Visible to Everyone

Proposed Creation of IO500 Production and Research Lists

Authors: IO500 Steering Committee
Date: Apr 13, 2022

tl;dr

This document outlines a proposal to split the io500 into 2 lists: one that ranks production systems and one that ranks research systems. Each list would maintain both a full list and a 10-node (client) list. All existing and future entries would be in one of these two lists (but not both). The proposed timeline is to do an initial non-binding trial of this at ISC22 and, barring any major concerns or issues, implemented at SC22 and for all future lists.

Next Steps

- Complete reproducibility effort
- Create a sample split of the SC21 or ISC22 list

Motivation

Splitting list entries into two sub-categories has been discussed for quite some time due to the realization that different storage systems have different missions, and this directly affects how they are deployed.

For example, storage systems deployed by organizations (e.g., national lab supercomputers, university research centers) to run applications such as fraud analysis, physical simulations, genomic analysis, ML training, industrial image analysis, might be designed, deployed, and configured to ensure not only high-performance, but also to ensure high levels of availability, durability, and security to avoid data loss, extended down-time and intruders.

Whereas, storage systems that are deployed for the purposes of computer science research or testing (e.g., run storage CI tests, explore storage design tradeoffs, evaluate new storage hardware or software) might be configured very differently, such as by reducing the availability, durability, and security aspects to achieve some other goal such as maximizing performance.

By combining both types of systems in a single list as is done in the current IO500 list, concerns have been raised that may be preventing some community members from participating in the IO500. Further, the lack of detailed information regarding each entry can confuse practitioners looking to understand the suitability of a storage system for their use case (e.g., some current IO500 entries would lose data if a single device failed but others could lose three or more devices without incurring any data loss).

Visible to Everyone

Goals

Therefore the goals of this effort are as follows:

- Increase IO500 participation by creating a fair playing ground for all IO500 participants by comparing systems that are similar in their purpose and mission.
- Ensure the IO500 is useful for end users to evaluate storage systems by increasing transparency in the type of systems ranked
- Encourage additional participants in IO500 by giving assurances that submissions will be generally ranked with systems.
- The goal is to create a mechanism where all existing entries will be placed into a sub-category in future lists. An important constraint to any implementation is that we do not wish to restart the list contents from scratch, invalidating current submissions as this may discourage participation.

Proposal: Defining the Production List

The proposal is to create 2 lists:

1. **Production** - Systems deployed for the purposes of executing production scientific, industrial, or business applications
2. **Research** - All other systems. Generally includes systems deployed for testing, benchmarking, or systems research.

The key task is to define the criteria through which a submission is included on the Production list (since all other entries are included on the Research list). The definition is difficult (see background section below) due to the fact that the actual definition of the word production, which is something being produced, is not that helpful to the HPC storage community.

Production System: Proposed Definition

It is important to note that the IO500 steering committee has final say on whether a submission does or does not meet the definition of a production system and/or related definitions of its subcomponents.

The spirit of the Production System definition is that it is truly in place for the physical scientists, quants, security teams, data scientists over an extended period of time. Generally this means the cluster has a batch scheduler and queues up incoming Production Application execution requests. It also further means that the duration of deployment is much longer than a few days, weeks or months…and is typically measured in years.

Definition: Production System

An IO500 submission that:

- consists of a '*Compute System'* and '*Storage System'* that on a '*regular'*, '*frequent'*, and '*ongoing'* basis executes '*Production Applications'* that generate '*Production Data'*
- achieves the highest 'reproducibility score'
- has no 'single point of failure' in its '*Storage System'*

Visible to Everyone

Where the terms mentioned are defined as follows:
Definition 1: Storage System - The set of nodes and storage devices used by **Production** Applications to store **Production** Data and against which the IO500 benchmark suite is executed.

Definition 2: Compute System - The set of nodes that execute **Production** Applications and execute the IO500 benchmark. These nodes may overlap with those utilized by the storage system.

Definition 3: Production Application - An application that is executed on the Compute System during normal operation. This application MUST solve specific scientific or business problems and CANNOT be benchmarks, storage system software, or any other application whose purpose is purely motivated by computer science questions. Note that a build farm would count as a production application since it is using production data (i.e., code).

