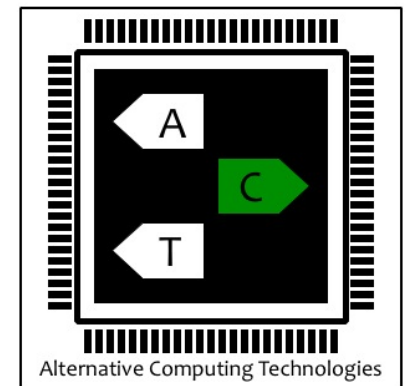


Scale-Out Acceleration for Machine Learning

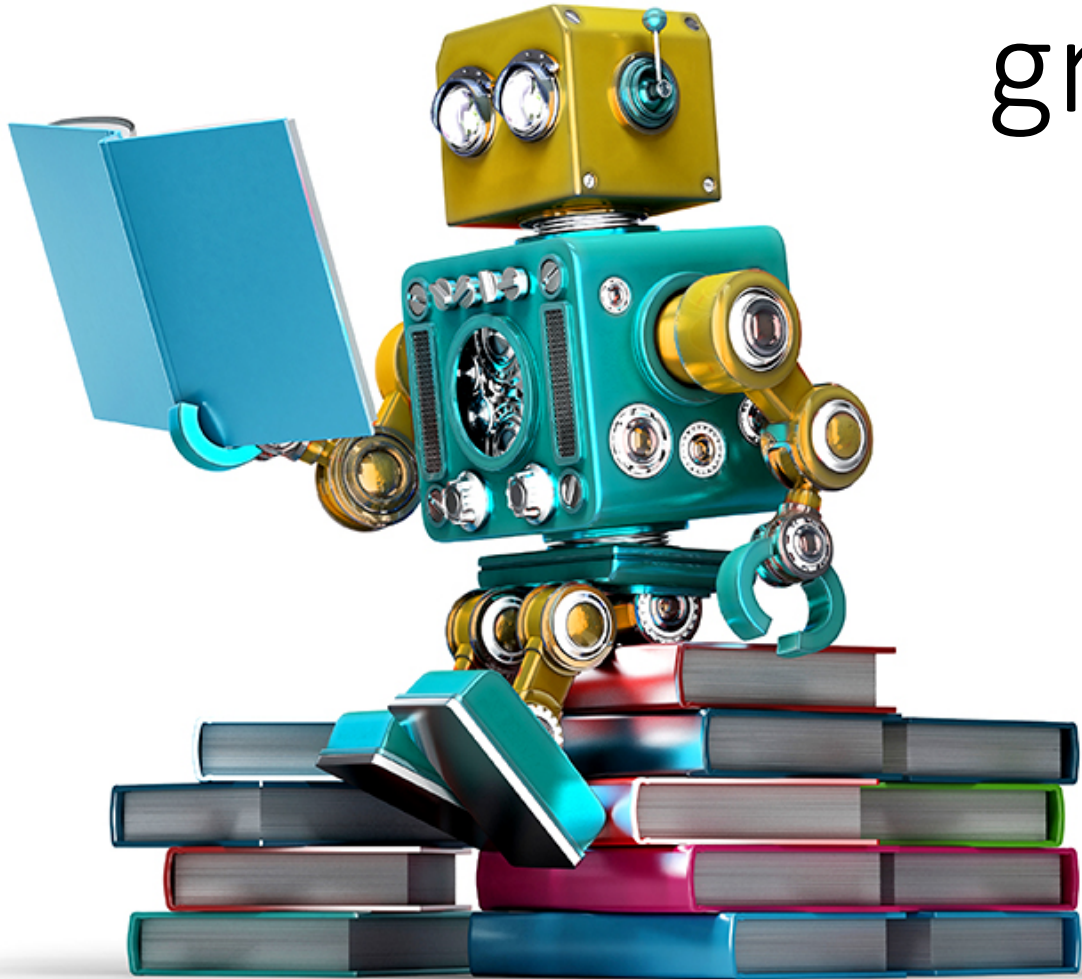
Jongse Park

Hardik Sharma
Divya Mahajan
Joon Kyung Kim
Preston Olds
Hadi Esmaeilzadeh

