AMD Scientific Libraries

Nicholas Malaya, René van Oostrum, Noel Chalmers, Damon McDougall, Paul Bauman, Nicholas Curtis, Noah Wolfe, Chip Freitag

Frontier Application Readiness Kick-Off Workshop

October 10th, 2019

AGENDA

Current state of AMD Libraries

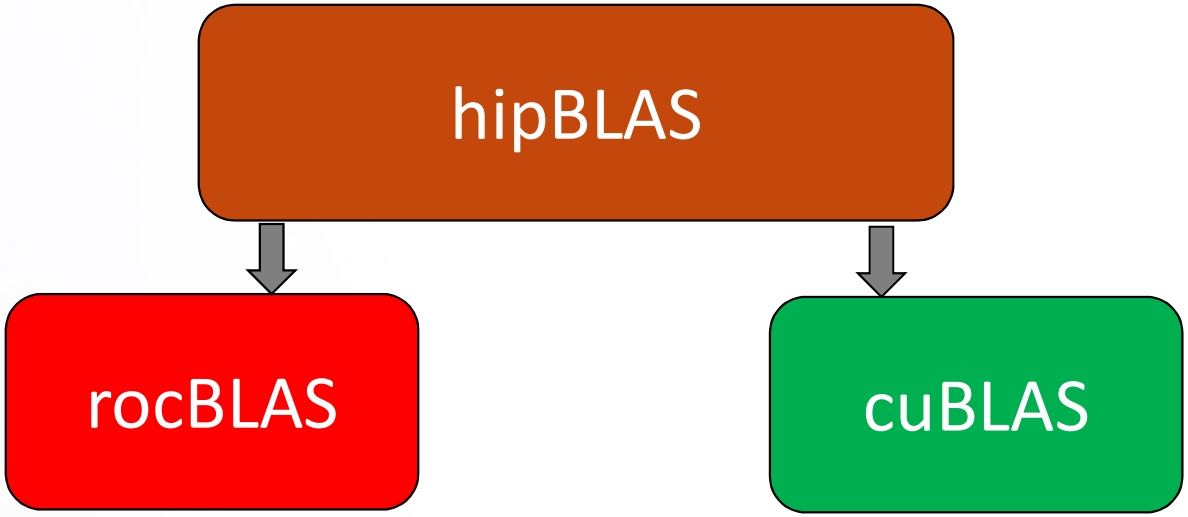
Mapping of CUDA to HIP

Future of AMD Libraries

What is missing, what do you need to succeed?

AMD GPU Library Organizational Pattern

- AMD ROCm has two classifications of libraries
  - roc* -> AMD GPU library, usually written in HIP
  - hip* -> Thin interface between AMD roc* and Nvidia cu* library
    - “Marshalling libraries”, “Shim libraries”
  - Targeting both CUDA and AMD devices: use the hip* libraries
  - Targeting only AMD devices, consider the roc* libraries



AMD Math Library Equivalents: “Decoder Ring” (1/2)

CUBLAS	**ROCBLAS**	Basic Linear Algebra Subroutines
CUFFT	**ROCFFT**	Fast Fourier Transforms
THRUST	**ROCTHRUST**	C++ Parallel Algorithms
CUB	**ROCPRIM**	Optimized Parallel Primitives
EIGEN	**EIGEN**	C++ Template Library for Linear Algebra

GITHUB.COM/ROCM-DEVELOPER-TOOLS/HIP → HIP_PORTING_GUIDE.MD FOR A COMPLETE LIST

AMD Math Library Equivalents: “Decoder Ring” (2/2)

CUSPARSE	**ROCSPARSE**	Sparse BLAS, SpMV, etc.
CUSOLVER	**ROCALUTION**	Solvers and preconditioners for GMG & AMG
CURAND	**ROCRAND**	Random number generators

GITHUB.COM/ROCM-DEVELOPER-TOOLS/HIP → HIP_PORTING_GUIDE.MD FOR A COMPLETE LIST

AMD Machine Learning Library Support

Machine Learning Frameworks:

- Tensorflow: https://github.com/ROCmSoftwarePlatform/tensorflow-upstream
- Pytorch: https://github.com/ROCmSoftwarePlatform/pytorch
- Caffe

Machine Learning Libraries:

