



LAWRENCE
LIVERMORE
NATIONAL
LABORATORY

LLNL-TR-417191

Level-2 Milestone 3244: Deploy Dawn ID Machine for Initial Science Runs

Milestone report for NNSA HQ

*Prepared by David Fox
September 21, 2009*

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.

This work performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract DE-AC52-07NA27344.

Contents

Introduction.....	1
Dawn System Architecture	2
Current Status.....	3
Installation and Integration Time Line	3
Testing and Acceptance Process	3
Attachment 1: Milestone Definition Text	5
Attachment 2: Project Plan	6
Attachment 3: Handoff Letter	7
Attachment 4: Science Run Results	8

Introduction

This report documents the delivery, installation, integration, testing, and acceptance of the Dawn system, ASC L2 milestone 3244: Deploy Dawn ID Machine for Initial Science Runs, due September 30, 2009. The full text of the milestone is included in Attachment 1. The description of the milestone is:

This milestone will be a result of work started three years ago with the planning for a multi-petaFLOPS UQ-focused platform (Sequoia) and will be satisfied when a smaller ID version of the final system is delivered, installed, integrated, tested, accepted, and deployed at LLNL for initial science runs in support of SSP mission. The deliverable for this milestone will be a LA petascale computing system (named Dawn) usable for code development and scaling necessary to ensure effective use of a final Sequoia platform (expected in 2011-2012), and for urgent SSP program needs. Allocation and scheduling of Dawn as an LA system will likely be performed informally, similar to what has been used for BlueGene/L. However, provision will be made to allow for dedicated access times for application scaling studies across the entire Dawn resource.

The milestone was completed on April 1, 2009, when science runs began running on the Dawn system. The following sections describe the Dawn system architecture, current status, installation and integration time line, and testing and acceptance process.

A project plan is included as Attachment 2. Attachment 3 is a letter certifying the handoff of the system to a nuclear weapons stockpile customer. Attachment 4 presents the results of science runs completed on the system.

Dawn System Architecture

The Dawn machine is an IBM BlueGene/P system composed of 36 compute racks. Each compute rack has 1024 compute nodes and 8 I/O nodes. Both compute nodes and I/O nodes have four PowerPC 450 processor cores and a total of 4 GB of memory. The system has a total of 147,456 compute cores, 288 I/O nodes, 148 TB total memory, and 501 TF peak performance.

The end-user access to the system is via the Front End Nodes (FENs). Dawn has 14 FENs, each with four Power6 processors running at 4 GHz with 8 GB of RAM. The FENs are used to compile, run, and monitor jobs and access file system resources.

Each of the 288 I/O nodes provides I/O services for 128 compute nodes. The I/O nodes mount the NFS and Lustre file systems. The Lustre file system deployed along with Dawn is 2.3PB in size and delivers a peak performance of 60 GB/s.

