Print Page

Submission Details: bof171s1

Submitted: 2017-07-28 19:19 · Last updated: 2017-08-07 22:04

Title

Title: XALT, OGRT and Related Technologies: Job-Level Usage Data on Today's Supercomputers

Session Leader Information

Session Leader 1:

Name: Dr. Robert T. McLay
Email: mclay@tacc.utexas.edu
Company/Institution: University of Texas at Austin
Company/Institution 2: Texas Advanced Computing Center, University of Texas
Country: United States of America
Biography:

Dr. McLay received his Ph.D. from the University of Texas at Austin in Engineering Mechanics. He joined TACC in 2008 and now leads the Software Tools Group. He has had a long interest in Software tools. He is the developer of Lmod, an environment module tool. He is also interested in Parallel 3D finite elements for incompressible fluid flow and heat transfer.

Photograph:
Is this person on the Birds of a Feather reviewing committee? No

Additional Session Leader Information

Additional Session Leader 1:

Name: Mr. Georg Rath
Email: rath.georg@gmail.com
Company/Institution: rath.io
Company/Institution 2:
Country: Austria
Biography:

Georg slipped into data intensive computing while studying software engineering. He held a part time position at a scientific research institute doing image processing in Vienna and got insight into the unique challenges this environment brings with it. After finishing his bachelor he had to come up with a way to handle the large influx of data and subsequent processing of that data in that institute. From there he went on to drag the legacy infrastructure of a life sciences institute into the 21st century and while doing so got interested in ways of improving user experience when working with centralized computing services.

Photograph:
Is this person on the Birds of a Feather reviewing committee? No

BOF Topic Area

BOF Topic Area: Performance Measurement, Modeling, and Tools

Abstract

Abstract (Maximum 100 words):

Wouldn't it be great to figure out what programs run on your cluster? To figure out who even uses that highly optimized library that you built? Which of the hundreds of programs in your job uses the most resources? We discuss the merits of monitoring jobs down to the program/library level by showing what XALT and OGRT do. We have a look into the insights gained by running such software on large installations, on clusters that run everything from physics to life sciences and discuss analytics of gathered data. Audience is operations and users of supercomputers.

Long Description

Long Description (Maximum 500 words):

Let's continue talking about real, high value cluster analytics at the level of each job. We're interested in what users are actually doing: from applications and libraries, to prevent things in the way of successful research. Moreover, we want to do this for every single job running on our systems. This year we're especially interested in some of the next challenges, including (1) understanding the needs of non-MPI workflows that comprise half the user community; (2) putting usage data in the hands of end users interested in records of their own job-level activity. Among the emerging needs: tracking individual usage within other frameworks such as Python; (3) combining job level activity with job level metrics to enable users and operations personnel to understand the resource requirements of applications and use resources more efficiently.

XALT (xalt.readthedocs.org) is a battle-tested tool focused on job-level usage data; it enjoys a history of helping decision makers manage and improve their operations. A current list of centers that run XALT includes CSCS, NCSA, UK NCC, KAUST, NICS and TACC. Version 2.0 is now ready to begin tracking non-MPI workflows.

OGRT (https://github.com/georg-rath/ogrt) started out as a way to apply the capabilities of XALT to non-traditional workloads, using non-traditional technologies. Its focus is on real-time tracking of job level activity with lowest possible overhead.

Join us a far-ranging discussion that will begin with an overview of new XALT and OGRT capabilities before it ventures into broader strategic and technical issues related to job-level activity tracking.

Session format

How much of the session will be used for interaction between audience and session leaders/presenters? 50%
What is the primary format for content that does not directly involve audience discussion? Sequence of presentations
Does the BOF topic deal with commercial technology? Vendor-neutral

Description of the session format

Description of the session format (Maximum 150 words):
The first part will be a 5 minute introduction followed by 25 minutes of presentation from the two presenters. The remainder of the time will be devoted to discussion.

BOFs at recent SCs

Has your BOF been held at recent SC conferences? SC13
SC14
SC15
If so, approximately how many attendees did your BOF attract the most recent year it was held? 50-74

Scheduling Information

Preferred date and time: Wednesday 1:30 - 3:00pm
Amount of time requested: 1.5 hrs
Expected Attendance: 50
Keyword/Phrase 1: Job level tracking
Keyword/Phrase 2: Job monitoring
Keyword/Phrase 3:

Conference Presentations

Can SC archive and distribute your conference presentation? Yes

Acknowledgement

Acknowledgement: yes