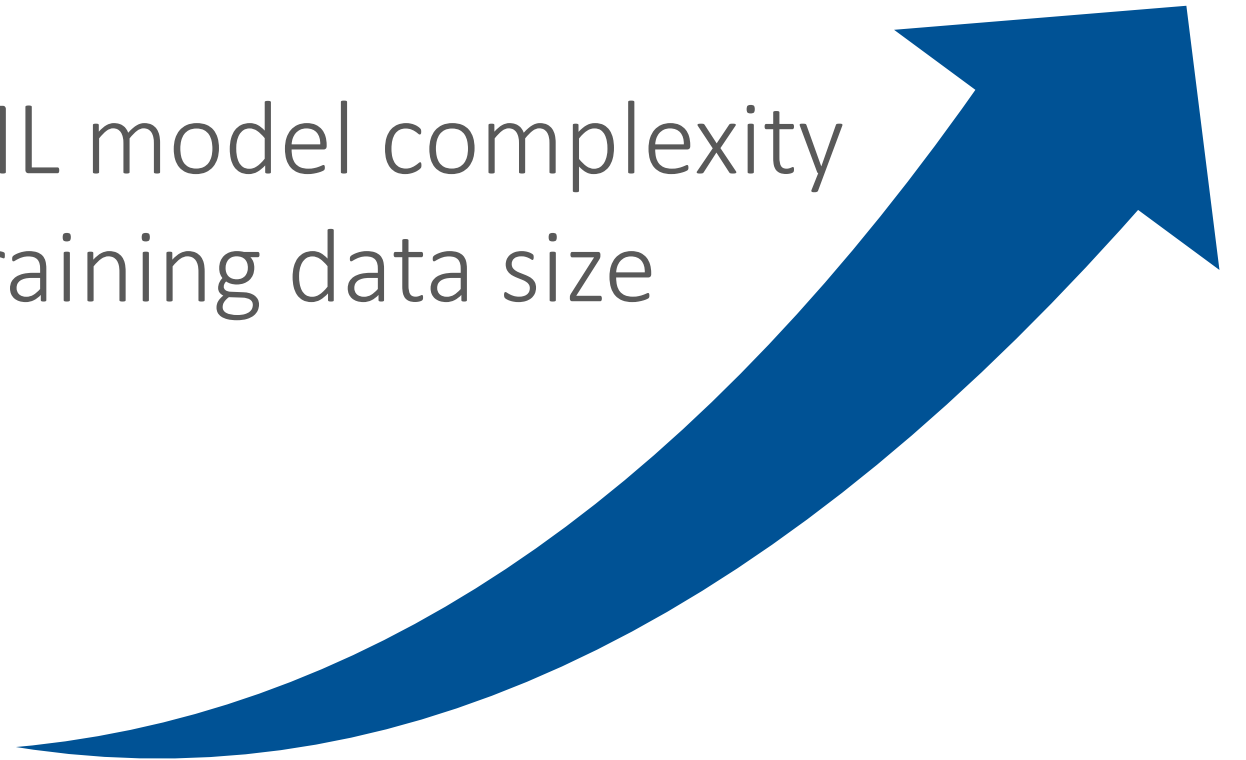
Alternative Computing Technologies (**ACT**) Lab
Georgia Institute of Technology
University of California, San Diego



Machine learning is growing exponentially!

