SAND2018-11378R

ECP EXASCALE COMPUTING PROJECT

ECP-U-2018-XXX

# Kokkos back-ends research, collaborations, development, optimization, and documentation

## WBS STPR 04 Milestone 4

**Authors**

Christian Trott, Dan Sunderland, Mark Hoemmen (1)

**Author Affiliations**

(1) Sandia National Laboratories

**September 19th 2018**

U.S. DEPARTMENT OF ENERGY | Office of Science

NNSA National Nuclear Security Administration

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

ECP EXASCALE COMPUTING PROJECT

# EXECUTIVE SUMMARY

This report documents the completion of milestone STPR04-4 Kokkos back-ends research, collaborations, development, optimization, and documentation. The Kokkos team updated its existing backend to support the software stack and hardware of DOE's Sierra, Summit and Astra machines. They also collaborated with ECP PathForward vendors on developing backends for possible exa-scale architectures. Furthermore, the team ramped up its engagement with the ISO/C++ committee to accelerate the adoption of features important for the HPC community into the C++ standard.

U.S. DEPARTMENT OF ENERGY | Office of Science

NNSA National Nuclear Security Administration

# 1. MILESTONE OVERVIEW

## 1.1 DESCRIPTION

Kokkos back-ends research, collaborations, development, optimization, and documentation. Back-end development emphasizing migration to OpenMP4 and CUDA9, quality/sustainability, and performance on ATS-1/ATS-2/ARM. Back-end documentation and research collaborations emphasizing formal publication, promoting Kokkos abstractions for incorporation into ISO/C++ Standard, strategic interaction with vendors to prepare for evolving advanced architectures.

## 1.2 EXECUTION PLAN

There are multiple aspects to this milestone which are handled in different ways:

1) Track compiler and runtime releases and adopt changes as necessary and appropriate.
2) Take part in ECP PathForward meetings to keep track of new hardware developments and possible backend requirements.
3) Identify features in Kokkos appropriate for standardization and contribute to related standard proposals or develop new ones.

## 1.3 COMPLETION CRITERIA

Document work via a Milestone Report, in GitHub issues and through contributions to proposals before the ISO C++ committee.

# 2. TECHNICAL WORK SCOPE, APPROACH, RESULTS

## 2.1 BACK-END MODERNIZATION

As runtimes evolve Kokkos backends have to adopt changes. This FY one of the major changes requiring work inside Kokkos was the adoption of CUDA 9 with Volta support. In fact, the combination of CUDA 9 and Volta GPUs broke significant portions of our CUDA backend, notably anything related to reductions and synchronization mechanisms. The root cause for the issues is a change in how Volta GPUs work compared to previous NVIDIA GPUs. In previous GPUs only a single instruction pointer existed per Warp (group of 32 threads which are roughly equivalent to vector lanes on CPUs). In the Volta architecture this was changed. Now every CUDA thread can have its own instruction pointer. That means instructions from two different branches visited by threads of the same warp can now interleave, which was not possible before. It also means that explicit synchronization of warps is now necessary.

To accommodate this change the interface of a number of CUDA functions related to warps changed. These include so called shfl operations as well as ballot functions, which are extensively used in Kokkos' CUDA backend to implement reductions and other synchronization operations. The functions were deprecated (and they are broken on Volta) and replaced by functions with a suffix "_sync" which takes a thread mask as its first argument.

In preparation for the Sandia Astra machine, which features ARM CPUs from Cavium, we added support for their ThunderX2 chips to Kokkos' build system and implemented nightly testing on Sandia's Mayer testbed. Thus, Kokkos is now fully qualified on this platform and ready for deployment on Astra as soon as it comes online.

## 2.2 VENDOR COLLABORATION

The Kokkos team collaborated with numerous ECP PathForward vendors. We cannot report on details here since these collaborations generally fall under NDAs. Notable interactions happened with Intel, NVIDIA, ARM and AMD.

For AMD we have been working on getting a backend working for their GPU systems. While it is mostly functional and it passes around 90% of Kokkos' unit test suite there are a number of outstanding bugs which prevent further progress. In particular one cannot capture classes with compile time sized array members into a parallel region. This bug was reported to AMD in Q1FY18 and is expected to be fixed in Q1FY19. The Kokkos team attended also all face to face meetings with AMD. At the last one in September 2018, an extra day was added for specific Kokkos discussions. One result of the discussion was that a new backend development will be started using the HIP front end. This front end is more stable than the native HCC since it is more widely used even within AMD. The native HCC backend will be maintained for now anyway, as a test case for AMDs compiler team. Currently Kokkos provides the most comprehensive testing ground for that front end, which is available to AMD. A staff member from AMD is expected to spend about half his time working on the Kokkos backends.

With NVIDIA we worked on exploring NVSHMEM as an option for the new Kokkos Remote Memory Spaces capability. We do not yet have permission to talk about this in a public document. That said, NVIDIA invited the Kokkos team to present their early experience with NVSHMEM at SC18 in their booth. After that we expect to be able to generally share results.