Definition 4: Production Data - The data stored in the Storage System during normal operation that is read by written by Production Applications. This data MUST have scientific and/or business value and CANNOT be a well-defined pattern (e.g., 0s, 1s, repeated hash) or algorithmically generated (e.g., random, a function without scientific/business value).

Definition 5: System Metadata - Any information tracked or stored regarding application or storage system execution. The point of defining System Metadata is to clarify that it is completely separate and not included as a type of Production Data since it is not directly generated by the Production Applications, but rather a set of information about the system and its behavior. Examples of System Metadata include logs (from applications like Splunk, the Production Application, or the Storage System), performance/operational metrics, etc.

Definition 6: Reproducibility Score - Quantifies the level of reproducibility of a submission. It is assigned by the IO500 Steering Committee to all IO500 submissions based upon the amount of information provided to enable others to reproduce the IO500 result. This information includes system metadata (e.g., number of compute nodes, storage device information), storage system metadata (e.g., RAID encoding, tuning parameters, find script) and answers to the reproducibility questionnaire on the details of how the benchmark was executed. For more information on the reproducibility score, please see the other proposal titled, “IO500 Submission Transparency and Reproducibility Proposal”.

Definition 7: Single Point of Failure - The ‘Storage System’ must be able to withstand any single failure of any component in its architecture. Upon a failure, while some amount of delay (on order of single digit minutes) is acceptable for the ‘Storage System’ to recover and become consistent, applications must not be disrupted and there must be no manual intervention. For example, a failure of a storage device, storage server, switch, network cable while executing the IO500 benchmark would in no way disrupt the execution of the benchmark, which would be able to finish successfully (with most likely a lower score than if the failure had not occurred).

It is worth clarifying a few terms in the definition of a Production System:

- *regular* - The system is (or will be) in place over an extended period of time and frequent large unexplainable gaps (e.g., days or longer) between the execution of Production Applications is unacceptable. A system that executed a few production applications last week and is planning to run more next week would not be a Production System. Note that maintenance periods during which production runs are paused is expected.
- *frequent* - Production Applications are continuously executed on the system, most likely using a scheduler and set of queues to stage the incoming application execution requests. Put another way, Production Applications must consume the vast majority of aggregate computational time on the compute system. This could include a few large and long running production application jobs, many smaller and short-lived production application jobs or anything in between.
- *ongoing* - The system is (or will be) executing Production Applications for the foreseeable lifetime of the machine. One time executions or short execution bursts would not qualify a system as production.

Production Systems and the Cloud

With 20% of HPC now estimated to be running in the Public Cloud, it is clear that there are at least some Public Cloud-based Production Systems. The goal of this proposal is to include cloud in a fair and reasonable way now that it is 20% and growing of the HPC community.

The definition of Production System above INCLUDES cloud, but there are several critical differences of note:

1. The Compute System may be extremely dynamic, growing from 1 node to 1000s of nodes in minutes. This doesn't change anything as long as those same types of compute nodes are executing Production Applications.
2. The Storage System may also be somewhat dynamic, and may even be shut down for short periods of time, but the general capacity and speeds/feeds should be similar to the Storage System that is used to execute Production Applications.
3. Many HPC cloud deployments are for burst use cases, where the Production System will vary in size and shape depending on the required compute/storage resources to augment the on-premise system. This continues to fit the above definition since Burst is not a one time activity of running Production Applications but rather a continuous activity that is by its very nature bursty.
4. To obtain the highest Reproducibility Score, any cloud-based submission must list all of the specific cloud vendor's compute/storage/networking offerings utilized so that anyone from the community could reproduce the IO500 results exactly assuming they could obtain the exact same storage system software.

Production Systems and Burst Buffers

Several emerging Compute Systems not only attach to a Storage System but also utilize a burst buffer, which is an intermediate storage layer positioned between the Compute System and the

Storage System. This could take the form of an additional shared storage system or utilizing storage devices installed on the compute nodes.

The key question of whether a Burst Buffer can be considered the **Storage System** in a Production System is whether or not it is utilized by the majority (if not all) of the time spent by Production Applications to read and write their Production Data. This in no way implies that a Production System couldn't have multiple Storage Systems, since a Production System typically is executing many jobs in parallel, and so the key is how idle is the Storage System in question.

