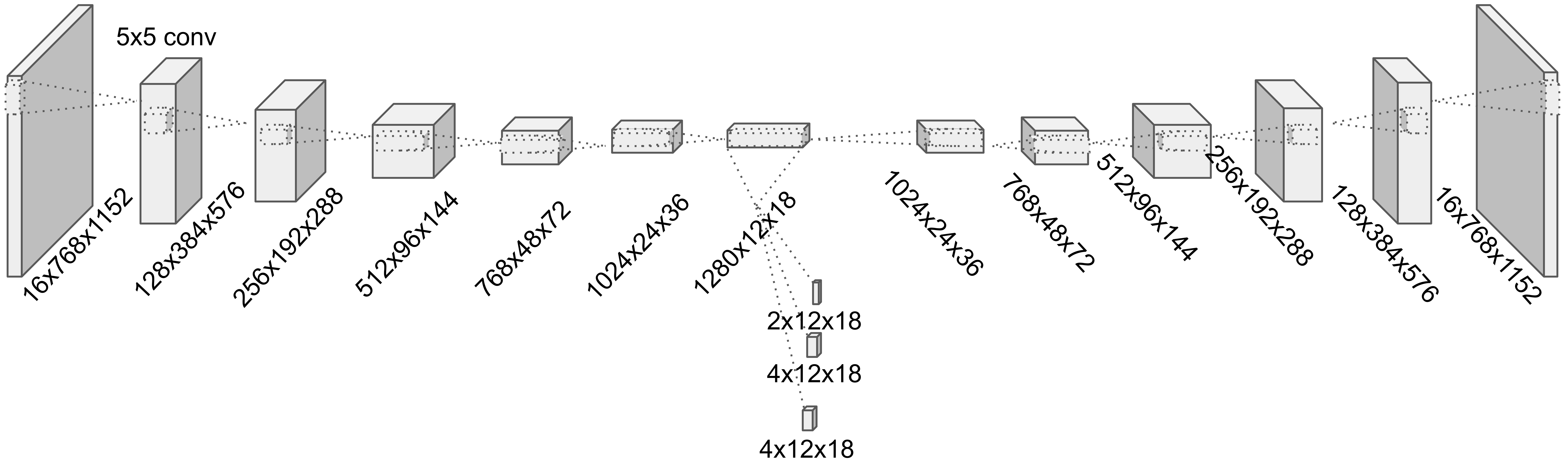
Climate Science

Finding Hurricanes

- locate and identify extreme weather phenomena in simulated climate data (CAM 25km resolution, 30 years)
- image size 768x1156 cropped to 768x768
- 16 channels (temperature, wind speed, pressure, etc.)
- not all images are annotated
- ~400K images (15 TB) data
- **semi-supervised bounding box regression+classification (7 classes)**



