Argo

The operating system provides necessary functionality to libraries and applications, such as allocating memory and spawning processes, and manages the resources on the nodes in an exascale system. The Argo project is building portable, open-source system software that improves performance and scalability and provides increased functionality to exascale libraries, applications, and runtime systems, with a focus on resource management, memory management, and power management.

Many exascale applications have a complex runtime structure, ranging from in situ data analysis, through an ensemble of largely independent individual subjobs, to arbitrarily complex workflow structures. To meet the emerging needs of exascale workloads, while providing optimal performance and resilience, the compute, memory, and interconnect resources must be managed in cooperation with applications, libraries, and runtime systems. The goal of Argo is to augment and optimize low-level system software components for use in production exascale systems, providing portable, open-source, integrated software that improves the performance and scalability of and that offers increased functionality to exascale applications, libraries, and runtime systems. The project focuses on resource management, memory management, and power management.

The Argo team is delivering resource management infrastructure to coordinating static allocation and dynamic management of node resources, such as memory and caches. By offloading system-specific aspects such as topology mapping and partitioning of massively parallel resources, this infrastructure will improve the performance and portability of exascale applications and libraries and their runtimes.

They are developing memory management libraries to provide flexible and portable memory management mechanisms that make it easier to obtain high performance and to incorporate nonvolatile memory into complex memory hierarchies using a memory map approach. These libraries will directly support new applications that analyze large, distributed data sets and make it easier to program heterogenous hardware resources.

They are providing fully integrated, end-to-end infrastructure for power and performance management, including power-aware plugins for resource managers, workflow managers, job-level runtimes, and a vendor-neutral power control library. This infrastructure addresses head-on the challenge of managing the performance of exascale applications on highly power-constrained systems.

PI: Pete Beckman, Argonne National Laboratory

Collaborators: Argonne National Laboratory, Lawrence Livermore National Laboratory

Progress to date

- The Argo team developed an initial version of the unified Node Resource Manager, which provides high-level control of node resources, including initial allocation at job launch and dynamic reallocation at the request of the application and other services. The Node Resource Manager integrates dynamic power control and provides support for tracking and reporting of application progress.

- The team released a first version of UMap, a user-space memory map page fault handler for nonvolatile memory that maps virtual address ranges to persistent data sets and transparently pages in active pages and evicts unused pages.

- The team developed AML, a memory library for the explicit management of deep memory architectures that features a flexible and composable interface, allowing applications to implement algorithms similar to out-of-core for deep memory. Multiple optimized versions of these memory migration facilities, using synchronous and asynchronous interfaces and single- and multi-threaded backends, were included.

- The team released an interface between the Node Power and Node Resource Manager services, which in turn allows their Global Resource Manager to control and monitor power and other node-local resources. Additionally, the team studied the effect of power capping on different applications using the Node Power interface and developed the power regression models required for a demand-response policy.