SC'19 Workshop Proposal

6th International Workshop on HPC User Support Tools (HUST-19)

Ralph C. Bording (IBM Research), Elsa Gonsiorowski (LLNL),
Olli-Pekka Lehto (Jump Trading, LLC)

Abstract

Supercomputing centres exist to drive scientific discovery by supporting researchers in computational science fields. To improve the productivity of the user and the usability of systems in an environment as complex in a typical HPC centre they employ specialised support teams and individuals. The broad support effort ranges from basic system admin to managing 100's of Petabytes with complex hierarchal data-storage systems, to supporting high performance networks, or consulting on advanced math libraries, code optimisation, and managing complex HPC software stacks. Often, support teams struggle to adequately support scientists as HPC environments are extremely complex, and combined with additional complexity of multi-user installations, exotic hardware, and maintaining research software, to support 100's or even 1000's of HPC users can be extremely demanding.

The HUST workshop, has been the ideal forum where new and innovative tools such as XALT, SPACK and Easybuild, have been widely announced to the broader HPC community that results into creating the communities and special interest groups about these tools to support entire BoFs, Workshops and Tutorials on these tools at SC, ISC and other HPC conferences We will continue to provide the necessary forum for system administrators, user support team members, tool developers, policy makers and end users. We will provide a forum to discuss support issues and we will continue to provide a publication venue for current support developments. Best practices, user support tools, and novel ideas to streamline user support efforts at supercomputing centres are in scope for the HUST workshop.

Detailed Description

Motivation

Every high-performance computing (HPC) site exists to support its users, be they academic, governmental or industrial. Users run complex applications that stress the limits of HPC resources and software environments. Users come from a wide variety of fields, and typically they can only maximise their productivity with the help of committed user support teams. Support entails a wide range of tasks, e.g., installing and building (scientific) software packages; providing performance analysis tools; user training, education, and documentation; testing systems, and providing frequent system and OS upgrades.

Over the past few years, we have spoken with members of many HPC support teams, and we found that many sites struggle with providing adequate user support. There are two reasons: (i) lack of funding and therefore manpower, and (ii) lack of good tools to automate ubiquitous, labor-intensive tasks. Given the limited manpower, and given the commonality of tasks across both large and small HPC sites, it surprised us that there was not more collaboration among HPC centers. This would allow effort to be leveraged across different HPC sites. For this reason, we started the HUST workshop at SC14.

The HUST workshop at SC provides the forum for the dissemination of knowledge about best practices and tools for HPC user support teams. Our goal now is to continue the development of a community to aid HPC teams with their user support efforts and to publish peer-reviewed papers through the ACM Digital Library [Bor14] [Bor15] [Bor16] [Bor17] .

The HUST-18 workshop was on the Sunday morning at SC18 and we saw an increase in the attendees up to over 115 compared to 105 at HUST17 . The HUST18 workshop again scored well compared to the conference average in 'Technical Content', 'Quality of Presentations','Overall Value' and 'Workshop Program', and based on the attendees comments they found the workshop and the presentations of value as HPC practitioners.

Scope

To ease the burden of user support, HPC center staff must adopt collaborative methodologies and practices already used by the scientists and researchers they support: continual sharing, collaboration, and evaluation of methods and tools. The scope of this workshop is to provide a forum to bring together HPC user support teams, allowing them to share their ideas and experiences across a range of topics directly related to improving HPC user support. This includes, but is not limited to, raising awareness and increasing uptake of existing production and research tools. As a community we should be able to further improve these tools, fix their current shortcomings and evangelise their existence in a much more efficient manner. HPC user support teams often work in isolation, and this workshop will allow them to benefit from knowledge sharing and to reduce duplication of effort across sites. This in turn allows the community as a whole to benefit from these tools and improve the user experience on HPC systems, which hopefully results in maximising the scientific output of each HPC centre.