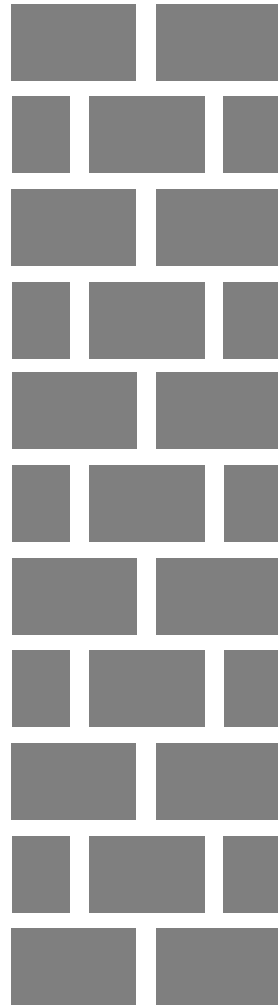


ML model complexity
Training data size

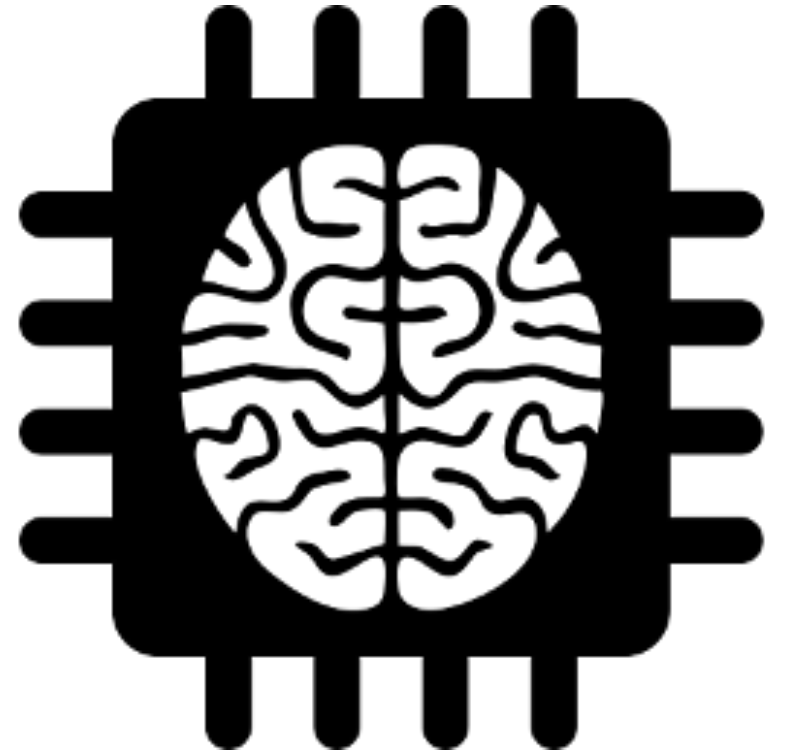


Two Disjoint Solutions

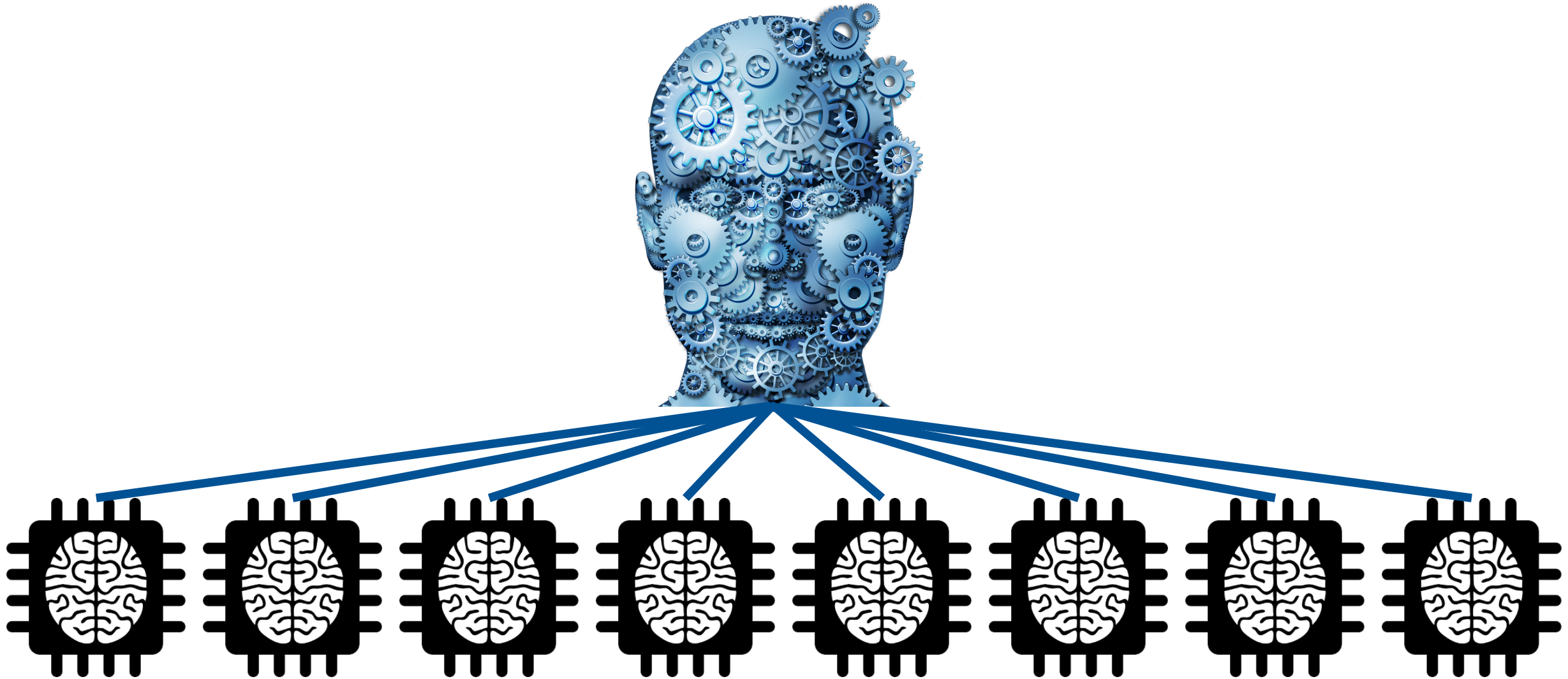
Distributed computing



Accelerators

