SAND2018-10738R

SANDIA NATIONAL LABORATORIES

SNL ADTM

FY18Q4 report for ATDM ST projects to ECP

September 14, 2018

Prepared by: Aaron Pennington

Prepared for:

ECP Confluence updates

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.

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology and Engineering Solutions of Sandia, LLC.

2.3.1.04 ATDM PMR

Accomplishments and Progress:

PERFORMANCE TO PLAN / MEETS

- Kokkos and DARMA teams worked together on several significant proposals for the C++ standards body, including the MDSpan and Exectutors proposals. This work has potential for long-term impact into the language standard reflecting core abstractions needed for our HPC programming and development efforts.
- Kokkos team collaborating with AMD via the Path Forward effort. This is advancing Kokkos support for [AMD] backends, necessary for future exascale/HPC architectures.
- Kokkos team continued to provide direct support for ASC codes Aria, SPARC, EMPIRE and Alexa, helping with algorithm design, performance analysis and software engineering issues. A significant feature providing tiled View layouts requested by SPARC was developed.
- DARMA team successfully completed work on DARMA-MPI interoperability mini-apps with PIC move kernel in DARMA and solver in MPI adapted from HPCCG.

EXCEEDS (these are not planned exceeds)

- DARMA team developed an MPI-based AMT runtime backend, and did initial performance testing on Trinity.

- **Q4: Kokkos back-end R&D and external collaborations**
  - The Kokkos Team reworked the MDSpan proposal for the C++ standard and implemented a standards conform reference implementation, which is getting evaluated by other committee members.
  - Continuing our collaboration with AMD, the team attended the AMD Path Forward Deep Dive, with an extra day added for specific AMD-Kokkos collaboration work.
  - The team took part in the Project 38 review, a collaboration between NSA and DOE to explore new hardware designs.
  - Status: complete

- Q4: **Kokkos R&D for SIMD types and remote memory spaces**
  - Remote spaces backends using MPI-3 one sided communication and using a prototype of NVIDIA's NVSHMEM library were developed.
  - An initial performance analysis using remote memory spaces for the particle halo in ExaMiniMD was performed.
  - SIMD types work was delayed due to coordination between the involved parties, and more performance design work required by one of the involved projects. That design work let to new requirements for the implementation on GPUs. We now expect to have an initial unified prototype by end of Q1 of FY19.
  - Listing this as complete since the SIMD work was a small component, and is expected to worked in the next quarter.
  - Status: complete

- Q4: **Kokkos support for ASC applications and libraries, and project management requirements**

- 46 pull requests were merged into Kokkos in Q4, addressing user reported bugs and usability enhancements.
- Direct support for ASC codes Aria, SPARC, EMPIRE and Alexa was provided, helping with algorithm design, performance analysis and software engineering issues.
- A significant feature providing tiled View layouts requested by SPARC was developed.
- This is essentially a level-of-effort activity, and the team remains actively engaged per the stated goals.
- Status: complete

- Q4: **DARMA interoperability with Kokkos, OpenMP, and MPI**
  - Completed DARMA-MPI interoperability mini-apps with PIC move kernel in DARMA and solver in MPI adapted from HPCCG. ECP reports submitted along with completion memos and highlights slide. Code repositories available on Github
  - Status: Complete

- Q4: **Development of agile components for backend development**
  - Updated all DARMA components to modern CMake system. DARMA backend now configurable with serialization, preprocessor, and meta-programming libraries and provides universal CMake package interface for all backend implementations and configurations. Published Spack packages repo. Code repositories available on Github. Submitted ECP completion memo and highlights slide.
  - Status: complete

2.3.2.04 ATDM Tools
Accomplishments and Progress:

PERFORMANCE TO PLAN / MEETS

- Tri-Lab Co-design L2 Milestone review completed successfully on Sept 12.
- Vectorization of SPARC significantly improved but the performance benefits are limited at best. (Mostly in the solvers)
- Trilinos continuous integration testing significantly improved this FY with benefit to the ATDM codes.
- Benchmarking of SPARC for Crossroads RFP as well as identification of "stress tests" including Kokkos testing and SPARTA.

