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

The research community in high-performance computing is organized loosely. There are many distinct resources such as homepages of research groups and benchmarks. The Virtual Institute for I/O aims to provide a hub for the community and particularly newcomers to find relevant information in a more straightforward way. It hosts the comprehensive data center list (CDCL). Similarly to the top500, it contains information about supercomputers and their storage systems.

I/O benchmarking, particularly, the inter-comparison of measured performance between sites is tricky as there are more hardware components involved and configurations to take into account. Therefore, together with the community, we standardized an HPC I/O benchmark, the IO-500 benchmark, for which the first list had been released during supercomputing in Nov. 2017. Such a benchmark is also useful to assess the impact of system issues like the Meltdown and Spectre bugs.

This poster introduces the Virtual Institute for I/O, the high-performance storage list and the effort for the IO-500 which are unfunded community projects.

The Virtual Institute for I/O

Goals of the Virtual Institute for I/O (VI4IO) are:

- Provide a platform for I/O researchers and enthusiasts for exchanging information
- Foster training and international collaboration in the field of high-performance I/O
- Track and encourage the deployment of large storage systems by hosting information about high-performance storage systems

The philosophical cornerstones of VI4IO are:

- Treat contributors/participants equally
- Allow free participation without any fee inclusive to all
- Independent of vendors/research facilities

Open Organization

The organization uses a wiki as central hub

- Registered users can edit the content
- Mayor changes should be discussed on the contribute mailing list
- Tag clouds link between similar entities
- Supported by mailing lists, e.g.
  - Call-for-papers
  - Announcements
  - Contributions / suggestions

Community Content

The wiki covers AI worldwide research groups that address high-performance I/O including:

- A taglist for available knowledge
- Research products such as file systems
- Ongoing research projects

Flexible equations It supports equations to compute derived metrics, here easy_create / client_made:

As we can see, scalability of metadata is limited.

This can be used to create arbitrary new rankings. For example, to rank systems offering most MD performance per IOE: 

~~~
metrics = create easy_rat_eASY_delays, hard_rat, hard_rat, hard_rat
~~~

IO-Easy write, e.g., 2 x KBRs per GB throughput

Our Ongoing Work

- Supporting standardization efforts
- IO-500 benchmark
- Lossy compression interfaces
- Data center representation

Data Center List

The comprehensive data center list with its system model describes how characteristics are assigned to components. Storage is difficult to assign to a single component as it is often shared across supercomputers, therefore, a flexible component-based model is used.

Supported components:

- Site: Describes the facility
- Supercomputer: A system
- Storage system
- Nodes
- Network
- Building

The schema is under active development – we aim to describe data center characteristics. The web page allows the creation of a topology for the facility to indicate the relation between the components – ultimately multiple views will be created.

Metrics: Most metrics can be determined with existing measurement and describe hardware and software characteristics that should be known to the site and vendor. A few metrics cover actually observed metadata and I/O performance, in this case the measurement procedure must be defined. The list stores data entered in the wiki into a database and converts data to a base unit.

The following is an example of the schema for the DKRZ system:

Derived Analysis

With the collected data many in-depth analysis becomes possible, for example, the relationship between storage and memory capacity:

Ongoing Work

- Supporting standardization efforts
- IO-500 benchmark
- Lossy compression interfaces
- Data center representation
- IO-500 agenda:
  - June/July, proposal for extension rules
  - Extending schema
  - More HPSL sites
  - Support training and teaching for storage

IO-500 List Nov 2017

The normal list:

All results are available. The individual measurements for the benchmarks are stored and can be accessed.

Flexible equations It supports equations to compute derived metrics, here easy_create / client_made:

As we can see, scalability of metadata is limited.

This can be used to create arbitrary new rankings. For example, to rank systems offering most MD performance per IOE:

~~~
metrics = create easy_rat_eASY_delays, hard_rat, hard_rat, hard_rat
~~~

IO-Easy write, e.g., 2 x KBRs per GB throughput

The rules for determining performance are relaxed due to the complexity of I/O measurements, but this is augmented by the IO-500.