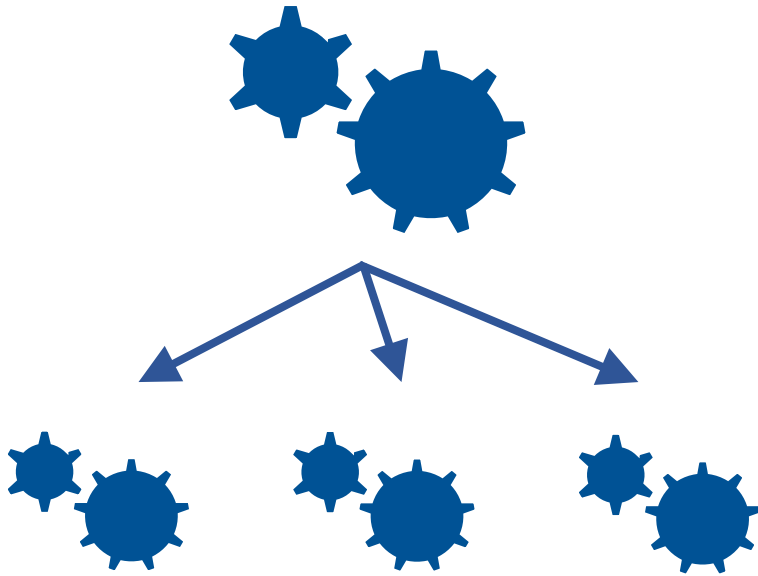


Scale-Out Acceleration for ML

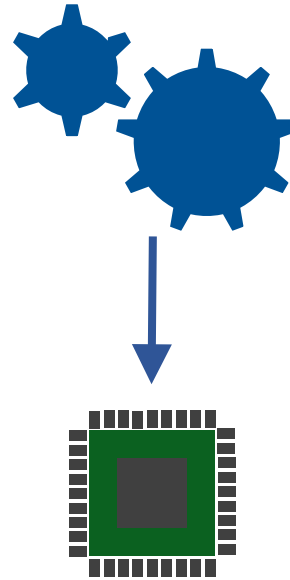


Research Challenges

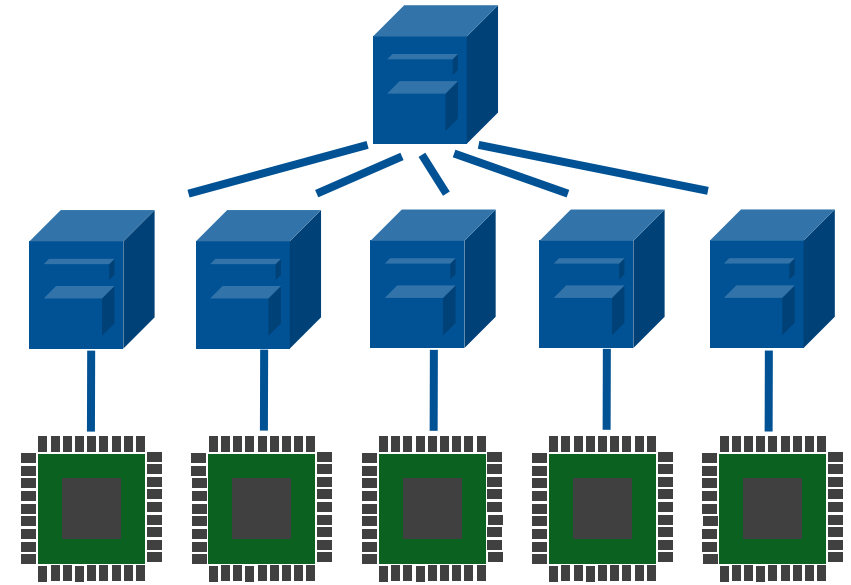
How to distribute?



How to customize?