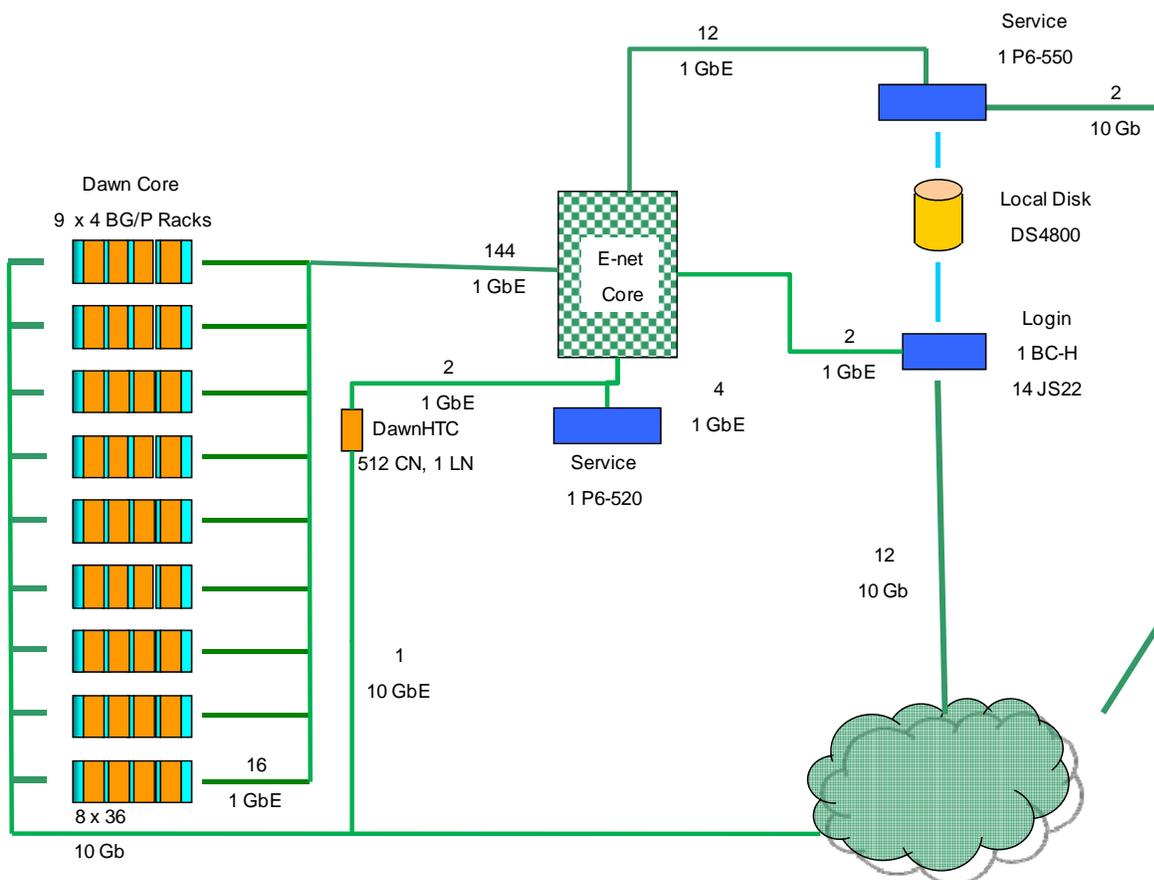


Figure 1: Dawn Architecture

Current Status

The science runs on Dawn ran for 22 weeks and were completed on August 31, 2009. At that time, the latest software updates from IBM were installed (version 1 release 4) and testing of this update continued until September 21, 2009. Version 1 release 4 of the software delivers bug fixes discovered during the acceptance tests and during the science runs. The software also delivers early technology enhancements to allow for testing and development that will be relevant to the Sequoia system. On September 21, 2009 the Dawn system was moved to the SCF environment.

Installation and Integration Time Line

The planning for the Dawn system began in October 2008. The layout of the Dawn machine and the cooling requirements necessitated a specially designed floor support system. The LC facilities team worked with IBM to develop a design for the power and floor support system. The floor support was built to specification by an outside contractor. The power modifications and installation of the floor support system were finished January 9, 2009.

The final contract with IBM was signed January 8, 2009, and the first set of 18 racks arrived on January 20. Between January 20 and March 2, 2009, the final 18 racks were delivered and all 36 racks installed and integrated into the LC environment. Special coordination between IBM and LC staff allowed for early system and application testing on racks as they were installed; therefore, testing had progressed up to more than one half the final system before installation was complete. This early system testing allowed for a compressed system test and verification phase. The full system was tested by IBM and LLNL for two weeks (March 3–March 17, 2009). The acceptance testing on the machine was completed and the machine accepted on March 27.

Testing and Acceptance Process

The acceptance tests for the Dawn machine were:

- The TPP Linpack benchmark
- The HPC Challenge benchmark
- Successful four hour runs of the five ASC marquee codes
 - UMT—3D unstructured-mesh photon transport
 - AMG—Algebraic multigrid linear system solver
 - IRS—Implicit radiation solver for 3D radiation diffusion equations
 - SPhot—A Monte Carlo photon transport code
 - LAMMPS—A full system classical molecular dynamics simulation code
- A successful eight-hour run of the IOR I/O benchmark
- The Synthetic Work Load (SWL) test

The Linpack run on Dawn was completed on March 12, 2009 and delivered 415 TF, which is 83% of peak. The HPC Challenge benchmark was completed March 19. The four-hour ASC marquee codes were run by IBM as part of the early full system testing. The performance results from the marquee code runs were 7.36% higher than the targets specified in the Sequoia Statement of Work. No single eight-hour run of IOR was done, but many days of Lustre testing were completed on the system, and the stability of the file system was tested as part of the SWL. In addition, during the 22 weeks of science runs, Lustre testing continued, and the Lustre team worked with the science run users to help tune their applications and debug problems. The result was increased performances of Lustre from 15 GB/s write and 9 GB/s read to 30 GB/s write and 44 GB/s read.