Climate Network Architecture

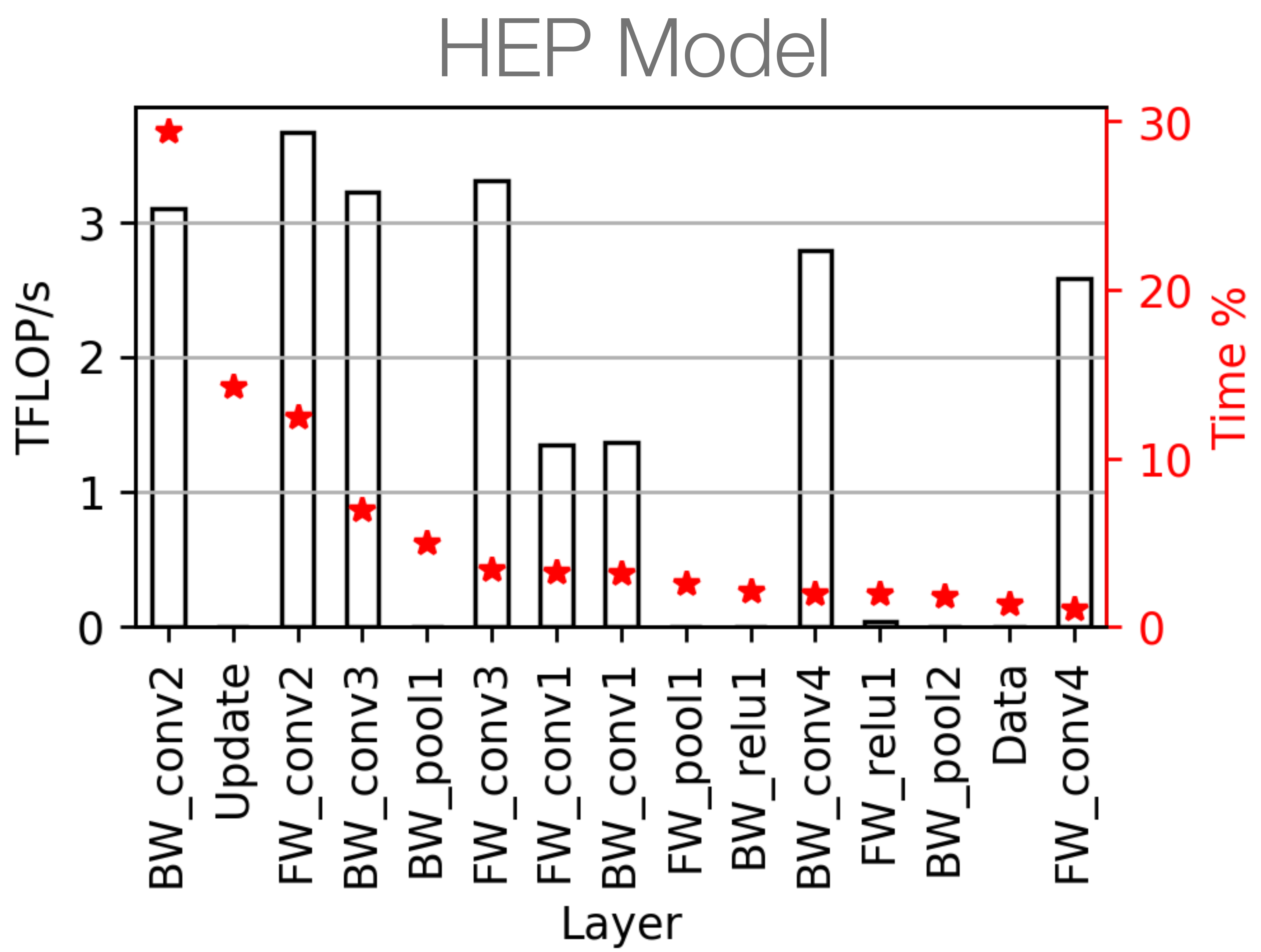


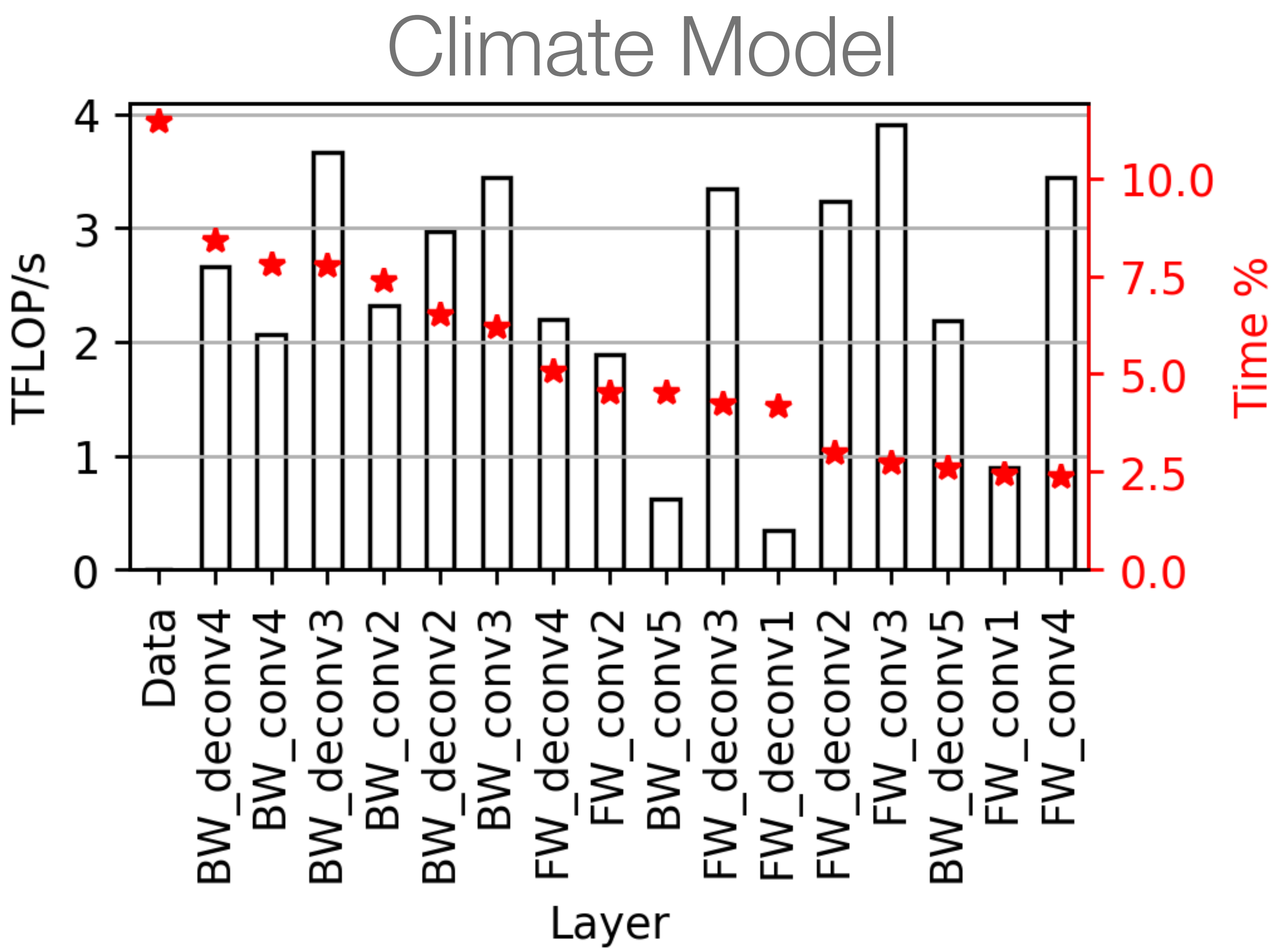
- only sparse layers
- regression/detection performed using YOLO (DOI: 10.1109/CVPR.2016.91)
- yolo-loss combined with euclidian L2 loss from autoencoder
- training: SGD + momentum/ADAM, weight decay

Software Stack and Hardware

- Intel® Distribution of Caffe
  - fast DL framework with distributed computing support
- Intel® Machine Learning Scaling Library (MLSL)
  - communication abstraction layer
  - here: use cray-mpich as backend with RDMA optimizations
- Intel® Math Kernel Library (Intel® MKL)
  - library with highly optimized DL primitives (soon to be replaced/merged with Intel® Math Kernel Library for Deep Neural Networks (Intel® MKL-DNN))
- Cori-KNL HPC system
  - 9688 Intel® Xeon Phi™ 7250 processor nodes (Knight's Landing)
  - 90 GB DDR + 16 GB MCDRAM memory per node
  - 68 cores with 272 threads, 1KB L1/core, 1MB L2 / 2 cores
  - high speed Aries interconnect w/ dragonfly topology