EXCEEDS (these are not planned exceeds)

- Vanguard/Astra Arm-based platform delivered in September 2018. This exceeds the L2 milestone completion which only required placement of the contract.
- Initial benchmarking/baselining of ATDM codes and Trilinos on testbed HW for Vanguard/Astra Arm-based system.

- **[DEV] Coordination and General DevOps support for ATDM**

- Running Coordinated DevOps for ATDM process, management of ATDM DevOps backlog and Kanban board, running meetings with team representatives, following up on issues, etc.
- Continued support for the ATDM JIRA site (addressing move to CEE LAN management, addressing permissions issue between ATDM JIRA projects and git repository browsing, notification emails, etc.)
- Status: complete

- **[DEV] Deploy scalable build, integration and test framework**
  - Increased effort needed to complete Trilinos testing and integration work will not allow funding and time to complete this in the FY. Targets/priorities being re-evaluated for next year
  - Status: delayed

- **[DEV] Trilinos integration and multi-platform CI testing for ATDM apps**
  - Set of Trilinos builds being done by EMPIRE now complete on all platforms.
  - Doing continued evaluation and bug fix efforts with Kitware for new versions of CMake/CTest/CDash critical for reporting Trilinos testing results for ATDM and set of CDash site for ATDM APPs.
  - Status: complete
- **[PERF] Performance analysis and assessment of SPARC application on ATS platforms**
  - Assessment of vectorization levels achieved in SPARC (with Trilinos solver stack) is continuing. Results for improved Kokkos Kernels (with manual SIMD-ization) show improvement in some key routines. Vectorization levels significantly increased in SPARC but only the solver performance has shown significant improvement.
  - Status: complete

- **[PERF] Performance analysis and assessment of EMPIRE application on ATS platforms**
  - Benchmarking of EMPIRE on Mutrino (ATS-1) and Serrano (CTS-1) is underway with feedback provided to the Trilinos development teams.
  - Optimized TPL builds provided on Mutrino (ATS-1) for Haswell and are in development for Knights Landing. This will enable binaries for the full EMPIRE code to be built on Mutrino in preparation for future RAFT-copying to the SCN.
  - Status: complete

- **[PERF] Assess ATS-1 and ATS-2 development toolchains for tri-lab co-design L2 milestone**
  - Analysis complete and reported at the Tri-lab Co-design L2 Review. Work continues on parameterization of SIERRA/IC builds which may require changes to the bake system. CTH evaluation using compilers from GCC, Intel and IBM completed.

2.3.3.04 ATDM Math Libraries

PERFORMANCE TO PLAN / MEETS

Successfully supported the SPARC and EMPIRE L2 milestones:

- Local and global sensitivity analyses, forward UQ, and Bayesian calibration calculations were completed to support validation activities as part of the L2 milestone, and documented in the milestone report.
- Deterministic ROL-based calibration for the double-cone problem was also demonstrated and compared to the Bayesian results. This work highlighted challenges with derivative-based inversion with SPARC due to lack of smoothness of SPARC output quantities-of-interest with respect to some model parameters, which will need to be addressed next year.
- Established interface between Dakota and EMPIRE for the purpose of performing parameter variation studies. This was used for sensitivity analysis in FY18 of and will be used in FY19 for sensitivity analysis and calibration studies. Sensitivity analysis of Langmuir oscillation problem. The sensitivity of plasma amplitude, frequency, and phase with respect to particle properties, such as mass and charge, as well as simulation specifications, such as number density and velocity perturbation, was calculated. This will be extended in FY19 to analyze data collected from the Z machine over the past 15 years and calibrate the corresponding EMPIRE model.
- Contributed to the successful completion of the EMPIRE L2 milestone by Integrating Tempus into EMPIRE.
- Intrepid2 has been updated to improve integration performance substantially for CPU runs (more than 2x speedup in overall integration time, in tests) by removing subviews.
- Ran test problems to demonstrate the trajectory tracking capability of LOCA within SPARC. Results and a description of the LOCA integration work were written into an ATDM milestone completion report. Currently exploring ways to improve robustness/performance, i.e. trying out the tangent predictor, line searches, pseudo-time step adapting.