- MIOpen (similar to cuDNN): https://github.com/ROCmSoftwarePlatform/MIOpen
- Tensile (_GEMM Autotuner): https://github.com/ROCmSoftwarePlatform/Tensile
- RCCL (ROCm analogue of NCCL): https://github.com/ROCmSoftwarePlatform/rccl
- Horovod (Distributed ML): https://github.com/ROCmSoftwarePlatform/horovod

Benchmarks:

- DeepBench: https://github.com/ROCmSoftwarePlatform/DeepBench
- MLPerf: https://mlperf.org

Missing

- Batched linear algebra:
  - dgetrf_batched, dgeqrf_batched, Xtrmm_batched, etc.
- Complex types:
  - zgetrf, zgetrs, etc.
- Portability Abstractions:
  - Kokkos, RAJA, etc.

Library availability is the first step
Expect tuned libraries


Requests?
What do you need?

Wrap up

Questions?

<nicholas.malaya@amd.com>

DISCLAIMER

DISCLAIMER
The information contained herein is for informational purposes only and is subject to change without notice. While every precaution has been taken in the preparation of this document, it may contain technical inaccuracies, omissions and typographical errors, and AMD is under no obligation to update or otherwise correct this information. Advanced Micro Devices, Inc. makes no representations or warranties with respect to the accuracy or completeness of the contents of this document, and assumes no liability of any kind, including the implied warranties of noninfringement, merchantability or fitness for particular purposes, with respect to the operation or use of AMD hardware, software or other products described herein. No license, including implied or arising by estoppel, to any intellectual property rights is granted by this document. Terms and limitations applicable to the purchase or use of AMD's products are as set forth in a signed agreement between the parties or in AMD's Standard Terms and Conditions of Sale. GD-18

DATA RIGHTS NOTIFICATION

(a) These data are submitted with limited rights under Government Contract No. DE-AC05-00OR22725 and DE-AC52-07NA27344 Subcontract No. 4000166089 and B626588. These data may be reproduced and used by the Government with the express limitation that they will not, without written permission of the Contractor, be used for purposes of manufacture nor disclosed outside the Government; except that the Government may disclose these data outside the Government for the following purposes, if any; provided that the Government makes such disclosure subject to prohibition against further use and disclosure: A. This "Limited Rights Data" may be disclosed for evaluation purposes under the restriction that the "Limited Rights Data" be retained in confidence and not be further disclosed; B. This "Limited Rights Data" may be disclosed to other contractors participating in the Government's program of which this Purchase Order is a part for information or use in connection with the work performed under their contracts and under the restriction that the "Limited Rights Data" be retained in confidence and not be further disclosed; and C. This "Limited Rights Data" may be used by the Government or others on its behalf for emergency repair or overhaul work under the restriction that the "Limited Rights Data" be retained in confidence and not be further disclosed.
(b) This notice shall be marked on any reproduction of these data, in whole or in part.

Key AMD Libraries

- BLAS
  - rocBLAS (https://github.com/ROCmSoftwarePlatform/rocBLAS)
  - hipBLAS (https://github.com/ROCmSoftwarePlatform/hipBLAS)
- FFTs
  - rocFFT (https://github.com/ROCmSoftwarePlatform/rocFFT)
- Random number generation
  - rocRAND (https://github.com/ROCmSoftwarePlatform/rocRAND)
  - hipRAND (https://github.com/ROCmSoftwarePlatform/hipRAND)
- Sparse linear algebra
  - rocSPARSE (https://github.com/ROCmSoftwarePlatform/rocSPARSE)
  - hipSPARSE (https://github.com/ROCmSoftwarePlatform/hipSPARSE)
- Iterative solvers
  - rocALUTION (https://github.com/ROCmSoftwarePlatform/rocALUTION)
- Parallel primitives
  - rocPRIM (https://github.com/ROCmSoftwarePlatform/rocPRIM)
  - hipCUB (https://github.com/ROCmSoftwarePlatform/hipCUB)

AMD Open Source Benchmarks

ROCHPCG	High performance CG
ROCALUTION	Solvers and preconditioners for GMG & AMG
ROCRAND	Random number generators

GITHUB.COM/ROCM-DEVELOPER-TOOLS/HIP → HIP_PORTING_GUIDE.MD FOR A COMPLETE LIST