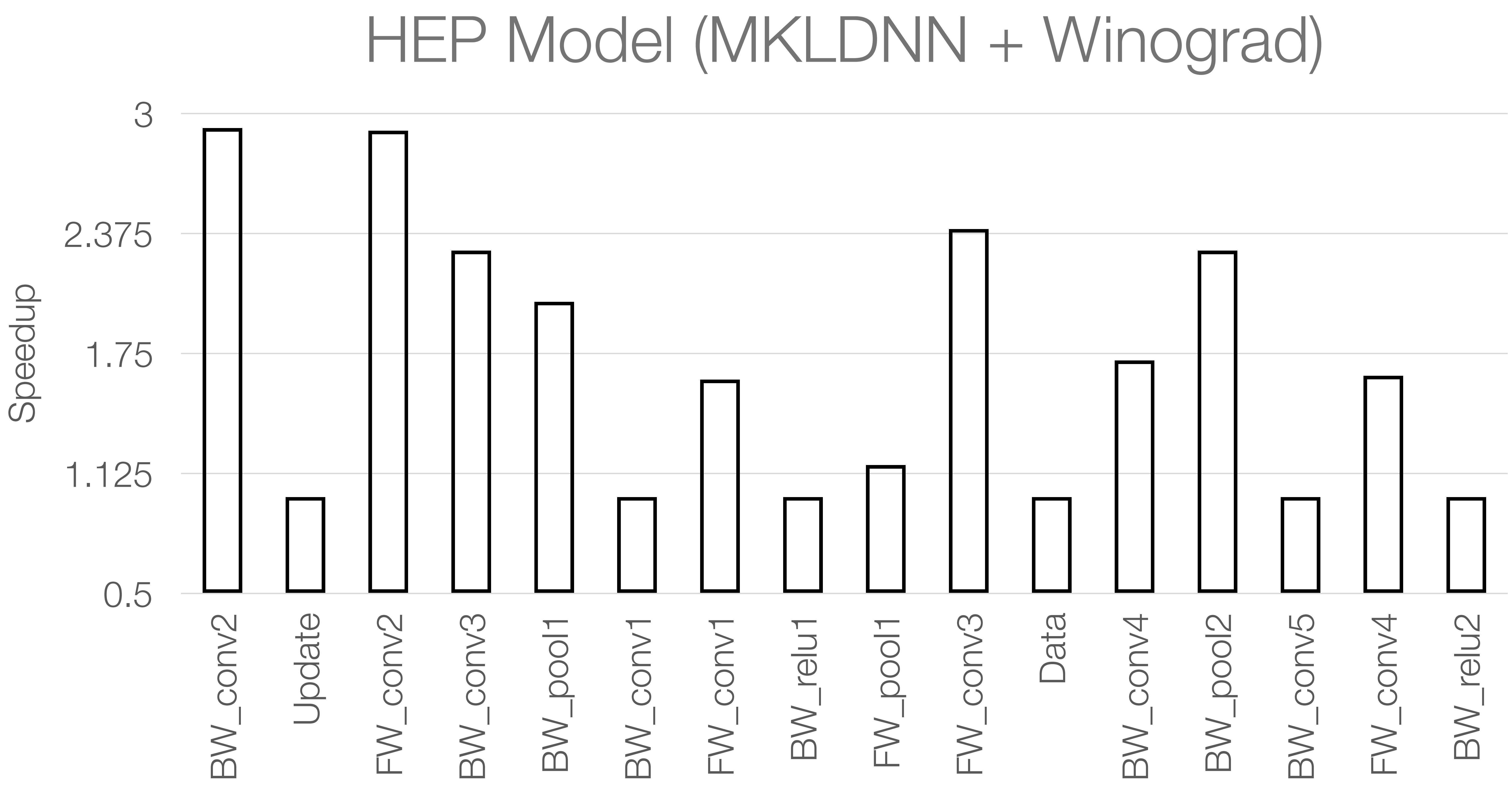
Single Node Performance





- local batch size of 8
- HEP: 1.90 TFLOP/s/node
- Climate: 2.09 TFLOP/s/node

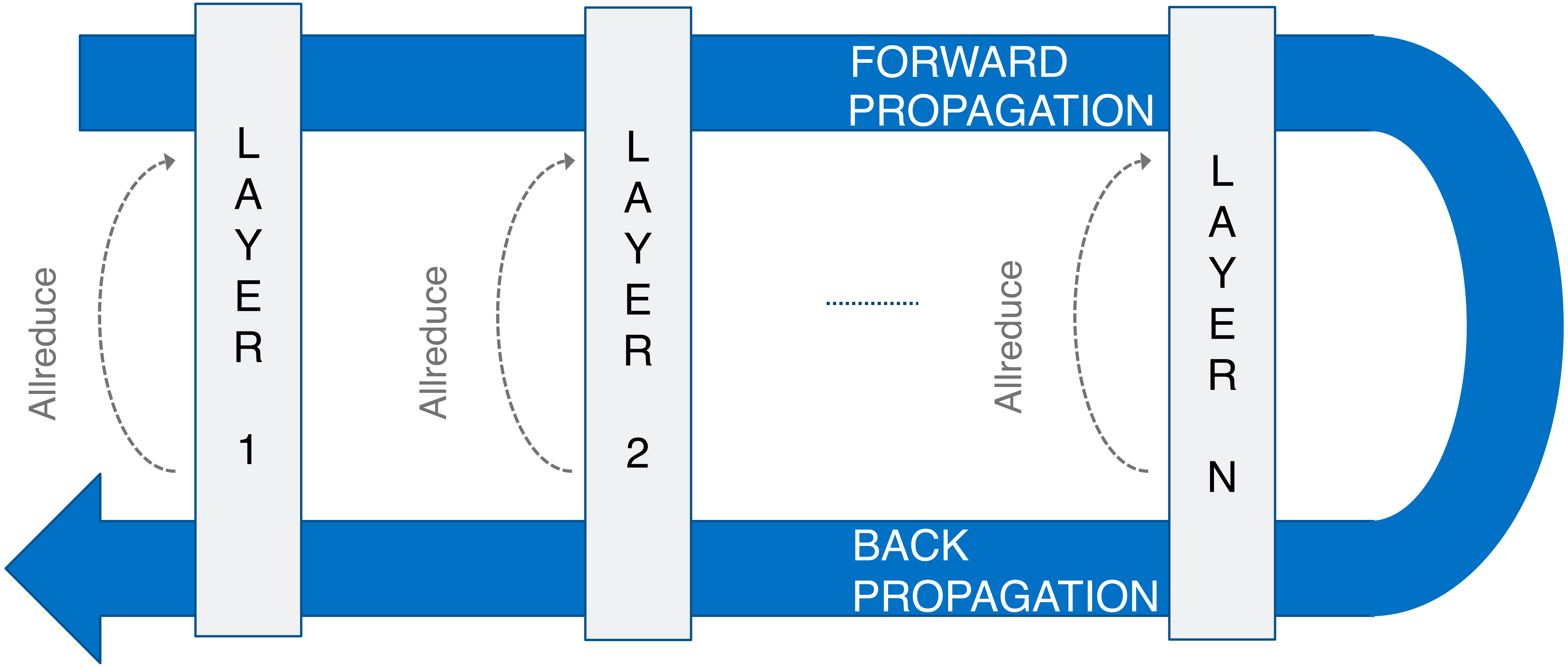
Single Node Performance Speedup using latest libraries