We had numerous meetings with ARM representatives as part of the preparations for Sandia's Astra super computer – slated to be the largest ARM installation in the world when it comes online in October 2018. These discussions included progress on compilers, as well as discussions of future hardware capabilities and how those may map to Kokkos. A group of ARM employers participated in a Kokkos boot camp organized by the Cambridge University in the UK in March 2018 and run by Kokkos team members.

## 2.3 ISO/C++ COMMITTEE

The team has considerably intensified its engagement with

### 2.3.1 Heterogeneity phone meetings

Key members of the ISO C++ Standard Committee and other experienced C++ developers teleconference once weekly to discuss how to increase support in the C++ Standard for heterogeneous computer architectures. This includes graphics processing units (GPUs), FPGAs, fixed-function hardware accelerators, and distributed-memory parallel computers. Call participants represent not just traditional high-performance computing (HPC) users and vendors, but also developers for consumer and embedded hardware. Sandia staff work with this group to broaden support for their proposals and to ensure HPC representation and input in other proposals by the group. Sandians have become coauthors on numerous proposals as a result of their participation.

### 2.3.2 LWG issues processing

Sandia has several proposals pending before the ISO C++ Standard Committee, in various stages of the acceptance process. In order to learn how better to craft proposals to increase likelihood of acceptance, Sandia staff joined the Library Working Group (LWG). LWG is responsible for the official wording of all C++ Standard Library features. Before a proposal reaches LWG, it must first pass through other Working Groups which must agree that the proposal is technically sound and desirable. LWG requires that the proposal's language conforms with that of the Standard and contains no ambiguities. Sandia staff is actively involved in vetting other proposals and learning from more experience LWG members how best to write our proposals.

### 2.3.3 Shared Repository

In order to facilitate better collaboration with other DOE partners, we moved the repository to develop proposals and track issues to https://github.com/ornl/cpp-proposals-pub
This serves as a Kokkos independent collaboration place for Sandia and Oak Ridge. We are engaging other laboratories to contribute too.

### 2.3.4 atomic_ref

Sandia staff have successfully added language to the C++ 2010 Standard Draft to ensure correct C++ support for high-performance hardware atomic updates. This feature is called "atomic_ref." Hardware atomic updates give scientific and engineering codes a quick path to thread-parallelize existing loops in a correct and performant way, but lack of support for atomic_ref in the C++ Standard risked making this feature not correct in the future. Sandia staff first proposed atomic_ref to the ISO C++ Committee (WG21) in 2015. Staff worked with other C++ Committee members and editors to add official wording. In June 2018, WG21 accepted the proposal into the C++ 2020 Standard Draft. A description of the feature is already available here: https://en.cppreference.com/w/cpp/atomic/atomic_ref

### 2.3.5 mdspan

Sandia's "mdspan" proposal adds multidimensional arrays to the C++ Standard Library. This would let compiler vendors optimize multidimensional array access in C++, just as they do now in Fortran. It also offers a path forward for heterogeneous computing in Standard C++. Sandia staff brought their draft of the proposal to the C++ Standard Committee's Library Working Group (LWG). LWG spent several hours giving extensive editorial advice. As a result, staff completely reworded the mdspan proposal, and produced a reference implementation of mdspan. The resulting proposal is much closer to being accepted into the C++ Standard. The current goal is to bring it forward for inclusion into a C++ TS at the next meeting, though scheduling issues in the committee may prevent that. The mdspan proposal lives at: https://github.com/ORNL/cpp-proposals-pub/tree/master/P0009

### 2.3.6 executors

Executors is the primary vehicle intended for heterogeneous hardware support in C++. A member of the Kokkos team is heavily involved in the proposal preparation and is representing

HPC concerns. Many other committee members are not interested in features, which are absolutely necessary to make executors useful to the bulk of DOE’s C++ applications. Thus having a strong involvement is critical for our community. Participation includes weekly telecons as well as preparation and modifications of proposals and code prototypes.

### 2.4 RESOURCE ACCESS

Further evidence and details for the performed work can be found in the relevant

| Feature | Location | Access Restriction |
|---|---|---|
| Kokkos Issues | https://github.com/kokkos/kokkos/issues | Open access. |
| Kokkos AMD Support | https://github.com/kokkos/kokkos | Open access. Use ROCm backend. |
| C++ Proposals | https://github.com/ornl/cpp-proposals-pub | Open access. |

## 3. RESOURCE REQUIREMENTS

The work performed here required 1.25 FTE.

## 4. CONCLUSIONS AND FUTURE WORK

Ongoing improvements for our existing backends will be necessary going forward. At the same time backends for the new exa-scale architectures must be developed in collaboration with the vendors. The work with the C++ committee is going to stay an important part of our work, with the Kokkos team developing into one of the primary voices for the HPC community on the committee.

## 5. ACKNOWLEDGMENTS

This research was supported by the Exascale Computing Project (ECP), Project Number: 17-SC-20-SC, a collaborative effort of two DOE organizations—the Office of Science and the National Nuclear Security Administration—responsible for the planning and preparation of a capable exascale ecosystem—including software, applications, hardware, advanced system engineering, and early testbed platforms—to support the nation's exascale computing imperative.

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC. a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy’s National Nuclear Security Administration under contract DE-NA-0003525.