Advanced Technology System Scheduling Governance Model

Authors:
**Jim Ang (SNL), Brian Carnes (LLNL),
Thuc Hoang (ASC HQ), Manuel Vigil (LANL)**

June 11, 2015
LLNL-TR-680598

Disclaimer

This document was prepared as an account of work sponsored by an agency of the United States government. Neither the United States government nor Lawrence Livermore National Security, LLC, nor any of their employees makes any warranty, expressed 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 Lawrence Livermore National Security, LLC. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States government or Lawrence Livermore National Security, LLC, and shall not be used for advertising or product endorsement purposes.

Contents

Introduction

In the fall of 2005, the Advanced Simulation and Computing (ASC) Program appointed a team to formulate a governance model for allocating resources and scheduling the stockpile stewardship workload on ASC capability systems. This update to the original document[1] takes into account the new technical challenges and roles for advanced technology (AT) systems and the new ASC Program workload categories that must be supported. The goal of this updated model is to effectively allocate and schedule AT computing resources among all three National Nuclear Security Administration (NNSA) laboratories for weapons deliverables that merit priority on this class of resource. The process outlined below describes how proposed work can be evaluated and approved for resource allocations while preserving high effective utilization of the systems. This approach will provide the broadest possible benefit to the Stockpile Stewardship Program (SSP).

The objectives of this governance model are to:

- Ensure that AT system resources are allocated on a priority-driven basis according to SSP Program requirements.

- Utilize ASC AT systems for the most demanding workload categories, for which they were designed and procured.

- Support the role of AT systems to prepare ASC resources (including its people, its applications, and its computing environments), for significant changes in *future* computer and system architectures.

Within the constraints of meeting these primary objectives, this model maximizes effective use of a system both by minimizing idle cycles and by enhancing the probability of productive and useful calculations. An important, but secondary, objective is to simplify the prioritization and allocation processes to assure that these do not impede successful attainment of the primary objectives. This paper describes and leverages existing review bodies, updates the character of work packages, and establishes a framework for the proposal process as well as procedures for prioritizing proposals, allocating resources, and collecting relevant data to measure progress.

AT systems may be based on a globally consistent architecture but may also include architecturally diverse system partitions based on the results of a competitive procurement. Systems that include diverse architecture partitions may be operated as separate application execution environments, one of which may be more architecturally advanced and therefore more of a preparation challenge for the applications. The ASC Program recognizes that as it transitions to next-generation system technologies, getting applications to run or run efficiently on new architectures will be quite challenging. In such cases, the workload categories will apply to the specific application execution environment, each of which should support the specific mission need for which the system was designed. Allocation of computing resources on the most advanced application execution environment must be prioritized on either demonstrated application

---

[1] *Capability Compute System Scheduling Governance Model Advanced Simulation and Computing Program*, April 20, 2006, NA-ASC-109R-06-Vol.1-Rev.0

readiness to use that particular architecture or on getting applications ready for the new architecture. The acknowledgement of distinct compute partitions does not preclude the opportunity or ability for jobs to span partitions, as may be required for advanced ASC workflows.

Workload Organization

An NNSA AT system is similar in value and uniqueness to a large experimental facility. For this reason, the process described here to request and review proposals to utilize these systems is similar to that of experimental facilities, while taking into account that these systems uniquely support the stockpile stewardship mission. Major programmatic computing efforts will be organized as computing work packages and will be reviewed and prioritized for relevance, importance, and technical rationale. Each AT system will be scheduled via a series of Advanced Technology Computing Campaigns (ATCCs). Each campaign will be six-months in duration.

The ATCC concept respects the proven-by-time strategy employed for running major computational efforts, in that several calculations, often building up in size, are run in support of one computing campaign. This approach provides the necessary verification of the calculation methodology to maximize the understanding and value gained from a full-sized major calculation. In addition, a major calculation is often followed by or accompanied by smaller supporting calculations (for example, coarser mesh, assumed symmetries, or physics approximations) to provide additional insight and explore sensitivities.

Under the proposed model, there are three workload categories relevant for inclusion in a proposed ATCC project:

- *Category 1* **(C1):** Jobs that require more computational resources than are available on commodity technology (CT) systems. This class of job taxes the capability of the AT system and represents the most computationally demanding work run in the complex. This class of job may also include scaling studies in preparation for long-running calculations on existing and future systems.

- *Category 2* **(C2):** Uncertainty Quantification (UQ) studies that are comprised of several thousand smaller jobs, none of which individually uses a large number of the available nodes on a system or application execution environment but in aggregate require compute resources that cannot be found in any other venue. This class of job may also consist of large parameter calculations and performance studies.

- *Category 3* **(C3):** Jobs to prepare production codes to run robustly at large scale using accelerators, advanced memory technology, or other unique architectural components of an AT system or application execution partition. This class of job may consist of an ensemble of jobs to support application development, application readiness, performance analysis, or development/test of new computing environment capabilities at scale.

Allocations

Each laboratory (Livermore, Los Alamos, and Sandia National Laboratories) will assume a one-third allocation on each AT system and will hold an internal process for each six-month campaign to review and prioritize proposed projects. Each laboratory will then send its recommend list of projects for each campaign to the Advanced Technology Executive Committee (ATEC) for final approval. If at any time there is program need to vary from the normal one-third, one-third, one-third lab allocation on any particular ATCC, then the Advanced Technology Planning Advisory Committee (ATPAC) will hold a tri-lab video conference to develop a recommendation to be made to the ATEC. The ATPAC must also be available to confer as necessary for any mid-course corrections.