- local batch size of 8
- HEP: 1.90 TFLOP/s/node in the paper
  -> improved to 3.21 TFLOP/s/node w/ Winograd optimizations in Intel® MKL-DNN

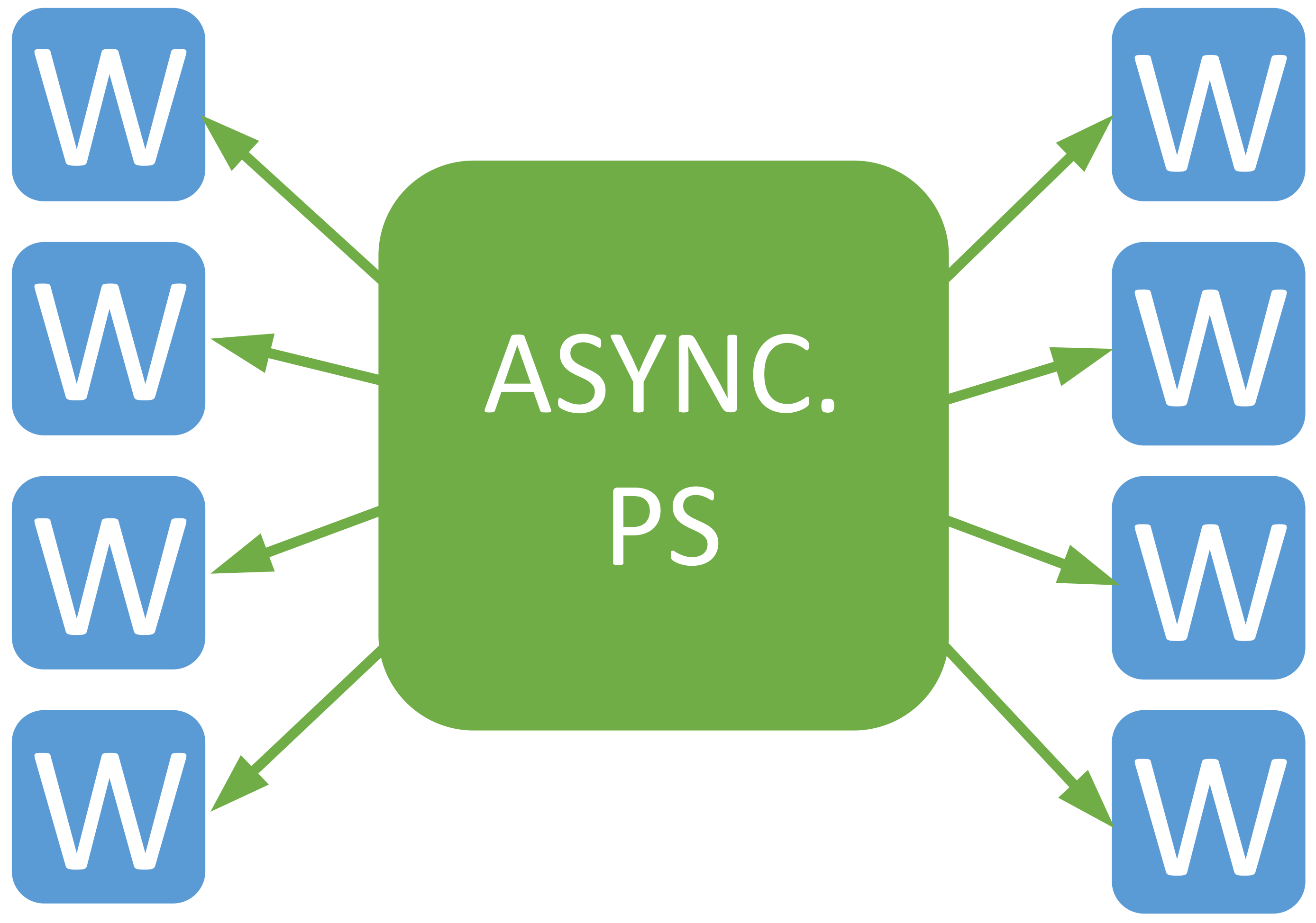
Multi Node Parallelization Scheme

- Use data parallelism for the Solver
  - Each node takes part of the data and computes model updates independently without communication
  - These updates are then collectively applied to the model



From Pradeep Dubey, "Scaling to Meet the Growing Needs of Artificial Intelligence (AI), IDF 2016
https://software.intel.com/en-us/articles/scaling-to-meet-the-growing-needs-of-ai

Synchronous and Asynchronous Update Strategies



SYNCHRONOUS

- understood convergence
- nodes have to wait for reduction to complete (stragglers slow everyone down)
- global (effective) batch size grows with number of nodes

