Capability Compute System Scheduling Governance Model
Advanced Simulation and Computing Program

April 20, 2006

Jim Ang, SNL, Brian Carnes, LLNL, Sudip Dosanjh, SNL,
Rick Martineau, LANL, Ken Perry, HQ DSW, Jim Rathkopf, LLNL,
Manuel Vigil, LANL, Cheryl Wampler, LANL/HQ rep

INTRODUCTION

In the fall of 2005, the ASC Program appointed a team to formulate a governance model for allocating resources and scheduling the stockpile stewardship workload on ASC Capability Systems (see Charter, Appendix A). This document describes a model that is devised to allocate capability-computing resources for weapons laboratory deliverables that merit priority on this class of resource and that cannot be reasonably attempted on other resources. The process outlined describes how capability work can be evaluated and approved for resource allocations, while also preserving high effective utilization of the systems. This approach will provide the broadest possible benefit to the Program.

The objectives of this initiative are:

- To ensure that the capability system resources are allocated on a priority-driven basis according to the Program requirements;
- To utilize ASC Capability Systems for the large capability jobs for which they were designed and procured

Within the constraints of meeting the two primary objectives, this model maximizes effective use of the machine both by minimizing idle cycles and by enhancing the probability of productive and useful capability calculations.

An important, but secondary, objective is to simplify the prioritization and allocation processes in order to assure that these do not impede successful attainment of the primary objectives.

This paper authorizes the creation of relevant review bodies, describes the character of work packages, establishes a framework for requesting and reviewing proposals as well as procedures for prioritizing proposals, allocating resources, and collecting relevant data to measure progress both from capability providers and users.

WORKLOAD ORGANIZATION

A capability class 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 ASC capability 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 proposed work package, called a *Capability Computing Campaign* (CCC), consists of at least one major calculation needing a significant proportion of an ASC capability system, together with related supporting jobs of smaller sizes. The portfolio of CCCs should provide a reasonable balance between the objective to ensure that the resources are successfully applied to the highest programmatic needs and the objective to use ASC capability systems effectively for capability jobs that cannot be run on any other computing resource.

The CCC concept respects the tried-by-time strategy employed for running major computational efforts, in that several smaller calculations, often building up in size, are run in support of one or more large calculations. This approach provides the necessary verification of the calculation methodology to maximize the understanding and value gained from the full-size calculation. In addition, a major calculation is often followed by or accompanied by smaller supporting calculations (coarser mesh, assumed symmetries, physics approximations, etc.) to provide additional insight and explore sensitivities. For example, on a peak 100 TF class system, a reasonable CCC proposal might request several months of access and include a 40 TF main calculation supported by a number of smaller calculations. The proposed CCC would demonstrate that all smaller jobs are essential for and in support of the capability run(s). The allocation process (described below), in turn, will ensure that the technical rationale includes an explanation of why these smaller jobs are impractical to run on capacity resources.

Under the proposed model, there are three regimes of job size relevant for inclusion in a proposed CCC:

- *Category 1* (C1): Capability jobs that use 75% or more of the available nodes on a system. This class of job fully taxes the capability of the machine, and represents the most computationally demanding work run in the complex. Some C1 work will also include scaling studies in preparation for long running calculations on existing and future capability systems. Due to its size, a single job in this category will effectively require dedicated use of the machine, and will require careful scheduling in order to allow other CCCs to advance.
- *Category 2* (C2): Capability jobs that use between 30% and 75% of the available nodes on a system. This class of job will typically consist of large production weapons calculations and performance studies, which depending on their exact sizes, would generally permit two C2 jobs to run on the system simultaneously. This is desirable in order to permit multiple CCCs to advance simultaneously.

- *Category 3* (C3): Jobs that use less than 30% of the available nodes that are part of a CCC. These smaller jobs are essential for carrying out a CCC and can be run only on the capability system. A CCC may not consist of C3 jobs only but must include at least one C1 or C2 job.

This approach will be successful only if sufficient alternative resources are available to carry out the capacity workload of the complex, characterized by tens-of-thousands smaller calculations. If this is not the case, then an alternative utilization scenario must be devised for the capability platforms, as, inevitably, these will be preempted frequently for high-priority capacity calculations.