The SWL test comprises three parts: the functionality tests (FT), the performance tests (PT), and the stability tests (ST). The SWL-FT and SWL-PT are each 24 hours of tests over two days, and the SWL-ST was 5 days. A large number of LC staff along with IBM staff were involved in the SWL tests. The SWL-FT and SWL-PT tests were run and verified March 17–20, 2009. The SWL-ST was run from March 23–27, 2009. Testing prior to and during the SWL uncovered a total of 27 bugs that were reported to IBM, however, none were considered critical such as to make the machine unusable. By the time the SWL was completed, 11 bugs were still open with fixes being worked for the V1R4 install.

Attachment 1: Milestone Definition Text

Milestone (3244): Deploy Dawn ID machine for initial science runs				
Level: 2		Fiscal Year: FY09		DOE Area/Campaign: ASC
Completion Date: Sep-09				
ASC nWBS Subprogram: CSSE, FOUS				
Participating Sites: LLNL				
Participating Programs/Campaigns: ASC				
<p>Description: This milestone will be a result of work started three years ago with the planning for a multi-petaFLOPS UQ-focused platform (Sequoia) and will be satisfied when a smaller ID version of the final system is delivered, installed, integrated, tested, accepted, and deployed at LLNL for initial science runs in support of SSP mission. The deliverable for this milestone will be a LA petascale computing system (named Dawn) usable for code development and scaling necessary to ensure effective use of a final Sequoia platform (expected in 2011-2012), and for urgent SSP program needs. Allocation and scheduling of Dawn as an LA system will likely be performed informally, similar to what has been used for BlueGene/L. However, provision will be made to allow for dedicated access times for application scaling studies across the entire Dawn resource.</p>				
<p>Completion Criteria: Dawn platform is (1) delivered, (2) installed, integrated and tested, (4) accepted, and (5) deployed at LLNL for initial science runs.</p>				
<p>Customer: SSP early science applications and customers.</p>				
<p>Milestone Certification Method:</p> <p>(1) Professional documentation, such as a report or a set of viewgraphs with a written summary, is prepared as a record of milestone completion.</p> <p>(2) The “handoff” of the developed capability (product) to a nuclear weapons stockpile customer is documented.</p>				
<p>Supporting Resources: Computational Systems and Software Environment projects, Facility Operations and User Support projects.</p>				
Supporting Milestones:				
Program		Title		Due Date
None				
<p>Codes/Simulation Tools Employed: Unclassified IC, PEM or V&V science codes.</p>				
<p>Contribution to the ASC Program: Overall Sequoia platform strategy is based on availability, for at least two or three years, of a significant ID system upon which simulations can evolve and adapt to the architectural evolution inherent in Sequoia platform planning process. This approach enhances the probability of effective use of Sequoia during its lifecycle.</p>				
<p>Contribution to Stockpile Stewardship: In addition to the above strategic approach for simulation code evolution and adaptation to Sequoia architectures, the Dawn ID system will also be used for urgent nearer-term SSP program needs, in particular for UQ, National Boost Initiative and energy balance PCF pegpost deliverables.</p>				
No.	Risk Description	Risk Assessment (low, medium, high)		
		Consequence	Likelihood	Exposure
1	Platform vendor parts commit and build is late, affecting the scheduled delivery timeline	High	Low	Moderate
2	Installation, testing, and/or acceptance is delayed due to unforeseen problems	High	Low	Moderate
3	Initial science applications are not ready to run on Sequoia ID	Moderate	Very Low	Low