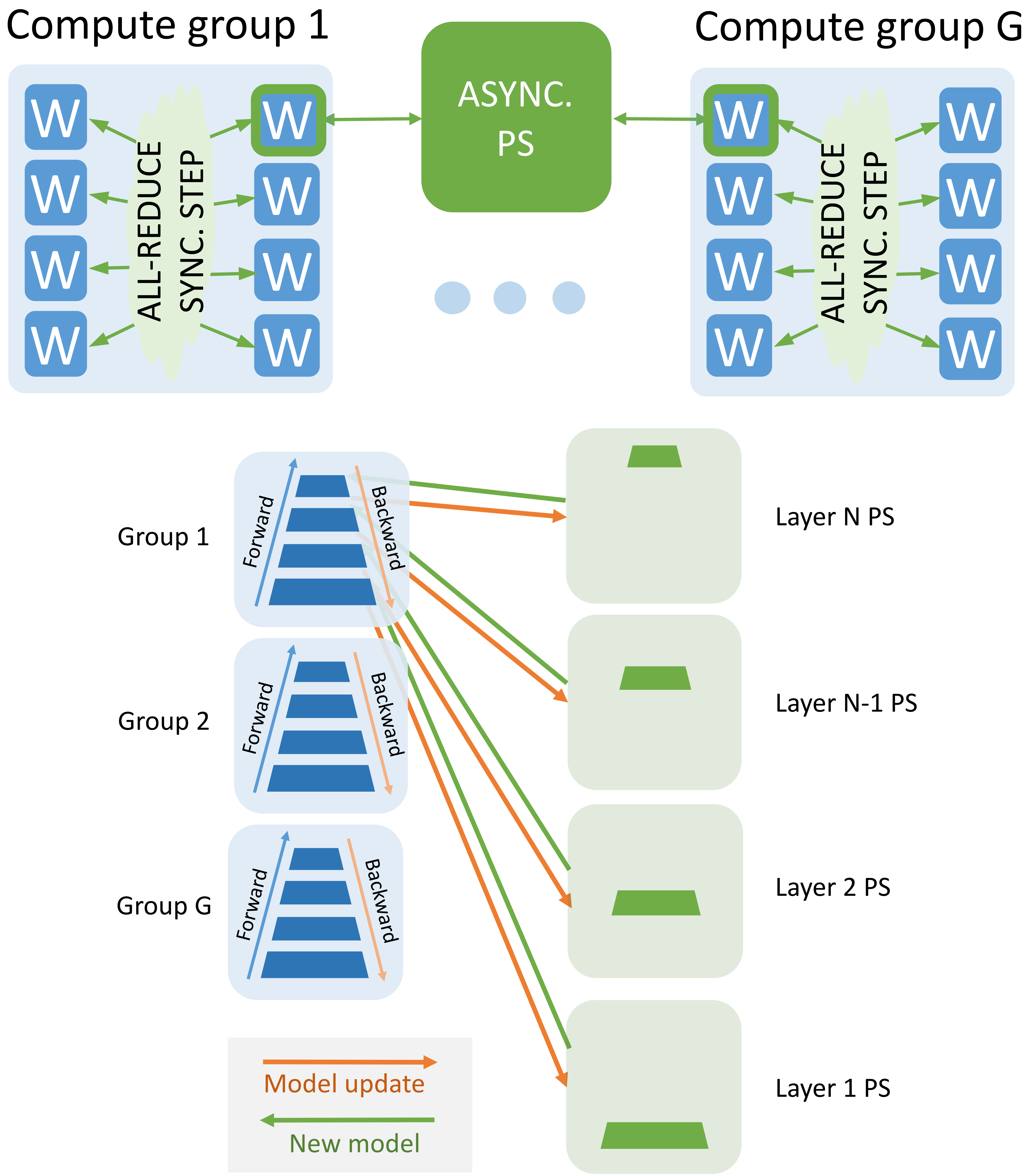
ASYNCHRONOUS

- no node waits for anybody
- resilient
- stale gradients can have impact on convergence rate
- parameter server can be bottleneck

