



LAWRENCE  
LIVERMORE  
NATIONAL  
LABORATORY

# Level-2 Milestone 4797: Early Users on Max, Sequoia Visualization Cluster

K. C. Cupps

January 2, 2014

## **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.

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

## *Table of Contents*

<i>Introduction</i>	<u>1</u>
<i>Max Description</i>	<u>1</u>
<i>Attachment 1: Milestone Definition Text</i>	<u>2</u>
<i>Handoff Letter</i>	<u>3</u>

## Introduction

This report documents the fact that an early user has run successfully on Max, the Sequoia visualization cluster, ASC L2 milestone 4797: Early Users on Sequoia Visualization System (Max), due December 31, 2013. The full text of the milestone is included in Attachment 1. The description of milestone is:

*The Max visualization and data analysis cluster will provide Sequoia users with compute cycles and an interactive option for data exploration and analysis. The system will be integrated in the first quarter of FY14 and the system is expected to be moved to the classified network by the second quarter of FY14. The goal of this milestone is to have early users running their visualization and data analysis work on the Max cluster on the classified network.*

The milestone completion criterion is:

*Racks are assembled in B453 and the system has been moved to the classified network. A visualization user will write a memo certifying that he/she has run successful visualization jobs on Max on the classified network.*

The milestone was completed in early December of 2013. Max was delivered in August of 2013. Integration took place during September and October and early users ran on Max beginning in November of 2013.

A letter certifying that Steven Langer has successfully run his code on Max is included as Attachment 1.

## Max Description

Max is a 324-node cluster on the SCF network that supports Sequoia users' visualization and data analysis needs. It is two TLCC2-like scalable units with 280 compute nodes and 20 additional compute nodes that contain two Kepler K20x GPUs each. There are 6GB RAM per Kepler. On the 280 compute nodes there are 16 cores per node and 256 GB RAM per node, for a total of 76,800 GB of total system memory. The 24 remaining nodes of Max are used for Lustre routers, NFS gateways and login nodes. Bandwidth to Lustre is a peak of 78 GB/s. Max's peak performance is 107TF/s. The bandwidth to FLOPs ratio for Max is .24, a substantial increase over the .09 bandwidth to FLOPs ratio of Graph, our previous generation classified visualization cluster.

## Attachment 1: Milestone Definition Text

<b>Milestone (4797): Early Users on Classified Sequoia Hardware</b>		
<b>Level:</b> 2	<b>Fiscal Year:</b> FY14	<b>DOE Area/Campaign:</b> ASC
<b>Completion Date:</b> 12/31/14		
<b>ASC nWBS Subprogram:</b> FOUS		
<b>Participating Sites:</b> LLNL		
<b>Participating Programs/Campaigns:</b> ASC		
<p><b>Description:</b> The Max visualization and data analysis cluster will provide Sequoia users with compute cycles and an interactive option for data exploration and analysis. The system will be integrated in the first quarter of FY14 and the system is expected to be moved to the classified network by the second quarter of FY14. The goal of this milestone is to have early users running their visualization and data analysis work on the Max cluster on the classified network.</p>		
<p><b>Completion Criteria:</b> Racks have been assembled in the Terascale Simulation Facility (TSF), the system has been moved to the classified network, and at least one user has ported their code. A user will write a memo certifying that their code has run on Sequoia.</p>		
<b>Customer:</b> ASC/LLNL		
<b>Milestone Certification Method:</b> Professional report and handoff to ASC program.		
<b>Supporting Resources:</b> TBD		

# Handoff Letter

Dec 13 2013 12:31PM HP Fax

page 2



Lawrence Livermore National Laboratory

December 12, 2013  
ICCD13-084

Kim Cupps  
Division Leader, Livermore Computing  
Lawrence Livermore National Laboratory  
PO Box 808, L-556  
Livermore, CA 94551-9900

Re: Max

Ms. Cupps,  
I have run several simple tests on max and it seems to be working OK.

I use the yorick parallel framework for a variety of post-processing tasks. I tested the framework by running a small (512 core) pf3d test. The run worked as expected, so I shouldn't have any issues running post-processing of NIC simulations on max. That is good, because some of those simulations will really benefit from the large per node memory.

I used Visit to look at a HYDRA dump set and didn't notice any problems. I launched the Visit data engine on 16 nodes. I didn't have a viz script ready to run, so I made a very simple plot. This was a 6144 process run, which is about as large as we go with HYDRA at this time.

I copied some large datasets to lscratch1 using max. I had a lot of problems when I first tried. Adam Moody chased this issue for quite a while and eventually told me that "synchronous copies" would work. Once I switched to synchronous mode, my copies completed in a perfectly acceptable amount of time. lscratch1 runs Lustre on top of ZFS. ZFS wasn't reliable when used in "asynchronous mode", even though asynchronous mode works fine with ext3 sitting underneath Lustre. I don't regard this as a problem with max - it was a Lustre/ZFS bug. The version of chaos\_5 installed this week should fix the problem. I won't be able to verify that personally until after the holidays, but Adam should have run a test equivalent to what I was doing.

I think max is ready to move into production.

Sincerely,

A handwritten signature in black ink that reads "Steven Langer".

Steven Langer  
Physicist - AX Division  
Weapons and Complex Integration

Cc: File

An Equal Opportunity Employer • Lawrence Livermore National Security, LLC • Operated for the US Department of Energy  
P.O. Box 808 Livermore, California 94551-0808 • Telephone (925) 422-1100 • <http://www.llnl.gov>