Accomplishments and Progress

- **Baseline performance data of existing linear solvers at scale and improve overall simulation time-to-solution be decreasing linear solver setup time by investigating multigrid methods, improving load balancing, and conducting cost-benefit analysis of effectiveness vs setup for a variety of solver capabilities (Technical Lead: Hu)**
    - Base-lining of solvers in EMPIRE was completed on up to 130K Haswell cores and up to 166M elements
    - Status: complete
- **Integrate sensitivity analysis and ROL optimization tools with Tempus in support of transient full-space optimization (Technical Lead: Phipps)**
    - Forward and adjoint sensitivity analysis method were implemented in Tempus in previous quarters. This quarter a ROL reduced-space optimization interface was implemented with Tempus and integrated into ROL. By the end of the quarter, a prototype full-space interface using Tempus will also be completed.
    - Status: complete

- **Create interoperability for Kokkoskernels with ASC, ATDM, and ASCR developed simulation codes via vector friendly data structures (Technical Lead: Rajamanickam)**
    - We have a SIMD data type that will work seamlessly between GPUs and KNLs and support complex vectorization use cases from Kokkos Kernels, ASC Mission apps such as SIERRA, externals apps such as Uintah (University of Utah), and embedded Uncertainty Quantification.
    - Status: complete

- **Strategic integration of agile components in SPARC and EMPIRE including finite element basis tools and discretization support, a LOCA continuation driver (for SPARC), time integration methods, and deployment of Kokkos (Technical Lead: Pawlowski)**
    - All of the sub-deliverables for this deliverable were completed successfully, with the exception of Tempus integration in SPARC
    - Status: complete

2.3.4.04 ATDM Data and Viz

PERFORMANCE TO PLAN / MEETS

- Good progress on IOSS capabilities for SPARC, including support for structured/unstructured grids, and addressing performance and scalability concerns.
- Completed prototype TuckerMPI/Catalyst in situ compression for SPARC. Waiting for opportunity to demonstrate to the team.

EXCEEDS

- IOSS: at the request of the users, added file-per-processor and single-file auto-decomposition. This goes beyond the scope of our original deliverable agreement with the application teams.

Accomplishments and Progress:

- **Develop and deliver input/output support for SNL ATDM Applications. FY18 focus is on checkpoint/restart capability and development of IOSS to support SPARC and EMPIRE.**
    - SPARC: Production IOSS-Based partitioning for structured/unstructured grids at scale on ATS-1: Worked with SPARC developers to identify issues in robustness and scalability in IOSS library and assist in adding IOSS capbilty to SPARC. All robustness issues were fixed and the scalability issues related to IOSS were fixed. I identified scalability issues in the underlying CGNS and/or HDF5 library and these are being fixed by the library developers with input from Sandia. Includes the newly requested file-per-processor input and output capabilities in addition to the single file auto-decomposition input and output.

    - EMPIRE: Initial checkpoint/restart of PIC structure. Checkpoint/restart code was developed for EMPIRE to manage particle data. Tests conducted on Mutrino demonstrated that FAODEL improves I/O by 1.8x-3x on DataWarp and 3-8x on Lustre.

- Checkpoint/restart for EMPIRE: A FAODEL interface for managing the checkpoint/restart of mesh field data was added to EMPIRE. A performance improvement of 4x was observed on Mutrino for the FAODEL implementation.

- Status: complete

- **Develop I/O capabilities for current and future ATS platform and ATDM application needs. FY18 focus is on burst-buffer support and I/O support for hybrid meshes**
  - While small scale tests were conducted on Ride, this work was postponed until FY19 due to hardware delays and a need for I/O support in the EMPIRE L2.

  - SPARC: Demonstrate IOSS-based hybrid mesh: Designed and implemented a serial representation of the hybrid mesh API in the IOSS library. Extending the implementation to provide parallel capability and document usage.

  - Status: complete

- **Develop and deliver scalable tools and algorithms for analysis and visualization (both *in situ* and post-processing) for next-generation, many-core architectures.**
  - We have brought in Todd Kordenbrock to continue this work on vectorization optimization for VTK-m. We have also done some additional performance metrics of external faces using some optimized sorts developed in Q3.