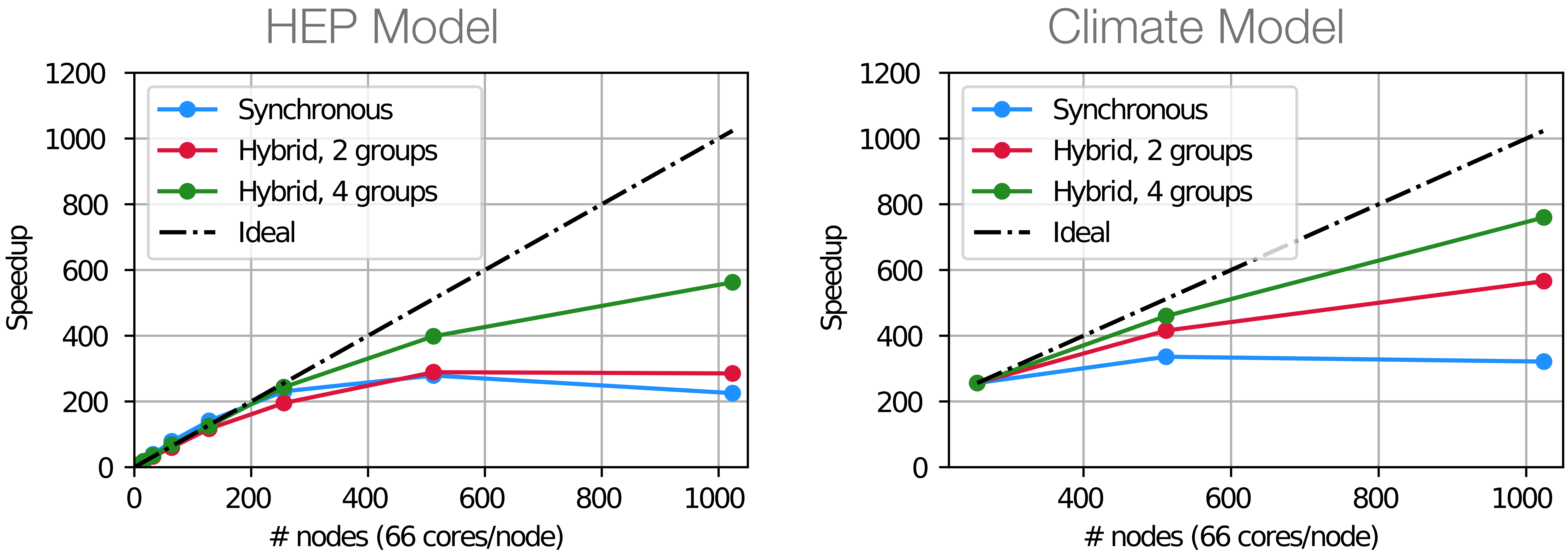
Hybrid Update

- impact of stragglers reduced compared to fully synchronous mode
- negative impact on stochastic convergence controlled
- finer control on total batch size
- group size needs to be tuned