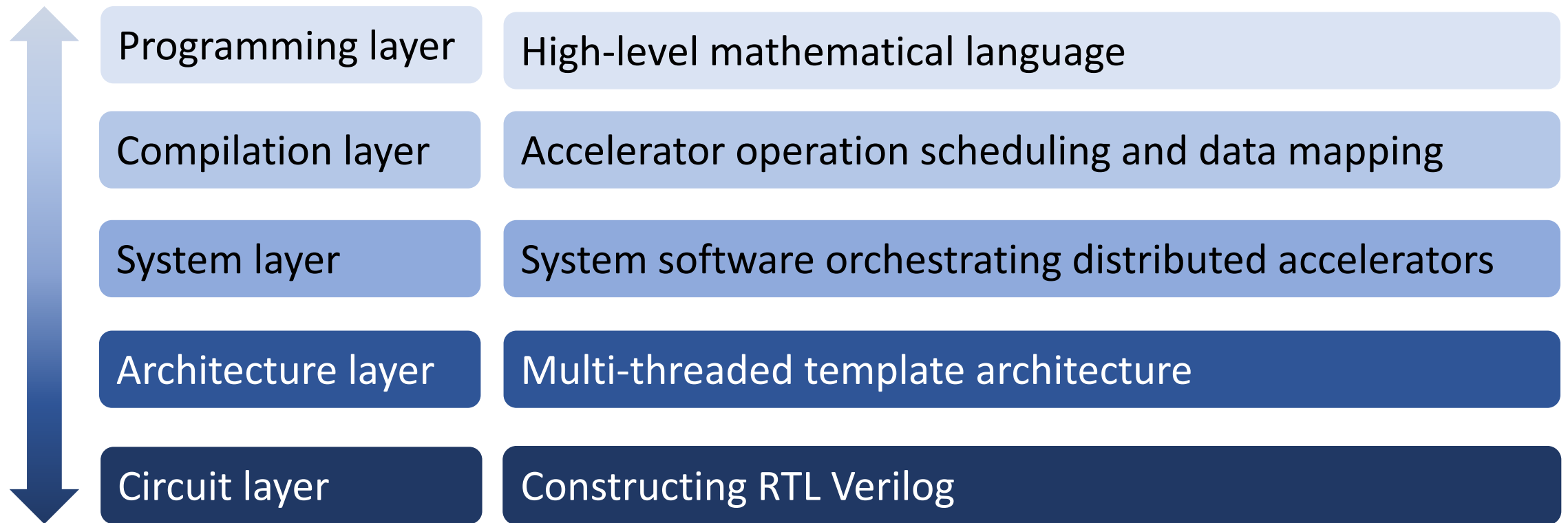
How to coordinate?



We need a full stack!

CoSMIC

Computing Stack for ML Acceleration In the Cloud





Session 4-A
At 2:30PM
Ballroom ABC

Come and learn about CoSMIC
phenomenon for ML!