

# HPC user tasks (weather prediction, search, deep learning based algorithms ...)

## HPC services

User front-end (cloud, GRID, supercomputer, etc)

Existing frameworks / algorithms

## Single nodes

Algorithm / source code

Available libraries

Compilers

Binary or byte code

Inputs

Various models

Run-time environment

Run-time state of the system

Hardware, simulators

Google Cloud, Microsoft Azure, AWS, XSEDE, PRACE, Watson...

TensorFlow, Caffe, Torch, Theano, TensorRT, CNTK, OpenCV ...

CUDA, MPI, OpenMP, TBB, OpenCL, StarPU, OmpSs ...

C, C++, Fortran, Java, Python, byte code, assembler ...

LLVM, GCC, ICC, Rose, PGI, (hundreds of optimizations) ...

cuBLAS, BLAS, MAGMA, ViennaCL, CLBlast, cuDNN, openBLAS, clBLAS, libDNN, tinyDNN, ARM compute lib, Intel libxsmm ...

diverse hardware: heterogeneous, many-core, out-of-order, reconfigurable (x86, ARM, PTX, AI/vector extensions) ...

Linux (CentOS, Ubuntu, RedHat, SUSE, Debian), BSD  
Windows, Android, iOS, MacOS ...

Researchers need to improve accuracy, execution time, energy/resource usage and other costs