Topics of interest include, but are not limited to:

- **Defining and customising the user environment**
  HPC system environments are complex, dynamic, and customised heavily. We will discuss automatic, flexible customisation of the user environment with environment modules and other tools.
- **Software build and installation tools**
  Software build and installation tools provide automatic configuration, compilation and installation of scientific software – this is far from trivial to do correctly, especially when reproducibility of built packages is desired. This includes tools that simplify the installation of new scientific software packages with many versions and configurations It also comprises of tools to quickly install a complex software stack onto new systems.
- **Workflow and pipeline tools**
  Quite often, researchers manage their workflow using scripts in Bash, Python, Perl, etc. There are many workflow tools available that might be a better fit for increasing user productivity. Such tools typically allow managing complex data analysis work, automating parameter sweep experiments, provide portals or GUIs through which users can submit their jobs on the HPC infrastructure, etc.
- **Collaborative development tools**
  HPC centers have traditionally provided fast, well-tuned hardware, a shell console, and not much more. Modern development environment make use of wikis, bug trackers, repositories, build and test software, and other hosted software solutions. These can be difficult to deploy when there are many users, particularly if there are security concerns. We welcome all papers on enhancing the HPC development environment with modern HPC web tools.
- **Supporting Hadoop and Spark clusters for *Machine Learning* and *Big Data***
  Though we want to avoid placing too much emphasis on Big Data as such, experiments involving huge amounts of data do pose challenges to user support teams, either to conduct on existing infrastructure or because new dedicated infrastructure needs to be deployed. We expect tools that facilitate Hadoop or Spark workflows to become common in HPC. We are interested in how centers can efficiently maintain two seperate software toolchains, for Big Data using Machine Learning applications and for more traditional HPC.
- **System testing and monitoring**
  Tools that help to test HPC systems and to optimise the performance of an HPC center as a whole are within scope. This could include continuous monitoring or performance measurement tools, or test suites used regularly to establish baseline cluster performance.

Half-day Workshop Format

The format of the workshop will be fairly straightforward. We envision the following the schedule for SC19:
9:00 introduction and survey announcement (5 minutes)
9:05 paper session 1 (2 papers, 15-20 minutes presentations + 5-10 minutes questions)
10:05 break (20 minutes)
10:25 paper session 2 (3 papers, 15-20 minutes presentations + 5-10 minutes questions)
11:25 panel discussion (30 minutes)
12:30 End of workshop

During the introduction, we will briefly explain the goal and the scope of the workshop, and outline the survey

we plan to conduct during the workshop. Throughout the workshop, we will ask users to fill out an interactive user support survey. We have previously used various online surveys as part of the workshops. The surveys have been informative and has showed how the HPC community has changed over time to adopt tools and methodologies. As well in helping us respond to new or underdeveloped areas of interest. We expect this year's survey to be similarly edifying.

We plan to have two sequential tracks for peer review paper presentations, with 5 presentations in total. Depending on the topics of received submissions, these may be organized into two sets of related talks. Each presentation will be 25 minutes long. We will spur the discussion and encourage a strong interaction between the audience and the presenter, so the presentation should be limited to 15 minutes at most, leaving 10 minutes for discussion. We will end the workshop with a panel of experts from major HPC centers, and we will discuss, in light of the presented papers, what the biggest user support issues are. Last year we selected two papers to give extend presentation to allow them give more comprehensive demos of the tools that they had developed. We plan to have the workshop last about 3.5 hours (half-day).

Expected outcome

The main outcome of this workshop is to continue to foster the creation of new communities to further development of the tools and best practices for HPC application support and admin teams with their overall user support efforts. Additionally, this workshop will help focus on ways to improve collaboration efforts by bringing together more HPC centres that wish to share their successes and lessons learned. By encourage the reuse of past and current efforts and to learn from each other, because every HPC site faces very similar challenges at the end of the day.

