BoF: The Virtual Institute for I/O and the IO-500

Julian M. Kunkel[1], Jay Lofstead[2], John Bent[3]

[1] Deutsches Klimarechenzentrum
kunkel@dkrz.de
[2] Center for Computing Research
Sandia National Laboratories
[3] Seagate Government Solutions

1 Abstract

Due to the increasing complexity of HPC data management, activities in the storage research community have increased over the last few years. The general purpose of this BoF is to foster this community and discuss the role of the international Virtual Institute for I/O (VI4IO, http://vi4io.org) in supporting, developing, and maintaining this community. The speakers will give talks covering VI4IO and issues of benchmarking storage systems. The direction of these efforts is then discussed with the participants. A specific purpose of the BoF is to elaborate whether the community would be well-served by an IO-500 benchmark similar to the Top-500. Goals of the BoF are to 1) advertise the community hub but also discuss and steer the direction of the community effort, 2) to specifically identify possibilities to create a relevant community I/O benchmark that can be tracked with an IO-500 list. Expected HPC audience are 1) I/O experts from data centers and industry, 2) researchers/engineers working on high-performance I/O for data centers, 3) domain scientists and computer scientists interested in discussing I/O issues. The outcome of this BoF will steer the direction of the community efforts.

2 Organization

The workshop was structured into two talks followed by a discussion framed by the talks. A summary of the talks and discussion is provided in the following. During the BoF, we counted roughly 45 attendees with many storage experts.

3 Talks

In the talk "The Virtual Institute for I/O and the HPSL", Julian Kunkel introduced the virtual institute and its community hub by giving a demo of its outreach on the web page. VI4IO hosts information about research groups, I/O tools and the High-Performance Storage List (HPSL). Technically, the open institute is organized in a Wiki and supported by mailing lists. The High-Performance

Storage List hosts characteristics of data centers (sites), the deployed supercomputers and storage/file systems including near-line storage. It allows to describe the characterizations of each system individually since most sites share storage resources across multiple compute infrastructures. The list offers means to summarize and visualize data according to several easy to determine characteristics such as netto storage capacity and peak performance, and may list individual systems or aggregate data center resources. A preliminary analysis of storage capacity vs. memory capacity on the 30 systems of the list gives some incentive about the kind of meta-analysis that is be possible with such a list. Naturally, the HPSL list would benefit from including observable throughput inspired by application needs, this lead to the talk of the IO-500.

In the talk "The IO-500 benchmarking effort", John Bent presented a few straw person proposals for the structure of an actual IO-500. This included recommendations for using IOR and mdtest as well as preliminary thoughts about rules, parameters, and ranking systems.

4 Discussion

In the discussion lead by Jay Lofstead, many experts in addition to the organizers fostered information sharing about previous, failed efforts, and has lead to a kick off discussion that is now continued on the email list:
`https://www.vi4io.org/listinfo/io-500`
Several concerns regarding the definition of a new benchmark, its goal and target audience have been raised and controversially discussed even within the audience. The audience was split between those favoring moving forward with current well-defined and accepted benchmarks such as IOR and mdtest or creating new benchmarks. The aggregation of several metrics returned by different access patterns into one "single value" is difficult to judge and the fair ranking of the list may be non-trivial. Regardless, the value of the effort was agreed upon by the audience with hopes that acceptable benchmarks can be found. Overall, this lead to the conclusion that agreeing on individual benchmarks first, keeping the current order of the HPSL by usable storage capacity and the individual benchmarking results. The aggregation of these metrics into a metric that reflects the usefulness of a storage system, i.e., its balance, is delayed and can be added later on. Eventually the goal is to maintain a list sorted not by capacity but rather by "winner" once winner can be defined.

Heavy interest in the topic demonstrates community interest and motivates us to hold the BoF with updates next year again. An additional goal is to have a benchmark defined with some early results from sites such as ORNL to present by ISC in June.