Hadjis et.al., Omnivore, 2016

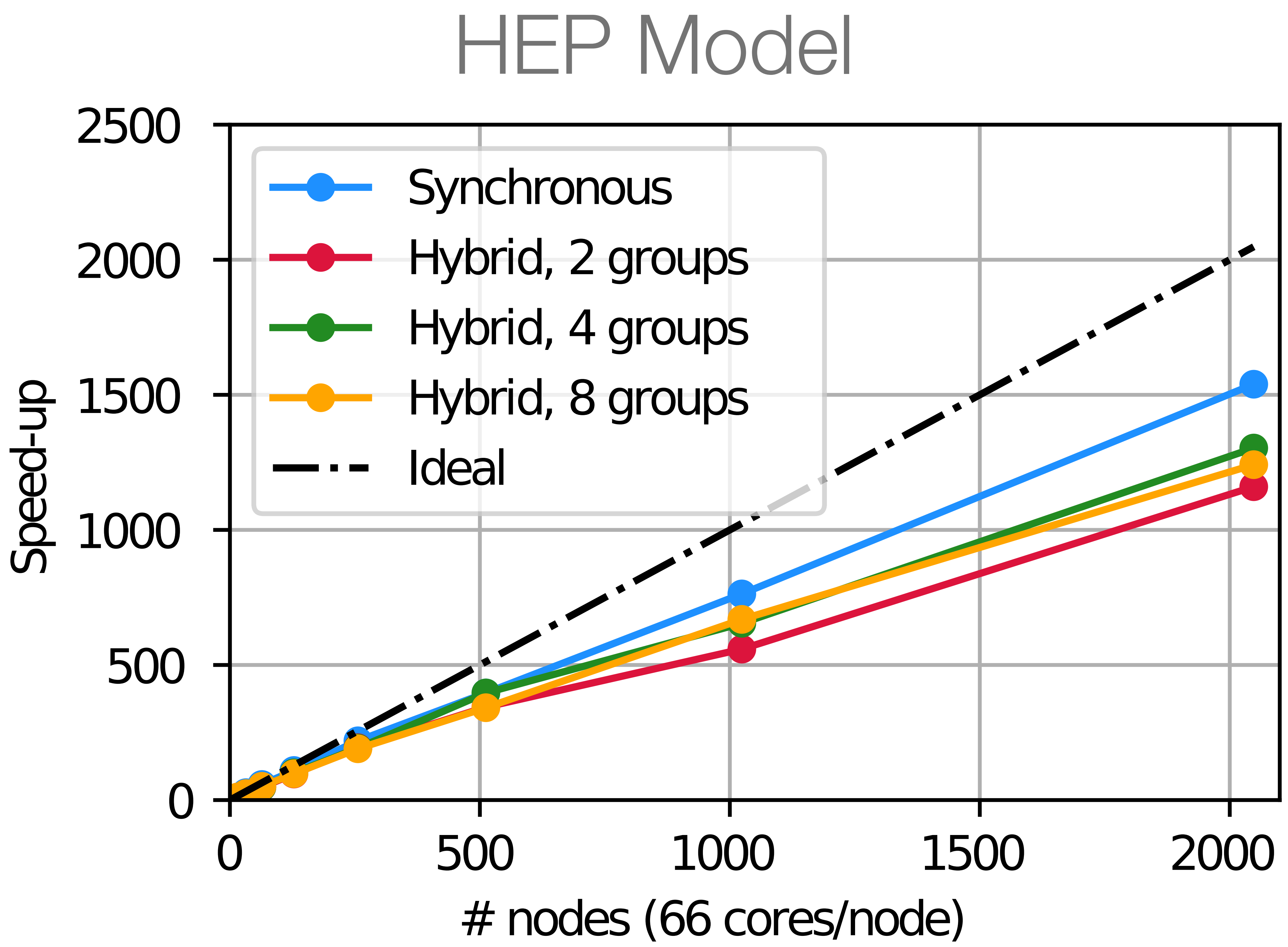


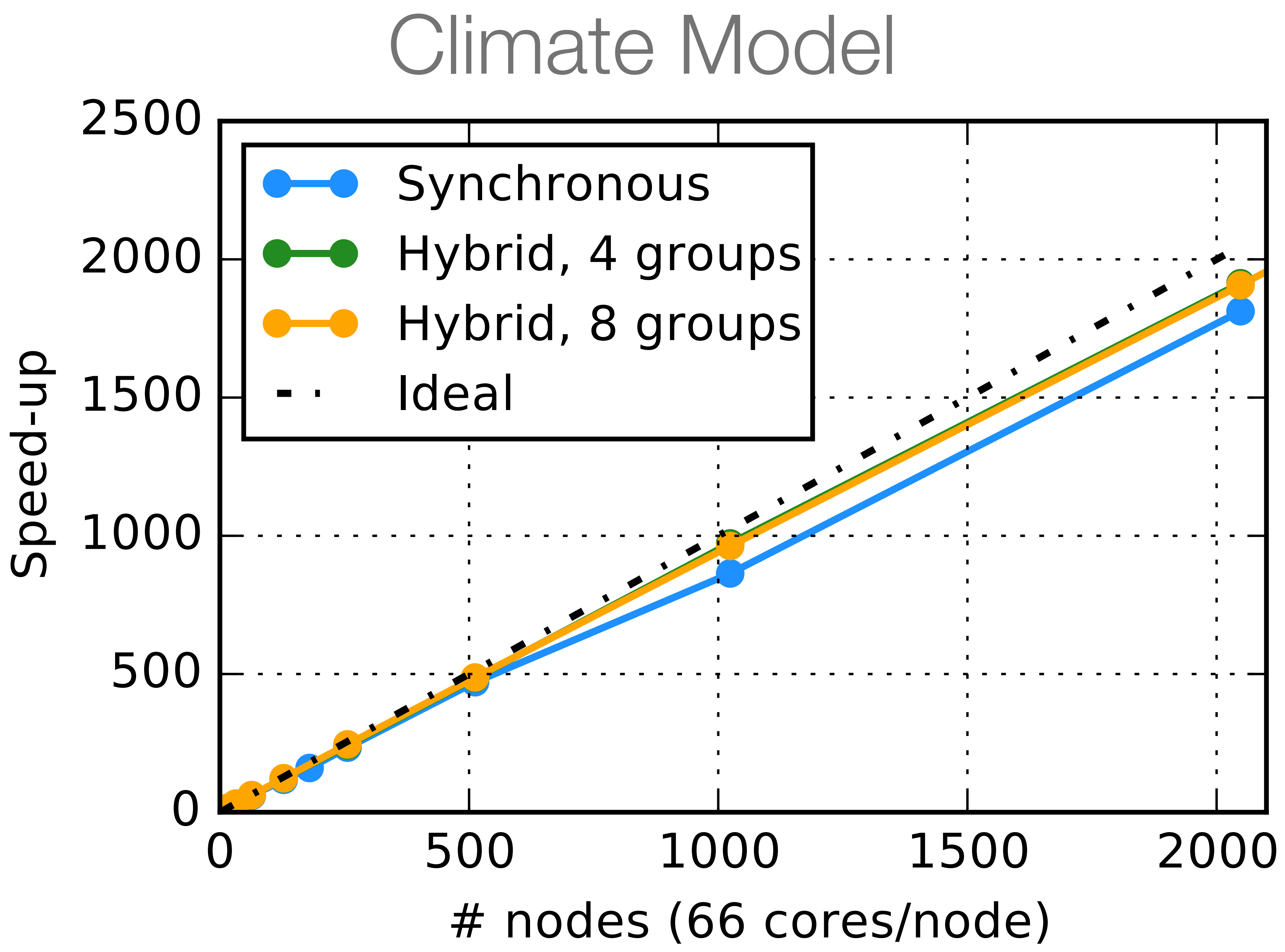
Strong Scaling Results



- batch size 2048 per group
- bad strong scaling beyond 512 nodes
- climate model better because compute/communication ratio is better