We want to publish 5-8 peer reviewed papers through the IEEE TCHPC, from either the full day or half-day workshop. Finally, a workshop report will be written and distributed. Based on the success of the past workshops, we plan to continue this event and as a hub for the HPC user support community. We feel that the HUST workshop has been key helping raising the awareness of new tools such SPACK, Easybuild, LMOD, XDMoD and other tools that has resulted the creation of their own special interest groups, conference tutorials, BoFs and promoting those tools. We feel that the HUST workshop is achieving the overall objective of the HUST organising community, the objectives of the SC Workshop organisers and the need of the HPC community at large,

The increase in the number attendees who joined us last year, and the positive feedback in the reviews and the discussion after the workshop we believe that we are achieving our goals is within and that this workshop will continue to encourage better collaboration and uptake of new tools.

Advertising and soliciting submissions

We intend to distribute the call for papers for the workshop via the usual channels, including:

- Mailing lists and forums specific to HPC communities;
- use EasyChair's new CFP feature
- Social networks: Twitter, Reddit, IRC, . . . ;
- Promotion at various HPC events, such as the Cray User Group 2019 (Montreal, Canada May 2019, ISC'19 (Frankfurt - Germany, June 2019) and other HPC Conferences.

In addition, we will also contact relevant peers in our professional network, including the PRACE consortium, the EasyBuild, Spack and Maali build system communities, DOE National Laboratories, and other groups that include HPC sites worldwide.

Timeline

We plan to adhere to the following timeline:

- **Call for papers** issued and **submissions open**: end of April 2019 footnoteThe exact date depends on the IEEE-TCHPC notification.
- **Submission deadline**: August 30, 2019 (optional extension: September 7, 2019)
- **Workshop paper reviews**: by September 20, 2019
- **Final program committee meeting**: week of 20th September 2019
- **Acceptance notifications**: September 25, 2019
- **Camera-ready papers**: October 11, 2019
- **Workshop**: At SC'19

We will open submissions as soon a possible after acceptance notification, taking into account time required for confirmation of IEEE-TCHPC cooperation status (4 weeks). With acceptance notification planned for late March 2019, we should be able to open submissions **end of April 2019**.

The submission deadline is tentatively **August 30, 2019**, optionally extended to September 7, 2019 in case we are not satisfied with the number of submissions. We hope to get at least 10 submissions again. The workshop should take place in **November 2019 in Denver, Colorado**.

Potential program committee members

The following is a non-exhaustive list of potential invitees:

- Daniel Ahlin - (KTH)
- David Bernholdt - (ORNL)
- Mozhgan Chimeh - (Sheffield University)
- Erik Engquist - (Rice U)
- Wolfgang Frings, (JSC)
- Andy Georges - (U. Ghent)
- Markus Geimer - (JSC)
- Chris Harris - (Pawsey Supercomputing Centre)
- Paul Kolano, (NASA-Ames)
- John Linford (ARM)
- Robert McLay (TACC)
- Dave Montoya (LANL)
- William Scullin - (Argonne National Laboratory)
- Karen Tomko - (Ohio Supercomputing Center)

References

[Bor14] Bording, C. and Georges,A. ,"Proceedings of the First International Workshop on HPC User Support Tools," *SC14 International Conference for High Performance Computing, Networking, Storage and Analysis*, New Orleans, LA, 2014.

[Bor15] Bording, C., Gamblin, T., and Hansper,V. ,"Proceedings of the Second International Workshop on HPC User Support Tools," *SC15 International Conference for High Performance Computing, Networking, Storage and Analysis*, Austin, TX, 2015.

[Bor16] Bording, C., Gamblin, T., and Hansper,V. ,"Proceedings of the Third International Workshop on HPC User Support Tools," *SC16 International Conference for High Performance Computing, Networking, Storage and Analysis*, Salt Lake City, UT, 2016.

[Bor17] Bording, C., Gamblin, T., and Lehto, O. ,"Proceedings of the Fourth International Workshop on HPC User Support Tools," *SC17 International Conference for High Performance Computing, Networking, Storage and Analysis*, Denver, CO, 2017.