  - In Situ Compression: Completed the final steps of integration between Catalyst and TuckerMPI for compressing SPARC data in-situ. Added code to the "TuckerMPI Writer" plugin to handle multi-block data sets, perform normalization of cell data and handle various corner cases. A separate python script reconstructs the compressed files and maps the cell data onto the original mesh (which is not compressed) to write out files for verification. Demonstrated all the pieces with a 144^3 grid SPARC performance test case. Writing up the work for an ASC news note.

  - TuckerMPI compression of non-rectilinear grids: Identified the core set of sub-problems, and candidate solutions, to perform canonical functional tensor approximation of generic unstructured mesh data: (1) identify the best basis along each mode, (2) "learn" the coefficients for the bases, (3) strategy for efficient sub-sampling during the learning stage.

  - Status: complete

- **Integrate, demonstrate, and deploy analysis and visualization libraries and tools for ATDM applications on the ASC Advanced Technology Systems (ATS).**
  - The SPARC/Catalyst integration has improved its data interface (better representing piece extents) and improved support for TuckerMPI compression. We have also been compiling

and running on ATS-1 platforms Mutrino and Trinity and scaling up representative examples of using SPARC.
    - Status: complete

2.3.5.04 ATDM SW ecosystem

PERFORMANCE TO PLAN / MEETS

- An ECP milestone not associated with a deliverable below was completed this quarter: "Development and evaluation of system software implementations and interfaces for operating systems, multithreading, and communication technologies." Work completed under this milestone included contributions to the OpenMP 5.0 specification, put out as a comment draft in July, and to the MPI Forum process. Additionally, work on the Qthreads multithreading library included improved support for IBM POWER and Arm64 architectures and porting nimbleSM to Qthreads.
- In conjunction with TD developers, the Sierra/SM team enabled successful compilation of the TD application (including contact) with Relocatable Device Code (RDC) turned on. RDC allows for modern code architecture in algorithms targeted at GPUs and is needed to integrate the desired Sierra products (e.g. Lame, Math Toolkit, Geometry Toolkit) into the TD application. In addition, Sierra/SM is addressing all issues found in integrated testing and implementing, and NimbleSM will run an explicit dynamics simulation with contact on both the GPU and CPU by the end of the FY. This will allow performance comparisons and provide information on next steps to design contact algorithms for production GPU use.

Accomplishments and Progress:

- **(AOSR, FY18 Q4) Prototype usage of containers and related technologies to support ATDM developer workflows**
    - The project has developed several containerized software builds and multiple images which provide a consistent software programming environment for production-based workloads. This includes a base container image, the SEMS environment with GCC 7.2, along with Trilinos and associated necessary third-party library dependencies. Using this container environment, the Nalu CFD application has been containerized and experimentation is currently ongoing using Singularity on a CTS-based cluster. Container images and associated build recipes are available within Dockerhub and Gitlab-ex, respectively. Results from the Nalu container experiment will be presented within our accepted abstract at the NECDC 2018 conference in Los Alamos, NM, along with a broader investigation of container usage for production mission codes.
    - Status: complete

- **(AOSR, FY18 Q4) Resource manager applied to Darma+Kokkos use case scenarios**
    - The implemented solution manages execution (thread) resources between Darma and Kokkos using the affinity layer of the OpenMP standard as defined in the OpenMP 4.0/4.5

specifications, because the OpenMP Kokkos back-end provides the primary CPU execution space. This same mechanism also provides for interfacing to Uintah and could be used for other asynchronous many-task runtimes in the future.
    - Status: complete

- **(AOSR, FY18 Q4) Prototype of message-based, open-source simulation framework capable of quantifying MPI resource usage for MPI-based ATDM workloads**
    - The toolkit is completed and available in gitlab-ex for internal Sandia users. It may be shared with external collaborators upon request on a case-by-case basis. Supported features include extrapolation of smaller traces into larger synthetic traces, configurable network parameters, decomposition of collective operations, optional noise injection, fault-tolerance modeling, MPI match queue statistics, and cost modeling. Input traces from SPARC and SNL legacy ASC codes have been analyzed using this framework.
    - Status: complete