For example, the current IO500 list contains several DDN IME entries, but the compute cluster may also have a Lustre or Spectrum Scale file system. If the Production Applications are utilizing both storage systems for a majority of the time then both storage systems would be valid entries on the Production List.. In another example, imagine a system does execute Production Applications with a Lustre-based storage system but its compute nodes also include NVMe devices and someone wants to write some software to enable the execution of the IO500 benchmarks against these compute node NVMe devices. Given that this new software is not being utilized by the majority of Production Applications, the submission would be listed on the Research list but they could list a submission with their Lustre system on the Production list.

Fault Tolerance Requirements

Production systems will have an additional fault tolerant requirement that there is "**No Single Point of failure in the Storage System**".

This information will be gathered via the metadata and questionnaire collection process that will be introduced as part of the reproducibility initiative. This information will be listed prominently on the IO500 list so it is clear to everyone what tradeoffs are employed to achieve the published score. Further, multiple submissions with different fault-tolerance/reliability mechanisms may be submitted and published in order to demonstrate the capabilities of a submission along different dimensions (although we may limit the total number that can be on a list).

While including this information is a major change from the existing list, the vast majority of production systems in existence can handle a single point of failure and so this requirement ensures the goals of splitting out the Production list are achieved. One type of system that could be affected are burst buffers, but the number of these being deployed appear to be dwindling and are in a small minority of the deployed production systems. The other type of system that will be affected is an ephemeral storage system that prioritizes performance over reliability, but generally that is also the tradeoff made by the systems on the Research list, and so that seems to be the better list for these submissions.

Examples of Production and Research Systems

Production System examples:

- A University Hospital, University department compute cluster running Slurm, LSF, HT-Condor, etc to execute research applications relating to medical discoveries, physics,

genomics, etc. The jobs could lead to drug discoveries, publications in Science or related journals, etc.
- A national laboratory supercomputer that executes mission critical applications by its own scientists and/or other scientific agencies or institutions. Many of these systems are on the Top500 (and hopefully all of them will submit to the IO500)
    - Note that in some cases a vendor may deploy a system at their own site prior to being installed at the final location. The temporarily deployed system would be considered a Production System as long as the system is planned to be re-deployed with the exact same configuration (e.g., fault tolerance mechanisms) at the production site and meets the goals. Note that the temporary system could be a smaller deployment of the larger upcoming production system.
- A commercial energy customer executes image analysis applications for themselves or on behalf of commercial customers seeking to find energy reservoirs. Note that energy is just one type of commercial customer. There are many different industries that run large compute and storage clusters that have Production Systems (e.g., chip design, aircraft/auto design, video creation/editing/rendering)

In each of these systems, the assumption is that the submission running on these systems utilized a primary storage system (there could be more than one) that was deployed with the compute system. Any other storage system utilized may result in the submission not being defined as a production system because it does not follow the definition above (i.e., only the primary storage system(s) would be executing 'Production Applications' that generate 'Production Data' on a 'regular', 'frequent', and 'ongoing' basis).

Research System Examples

- Systems utilized by university or industrial computer science research groups to build/evaluate/experiment with systems research. Generally these systems tend to mostly execute macro or micro benchmarks (e.g., IO500, IOR/MDTest, SpecSFS, Blast, Resnet-50) and tend to use data has no specific significance beyond the test or system operation (e.g., all 0s, random data, generated data to be non-compressible, patterns to detect corruption)
- System primarily executes applications that generate data that is thrown away and/or has no business or scientific significance. This system has no Production Data.
- The primary data generated on the system from executing applications is run time or data describing effects on the system. This system has no Production Data.

Production List and Existing IO500 Submissions

By default, all existing submissions will be placed on the Research list. If the owners of a previous submission want to have their entry moved to the Production list, then they must submit the additional required information to achieve the required reproducibility score and fill out the questionnaire to demonstrate their system meets the definition of a Production System. Given some time has passed and some info may no longer be available, the IO500 steering committee will work with submitters to ensure the best possible outcome.