Attachment 2: Project Plan

ID	Task Name	Duration	Start	Finish	Predecessors	% Complete
1	Dawn Site Prep	78 days	Thu 10/2/08	Mon 1/19/09		100%
2	Power, Seismic, network -under floor	75 days	Thu 10/2/08	Wed 1/14/09		100%
3	Design power layout	20 days	Thu 10/2/08	Wed 10/29/08		100%
4	Design floor support	20 days	Thu 10/2/08	Wed 10/29/08		100%
5	Order and receive power/support material	15 days	Thu 10/30/08	Wed 11/19/08	4	100%
6	Scan floor	1 day	Mon 11/24/08	Mon 11/24/08	3,5	100%
7	Design seismic	23 days	Thu 10/2/08	Mon 11/3/08		100%
8	Order and receive seismic material	9 days	Tue 11/4/08	Fri 11/14/08	7	100%
9	Seismic fabrication	14 days	Mon 11/17/08	Thu 12/4/08	7,8	100%
10	Seismic install	29 days	Fri 12/5/08	Wed 1/14/09	9	100%
11	Power and cable tray -- concurrent with seismic	29 days	Tue 11/25/08	Fri 1/2/09	3,4,6	100%
12	Floor tile cuts	5 days	Mon 1/5/09	Fri 1/9/09	11	100%
13	Network Install	3 days	Thu 1/15/09	Mon 1/19/09	12FS+1 day,10	100%
14	Dawn Installation and Testing	57 days	Thu 1/8/09	Fri 3/27/09		100%
15	Contract Signed	0 days	Thu 1/8/09	Thu 1/8/09		100%
16	18 racks shipped/delivered	6 days	Tue 1/13/09	Tue 1/20/09		100%
17	Support racks shipped/delivered	6 days	Wed 1/21/09	Wed 1/28/09		100%
18	First 8 racks installed (Row 0/1)	13 days	Wed 1/21/09	Fri 2/6/09	16	100%
19	Testing/SWL prep on 4 racks	3 days	Mon 2/9/09	Wed 2/11/09	18	100%
20	Second 8 racks installed (Row 2/3)	4 days	Mon 2/9/09	Thu 2/12/09	18	100%
21	Row 4 shipped/delivered	4 days	Tue 2/3/09	Fri 2/6/09		100%
22	Row 4 installed	5 days	Mon 2/9/09	Fri 2/13/09	21	100%
23	Testing/SWL prep on 8 racks	5 days	Fri 2/13/09	Thu 2/19/09	20	100%
24	Row 5 shipped/delivered	4 days	Tue 2/10/09	Fri 2/13/09		100%
25	Row 5 install	6 days	Mon 2/16/09	Mon 2/23/09	24	100%
26	Row 6 shipped/delivered	4 days	Thu 2/12/09	Tue 2/17/09		100%
27	Testing/SWL prep on 16 racks	5 days	Fri 2/20/09	Thu 2/26/09	18,23	100%
28	Row 6 install	4 days	Wed 2/18/09	Mon 2/23/09	26	100%
29	Row 7 shipped/delivered	5 days	Mon 2/16/09	Fri 2/20/09		100%
30	Row 7 install	6 days	Mon 2/23/09	Mon 3/2/09	29	100%
31	Testing/SWL prep on 24 racks	5 days	Tue 2/24/09	Mon 3/2/09	25,27	100%
32	Row 8 shipped/delivered	3 days	Mon 2/23/09	Wed 2/25/09		100%
33	Row 8 install	3 days	Thu 2/26/09	Mon 3/2/09	32	100%
34	Testing/SWL prep on 36 racks	2 wks	Tue 3/3/09	Mon 3/16/09	33	100%
35	SWL Functional Tests	2 days	Tue 3/17/09	Wed 3/18/09	34	100%
36	SWL Performance Tests	2 days	Thu 3/19/09	Fri 3/20/09	35	100%
37	SWL Stability Tests	5 days	Mon 3/23/09	Fri 3/27/09	36	100%
38	System Accepted	0 days	Fri 3/27/09	Fri 3/27/09	37	100%
39	Dawn Science Runs	111 days	Mon 3/30/09	Mon 8/31/09	38	100%

Attachment 3: Handoff Letter



TO: LLNL ASC Office

FROM: Bradly K. Wallin

SUBJECT: **Completion of LLNL ASC Level 2 Milestone 3244**

As the identified LLNL nuclear weapons stockpile program representative, I certify the LLNL ASC Level 2 Milestone 3244, "Deploy Dawn ID Machine for Initial Science Runs", was completed by September 30, 2009.

During the Dawn science run period two application codes were ported to and run successfully on the system. The development environment, run-time system, and support subsystems such as the networks and file systems were sufficient to run the applications to completion and to store and analyze the results.

The ddmCD code was run on the full Dawn machine and used to study important plasma-beam interactions. The calculations provide insight into interactions that might occur in a fusion fast ignition experiment.

The ARES run on the full Dawn system (147,456 processors) shows the viability of threading across many processors, and is a proof of principle for the type of improvements that will be needed to run on Sequoia.

The details of the efforts to meet this milestone are documented in the "Level-2 Milestone 3244: Deploy Dawn ID Machine for Initial Science Runs," LLNL-TR-417191 (September 2009). A summary of results of the application runs on Dawn are included in the document titled "Selected Dawn System Initial Science Results", LLNL-PRES-416880 (September 2009).

Date signed:

Bradly K. Wallin

Manager Dynamic Material Properties Campaign &
ASC Physics and Engineering Modesl SubCampaign

Attachment 4: Science Run Results

See “Selected Dawn System Initial Science Results,” LLNL-PRES-416880 (September 2009).