# AMD

# **Accelerators in FLAMEs**

Johannes M Dieterich Strategic Leadership Optimization Team (HPC) Machine Learning Software Engineering Radeon Technologies Group Advanced Micro Devices, Inc.

2

#### **OLCF-5: Frontier**



- World's first exascale computer June 2022 TOP500 with 1.102 Exaflop/s
- Top of June 2022 GREEN500 at 62.68 Gigaflops/Watt
- 9,472 AMD EPYC<sup>™</sup> 64 core 2 GHz CPUs
- 37,888 AMD Instinct<sup>™</sup> MI250X accelerators
- 1 EPYC : 4 MI250X
- Places 1-4 on June 2022 GREEN500 are AMD EPYC<sup>™</sup> with AMD<sup>™</sup> MI250X systems.

Photo under CC BY 2.0 from <a href="https://www.flickr.com/photos/olcf/52117623843/">https://www.flickr.com/photos/olcf/52117623843/</a> TOP500: <a href="https://www.top500.org/lists/top500/2022/06/">https://www.top500.org/lists/top500/2022/06/</a> GREEN500: <a href="https://www.top500.org/lists/top500/2022/06/">https://www.top500.org/lists/top500/2022/06/</a>

#### **Peak performance differences**





#### 2.0 GHz 64 core AVX256 FMAs Double precision peak: approx. 1 TFLOPs

https://www.amd.com/en/processors/epyc-server-cpu-family https://www.amd.com/en/products/server-accelerators/instinct-mi250x Single AMD Instinct<sup>™</sup> MI250X Double precision peak: 47.9 TFLOPs Double precision matrix peak: 95.7 TFLOPs

#### Illustration of a CPU+GPU Heterogeneous System



#### **AMD ROCm<sup>™</sup> platform – open-source acceleration**



Full OSS kernel to application stack

\_\_global\_\_ void saxpy (unsigned N, float a, float \*x, float \*y) { unsigned i = blockIdx.x \* blockDim.x + threadIdx.x; if (i < N) y[i] = a\*x[i] + y[i];</pre>

HIP – SIMT programming model

## Acceleration (in a very small nutshell)

kernel

Accelerator

#### CPU

 Ensure high arithmetic intensity of kernel and ideal hardware use (e.g., occupancy, matrix extensions, ...)

## Acceleration (in a very small nutshell)



- Optimize data transfers / memory operations
- Ensure high arithmetic intensity of kernel and ideal hardware use (e.g., occupancy, matrix extensions, ...)
- Minimize host overhead
- Maximize asynchronous chain length

Execute kernels back-to-back

# libFLAME/FLASH – how accelerator programming fits into classic dense LA



Journal of Parallel and Distributed Computing Volume 72, Issue 9, September 2012, Pages 1134-1143



The FLAME approach: From dense linear algebra algorithms to high-performance multi-accelerator implementations

Francisco D. Igual <sup>a</sup>  $\stackrel{\otimes}{\sim}$   $\boxtimes$ , Ernie Chan <sup>b</sup>  $\boxtimes$ , Enrique S. Quintana-Ortí <sup>a</sup>  $\boxtimes$ , Gregorio Quintana-Ortí <sup>a</sup>  $\boxtimes$ , Robert A. van de Geijn <sup>b</sup>  $\boxtimes$ , Field G. Van Zee <sup>b</sup>  $\boxtimes$ 

- LAPACK-style algorithms by tiles as a DAG of dense matrix tasks
- Acceleration of BLAS-level tasks on (multiple) CUDA accelerators
- Discrete memory model tracking of "dirty" tiles and minimal traffic to/from discrete tile pool
- Synchronous memory operations

### 4x4 tile Cholesky decomposition on single accelerator (non-HIP)



#### **HIP backend features**

#### At compile time:

 --enable-supermatrix --enable-hip -enable-multithreading=pthread

#### At runtime:

- Choice of discrete and managed memory HIP backend
- Match #SuperMatrix threads to #accelerators
- Select tile size for SuperMatrix
- Select #tiles in device cache (discrete backend)
- Reads are synchronous
- FLA\_Obj host buffers are allocated pinned (discrete) or managed memory

#### Implemented FLAME tasks for all [s,d,c,z]:

- Through rocSOLVER:
  - Chol, Trinv, Eig\_gest, LU\_piv, LU\_piv\_copy
- Through rocBLAS:
  - Gemm, Hemm, Herk, Her2k, Symm, Syrk, Syr2k, Trmm, Trsm, Gemv, Trsv, Axpy, Scal, Scalr
- Through HIP APIs:
  - Copy, Copyr

# 4x4 tile Cholesky decomposition on single accelerator – discrete HIP backend



# 4x4 tile Cholesky decomposition on multiple accelerators – discrete HIP backend



A cross-device memory operation (synchronous, via host)

## 2x2 tile Lyapunov equation solver using a single accelerator



Managed memory HIP backend



## Work in progress

**General Improvements** 

Implement more FLAME operations for HIP

lapack2flash wrappers to access HIP acceleration automatically

#### **Multi-Accelerator**

Cross-device communication should not be synchronous via the host Use device peer2peer communication on AMD Infinity Fabric<sup>TM</sup> and dirtying of respective host blocks

#### **Copyright and disclaimer**

©2022 Advanced Micro Devices, Inc. All rights reserved.

AMD, the AMD Arrow logo, EPYC, Instinct, MI250X, Infinity Fabric, ROCm and combinations thereof are trademarks of Advanced Micro Devices, Inc. PCIe is a registered trademark of PCI-SIG Corporation. Other product names used in this publication are for identification purposes only and may be trademarks of their respective companies.

The information presented in this document is for informational purposes only and may contain technical inaccuracies, omissions, and typographical errors. The information contained herein is subject to change and may be rendered inaccurate releases, for many reasons, including but not limited to product and roadmap changes, component and motherboard version changes, new model and/or product differences between differing manufacturers, software changes, BIOS flashes, firmware upgrades, or the like. Any computer system has risks of security vulnerabilities that cannot be completely prevented or mitigated. AMD assumes no obligation to update or otherwise correct or revise this information. However, AMD reserves the right to revise this information and to make changes from time to time to the content hereof without obligation of AMD to notify any person of such revisions or changes.

THIS INFORMATION IS PROVIDED 'AS IS." AMD MAKES NO REPRESENTATIONS OR WARRANTIES WITH RESPECT TO THE CONTENTS HEREOF AND ASSUMES NO RESPONSIBILITY FOR ANY INACCURACIES, ERRORS, OR OMISSIONS THAT MAY APPEAR IN THIS INFORMATION. AMD SPECIFICALLY DISCLAIMS ANY IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY, OR FITNESS FOR ANY PARTICULAR PURPOSE. IN NO EVENT WILL AMD BE LIABLE TO ANY PERSON FOR ANY RELIANCE, DIRECT, INDIRECT, SPECIAL, OR OTHER CONSEQUENTIAL DAMAGES ARISING FROM THE USE OF ANY INFORMATION

#