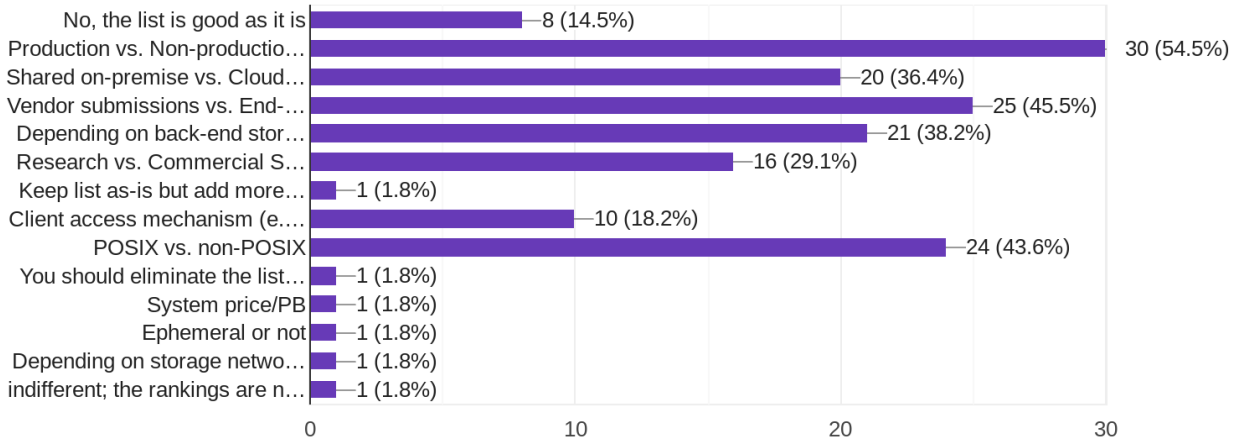
Visible to Everyone

Background: 2021 Survey Results

IO500: Accurate System Classification

Should we split the IO500 list somehow into sub-categories? (More lists may lead to more awards)

55 responses



In the 2021 IO500 community survey, 85.5% of respondents believe the list should be split into sub-categories. The exact path forward though is less clear, as no single option has a super majority (>66%) of support.

The most popular option, splitting based upon production vs non production, has 54.5% support among respondents.

The next most popular option, vendor vs end-user, in many ways could be considered a variant on prod vs non-prod. For example, if a vendor submits a test system, then it would end up in a non-prod category. If the system is for production users, then it really doesn't matter if the user or vendor submits the results as the same high level of integrity is required for either submission. Similarly with end users, if an end user submits a test system, it really isn't any more valid than a vendor doing so...and the same with production systems.

With the third most popular option, POSIX vs non-POSIX, this gets to the heart of what the IO500 is trying to achieve. A core tenant is to improve storage performance for HPC users, and the IO500 is in no position to dictate the best architecture. It is clear that HPC users are starting to explore many different types of storage systems (e.g., KV stores, object stores), with a variety of client access mechanisms and a wide range of storage hardware. The question here is more about whether the IO500 represents realistic I/O patterns within the HPC community and not whether they support POSIX or not. So while we respect the community's voice on this option, we must try to refer them to focusing on the workloads and I/O patterns and not the API or architecture.

Visible to Everyone

So the community generally is focused on creation of a production subcategory. The other submissions not meeting the requirements for the “Production” list will be in another list named “Research”.

When looking at the key features that define a production system, they are:

Feature	% of respondents
System for production applications (not just for testing)	90.70
System exists for reasons beyond just testing IO500 (and other benchmarks)	81.50
System for long-term usage (e.g. years, not necessarily the data retention lifetime)	68.50
System provides data redundancy/fault-tolerance to handle storage/server/network failures	64.80
System available to/used by end users	64.80
Software/hardware currently available to general public (i.e. can purchase/deploy a similar system)	61.10
System planned to operate for longer than one year,	40.70
Software/hardware currently installed and used by 3rd parties (beyond the developers)	31.50
System accessed by people outside a single organization	16.70
Software/hardware currently used at least in one Top500 site for production	14.80
Software/hardware currently deployed in at least two separate sites	14.80
System planned to operate for longer than one year	5.60
Instead of "available to general public" how about "can be independently verified in theory" - this precludes closed-source black box software without requiring open-source or commercial support	1.90

Top 6 Features

It is clear in the definition of a production system, we should include the top 2 items that each received over 80% support on ensuring the system isn’t just for testing or for the sole purpose of running the IO500. Past these 2 top items, there are 4 more characteristics that each got over 60% support.

Visible to Everyone

We tried to take all 6 of these into account in this proposal (or at least their intention and or the spirit in which they were selected). This includes ensuring that the Production list is not just for testing, is running user jobs on a regular, frequent, and ongoing basis, and that it has a high reproducibility score.