Weak Scaling Results

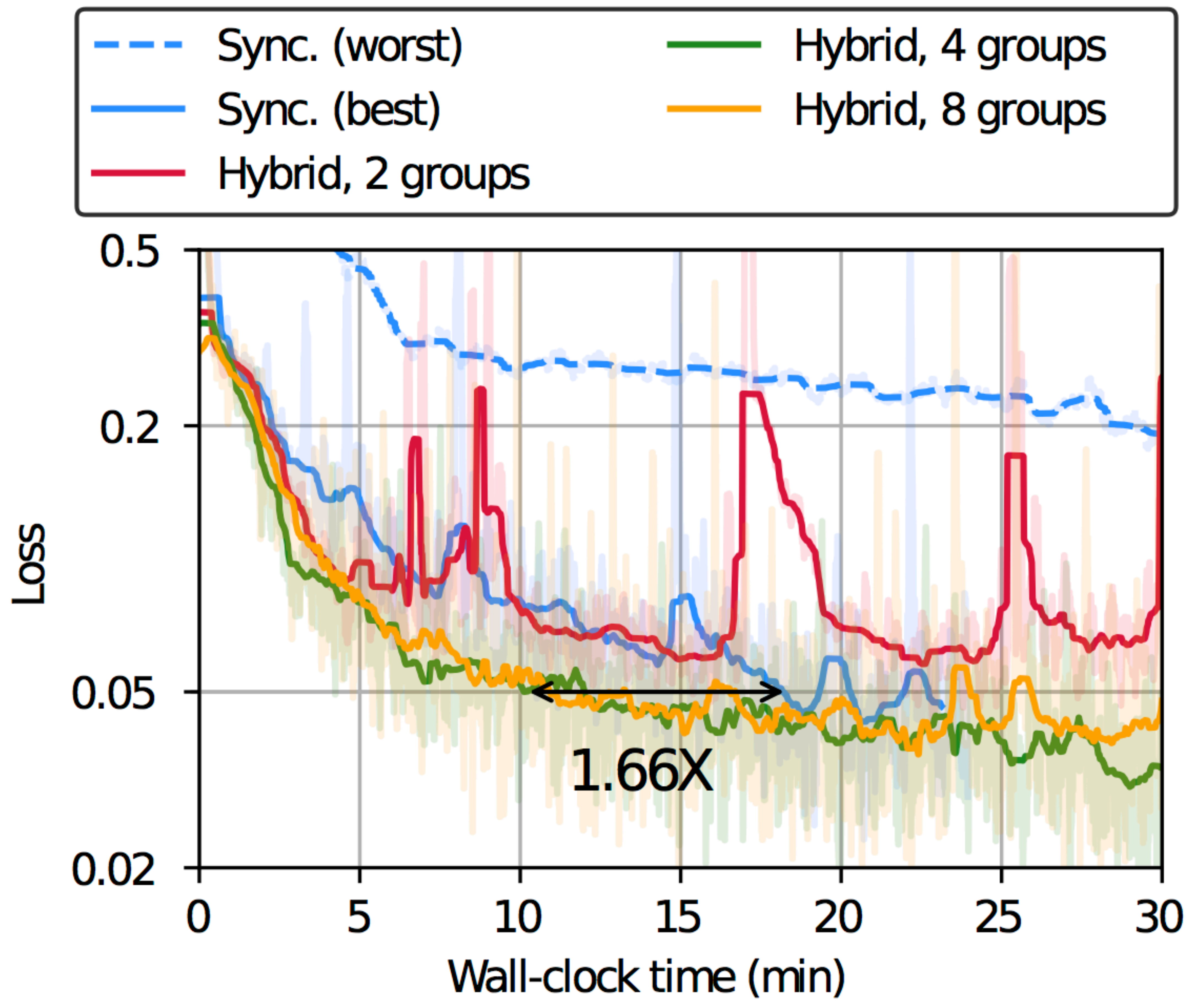




- batch size 8 per node
- good weak scaling properties
- climate network shows almost ideal scaling

Impact of Asynchronous Updates

- asynchronous groups
  - decrease time per update
  - increase number of iterations to reach the same loss
- plot shows that performance can still be gained for reasonable number of async groups
- much better improvement over worst synchronous case (improved resiliency)

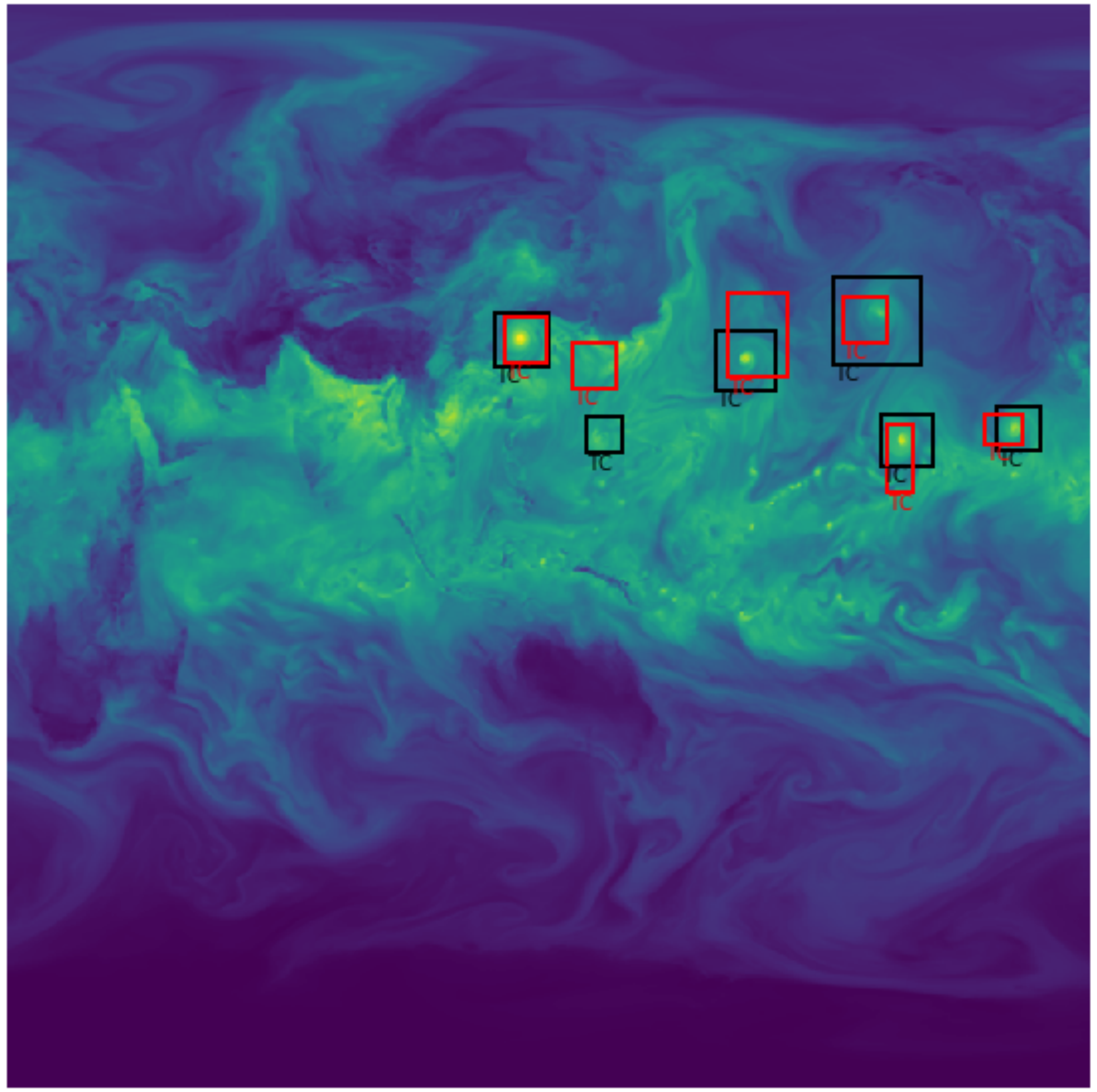


Full System Scale

- HEP Model
  - 9594 worker nodes + 6 parameters servers
  - 9 groups, 8528 examples/group
  - **11.4 PFLOP/s sustained, 11.7 PFLOP/s peak performance**
  - **1.3x improvement over baseline in ~12 minutes**
    (improvement in signal (true positive rate) over standard selection cuts for given background rejection (false positive rate)
- Climate Model
  - 9608 worker nodes + 14 parameters servers
  - 8 groups, 9608 examples/group
  - **13.3 PFLOP/s sustained, 15.1 PFLOP/s peak performance**

Summary

- presented large scale deep learning results for two selected science problems
- combination of synchronous and asynchronous algorithms, on-node optimizations and tuning network topology essential
- deep learning is well suited for large scale HPC systems
- Hyperparameter optimization at scale is difficult



Thank you