ALLOCATION

Allocation of capability resources will be achieved using a two-step process. First, a Capability Planning Advisory Committee (CPAC) will review the proposed CCCs and make a recommendation to the Capability Executive Committee (CEC). Second, the CEC will review this recommendation and make the final allocations.

The CPAC will consist of nine representatives, three from each of the three Laboratories appointed by the respective ASC Executive. CPAC members must be capable of representing the National Stockpile Stewardship program priorities and understand the DSW workload. Each Laboratory should include at least one computer-knowledgeable representative among its members. The CPAC will meet on a biannual basis to evaluate proposals and will be available to confer monthly if necessary for course corrections.

The CPAC will define the details of the proposal process, including the format of the request for Proposals and the Response to request for Proposals. It is important that it assure a “light-weight” proposal process for mid-year resubmissions of a continuing project.

After evaluating proposals, the CPAC will prioritize these based on the priorities of the Weapons Program, including relevance to Level 1 and Level 2 milestones, relevance to stockpile (DSW) deliveries, and importance in the progression towards predictability (e.g. removing the “knobs” as elucidated in the ASC Roadmap, in particular when such efforts are closely coordinated with the Campaigns). Since multiple capability platforms may be available to the complex at any given time, the CCCs will be prioritized based upon the aggregate capability available. In addition, the CPAC will be responsible for reviewing the technical rationale to ensure that the CCC is sound, ready to make effective use of the machine and is free of extraneous work. It is important that the review process ensure that the associated codes are validated and run efficiently. An important component of this technical review will be an examination of the proposed C3 class jobs to assess their potential to run on available capacity systems. The CPAC will present the recommended list of prioritized CCCs, directions to the host site(s) for executing these priorities, and the associated machine allocations, to the CEC.

The CEC will consist of the ASC Executive (or designee) from each Laboratory and representatives from the NNSA ASC office. It is also recommended that the CEC include at least one representative from federal DSW management. The CEC will review the prioritized list and allocations sent forward by the CPAC, and will make adjustments as necessary. The CEC will approve the final list of CCCs with associated machine allocations, as well as the list of CCCs approved for execution on targeted platforms for the review period and deliver this information to the sites. Note that in the event that the reviewed and approved CCCs exceed available resources, the prioritized list will contain CCCs that cannot initially be accommodated immediately. These CCCs will be viewed as waiting for resources that may become available in the event that previously approved CCCs are cancelled or complete ahead of schedule.

Outside of the normal review period, the CEC is empowered at any time to advise the CPAC of CEC-directed changes to CCC approvals, priorities, and allocations to accommodate unanticipated and critically important programmatic work (preempting the normal CCC approval process). The CPAC will forward notification of these changes to the Tri-Lab EPR.

The tri-lab technical Expedited Priority Run (EPR) body will continue to meet weekly to manage emergency situations, report up to the CPAC, and address user issues in the tri-lab community.

It is the responsibility of each site hosting a capability system to implement and enforce CEC directions and allocations according to the approved prioritization. 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 CCCs in accordance with CEC direction. It is the responsibility of the submitting site to ensure execution of the prioritized CCCs. 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 CCC within its portfolio, the hosting site will enable CCC(s) waiting for resources according to CPAC directions, with rapid notification to the CPAC. Complications that arise from this adjustment will be forwarded to the CPAC as well. If these complications result in the need to modify priorities or directions, the CPAC will seek approval for these modifications from the CEC. A CCC will be removed from the approved list when the allocation is exhausted, or 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 CPAC or to the relevant ASC Executive (see 5% provision below) for approval. At subsequent CPAC and CEC meetings, the Laboratories shall report CCC status and utilization.

In the event of urgent but unanticipated, high priority need for capability resources, the CEC has the authority to preempt the resources. In addition, the CEC may invoke a process to “renormalize” the allocations as necessary.

An allocation of 15% of the compute cycles per annum will be explicitly reserved for use at the discretion of the ASC Execs to cover urgent but unanticipated needs not explicitly met through active CCCs. This will be achieved by allocating 5% to each Laboratory. This allocation is intended for capability computing; however it is possible that pressing programmatic work may not meet the stringent Category 1 or 2 criteria, so no conditions will be explicitly imposed to enforce this outcome. Nonetheless, it is expected that Executives will be ready to explain and justify their interventions and choice of platforms at CEC meetings.