The ATPAC will consist of six representatives, two from each of the three laboratories appointed by the respective ASC executive. ATPAC members must be capable of representing the national SSP priorities and understand the Directed Stockpile Work (DSW) workload. Each laboratory should include at least one computer-knowledgeable representative among its members.

Each laboratory should prioritize its proposals based on the priorities of the Weapons Program, including relevance to Level 1 and Level 2 milestones, relevance to stockpile (DSW) deliveries, importance in the progression towards Predictive Science goal, and preparation for future computer and system architectures. Since multiple AT platforms may be available to the complex at any given time, the ATCC projects should be proposed to run on the system that best matches the needs of a particular campaign. In addition, each laboratory will be responsible for reviewing the technical rationale of its proposals to ensure the ATCC projects are sound and free of extraneous work. It is important that the review process ensure the associated codes are validated and can effectively use the selected AT system. The laboratories will provide directions to the host site(s) for executing their approved projects and their associated machine allocations.

The ATEC will consist of the ASC executive (or designee) from each laboratory and representatives from the NNSA ASC office. ASC Headquarters (HQ) will keep other HQ stakeholders informed of the semiannual allocations and the results of past campaigns. The ATEC will review the prioritized list of projects and allocations sent forward by the three laboratories and will make adjustments if necessary. The ATEC will approve the final list of ATCC projects for each AT computing campaign with the associated machine allocations.

Outside of the normal review period, the ATEC is empowered at any time to advise the ATPAC representatives at each laboratory of ATEC-directed change to ATCC approvals, priorities, and allocations to accommodate unanticipated and critically important programmatic work. The ATPAC will forward notification of these changes to the tri-lab technical Expedited Priority Run (EPR) body. The EPR body will continue to meet weekly to manage emergency situations, report up to the ATPAC as necessary, and address user issues in the tri-lab community.

It is the responsibility of each site hosting an AT system to implement and enforce ATEC directions and allocations according to the approved prioritization. In addition, the host

site will retain use of the computer as required for system maintenance, upgrades, and software development for upgrades and enhancements of the system. Although the host site will not be responsible for managing the work portfolio itself, it must be aware of the priorities established and coordinate efforts with the submitting site to successfully execute the ATCC projects in accordance with ATEC direction. It is the responsibility of the submitting site to ensure execution of the prioritized ATCC projects. If the submitting site is not able to use its full allocation within the period of approved access, either due to the failure or early completion of one or more ATCC projects within its portfolio, the hosting site will enable other ATCC projects waiting for resources according to ATPAC directions. Complications that arise from this adjustment will be forwarded to the ATPAC for resolution. If these complications result in the need to modify priorities or directions, the ATPAC will seek approval for these modifications from the ATEC. Any ATCC project will be removed from the approved list when the time period for access has passed or the project team reports that the project is completed, whichever occurs first. If the project needs extra computing time to finish, a request must be made to the ATPAC or to the relevant ASC executive (see "ASC Executives' Reserve" below) for approval.

ASC Executives' Reserve

An allocation of 15% of the compute cycles per annum will be explicitly reserved for use at the discretion of the ASC executives to cover urgent but unanticipated needs not explicitly met through active ATCCs. This will be achieved by allocating 5% to each laboratory.

Implementation

Site-local resource management tools will provide a means of implementing and enforcing allocations according to the priorities established by the ATEC. Since ATCC projects will progress at differing rates (due to, for example, idea gestation, extended analysis of results, and bug searches), multiple ATCC projects will be approved to access the machine, effectively alternating use of the machine, to ensure continued progress for all ATCC projects and to maximize utilization of the computer.

As an additional mechanism to maximize utilization of the computer, there will be a *separate* process to gain access to the computer, called Standby, where a fair-share bank will be given to each laboratory to be managed by that laboratory. Jobs will be submitted using these banks and are free to run unless they are preempted by an ATCC calculation. Some opportunity will be given to a Standby job to checkpoint itself; however, the intent is to minimize the frictional effect and frustrations encountered by the ATCC team seeking immediate access to capability cycles. Utilization by Standby jobs will not be counted as part of any approved ATCC project.

Reporting and Accountability

Host laboratories will prepare ATCC utilization reports on a monthly basis. A tri-lab utilization report will be generated for each ATS resource, distributed to both the ATPAC and ATEC, and will include utilization broken down by project. A common reporting mechanism has been developed for validating the ATCC project requests against their historical usage. This helps improve accuracy in the resource requests and ensures accountability in the use of the resources.

Once a year the host lab ATEC representative or delegate will brief NNSA HQ staff on the simulations that were carried out in the previous two ATCCs on the AT system. The briefs will alternate in time when multiple AT systems are simultaneously in production. The ATPAC representatives of each laboratory are responsible for collecting slides from each of their laboratory's computational efforts on each AT system, and forwarding the collected slides to the entire ATPAC. The simulation brief along with the relevant utilization reports constitutes the final report for each ATCC.