In addition, it must be acknowledged that 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.

IMPLEMENTATION

Site-local resource management tools will provide a means of implementing and enforcing allocations according to the priorities established by the CEC. Since CCCs will progress at differing rates (due to idea gestation, extended analysis of results, bug searches, etc.), multiple CCCs will be approved to access the machine, effectively alternating use of the machine, in order to assure continued progress for all CCCs 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: a fair share bank will be given to each laboratory to be managed by that laboratory. Site allocations for the standby resources will be determined by the CEC. Jobs will be submitted using these banks and are free to run unless they are preempted by a CCC 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 CCC team seeking immediate access to capability cycles. Utilization by Standby jobs will not be counted as part of any approved CCC. An example illustrating the envisioned utilization of the machine when in full capability mode is to be found in Figure 1, showing approved CCCs in various Categories and the Standby load.

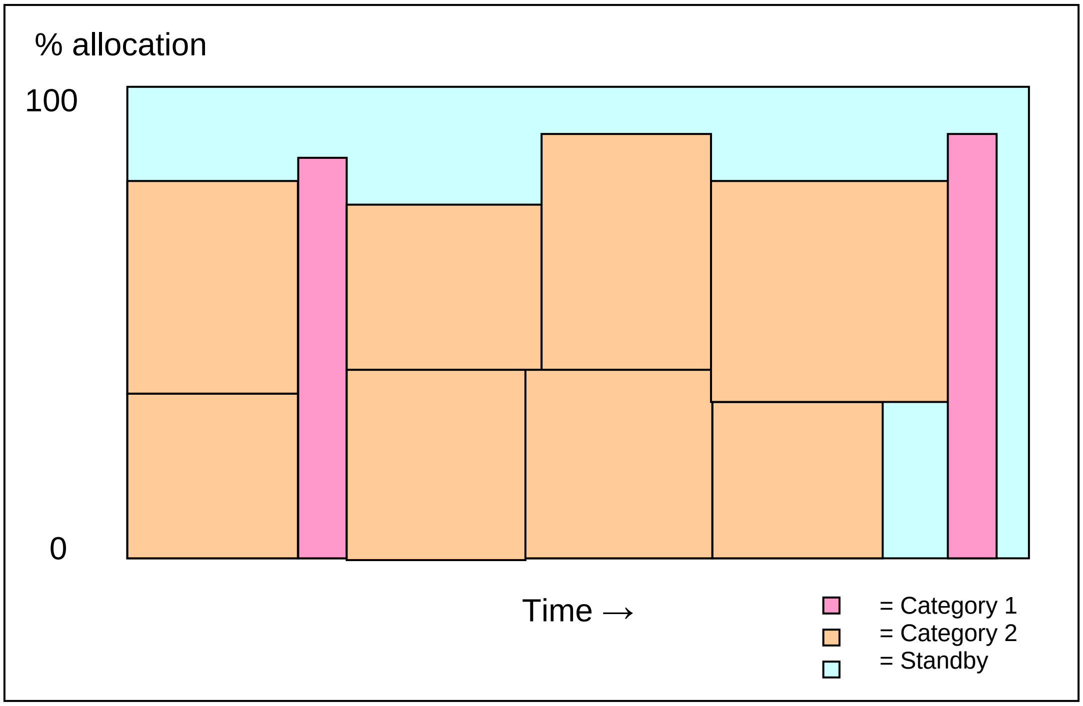

Figure 1: An illustration of the usage of an ASC Capability System versus time

REPORTING AND ACCOUNTABILITY

An annual ASC Capability Systems Tri-lab Review will be held to review the results of this governance model, providing an opportunity to adjust the process if necessary.

Host labs will prepare CCC utilization reports on a semi-annual basis, to coordinate with the semi-annual CPAC and CEC proposal review schedule. A single tri-lab report should be generated and distributed to both the CPAC and CEC and should include utilization broken down by CCC and by C1 through C3 categories for each capability class machine in the complex. It may be desirable to include additional information on job size or runtime breakdown.

A common reporting mechanism is needed for validating the CCC requests against their historical usage. This will improve accuracy in the resource requests and will ensure accountability in the use of the resources. It is recommended that this effort be coordinated with the Workload Characterization Project, with the aim that a tool be available for tri-lab common reporting.

UCRL-WEB-222594