

Establishing the IO-500 Benchmark

Julian M. Kunkel, John Bent, Jay Lofstead, George S. Markomanolis

2017-11-13

<http://www.io500.org>

IO⁵⁰⁰

v4io

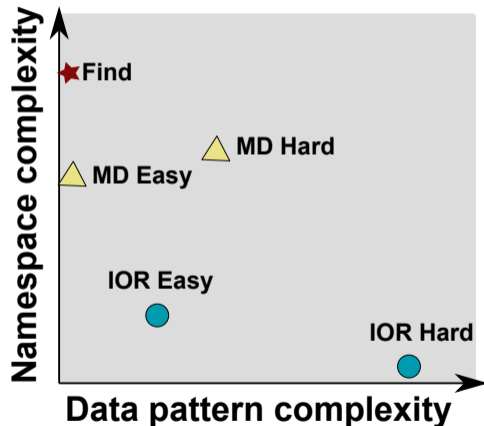
The IO-500

Goals

- Tracking storage performance
- Sharing best practices

Benchmarking Approach

- Community driven effort
- Patterns: metadata, data, search
 - Easy for optimized patterns
 - Hard for naive patterns
- Relies on community benchmarks

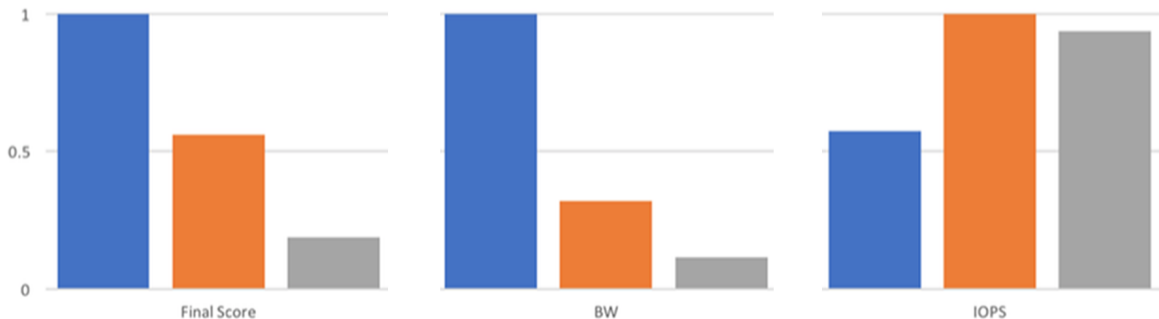


List Results from BeeGFS, DataWarp, IME, Spectrum Scale, Lustre

IO-500 Normalized

System ranked according to score

■ #1 ■ #2 ■ #3



Challenges of Establishing the Benchmark

This is a short summary of experience gained by

- Feedback from discussions
 - From SC/ISC BoFs
 - Peers
- Feedback of people executing the IO-500 on different systems
- *Thanks to everybody contributing*

Challenges & Approach

Representative of applications and user requirements

- Supply workloads providing
 - Upper bound for optimized applications
 - Performance expectation for non-optimized applications
- *More workloads and concurrent execution to be integrated*

Understandable and human comprehensive results

- Report meaningful metrics
- Imply low variability of repeated measurements
- Computing of an overall score for ranking but retain individual values

Challenges & Approach

Portable

- Ran into Python (Shell) portability issues
- C-APIs: `readdir()` does not return type on DataWarp
- Non-POSIX `stat()` call on one system

Inclusive: cover various storage technology and non-POSIX APIs

- Allow vendors to use specific optimizations (for easy runs)
 - Enable replacement for `find` (IBM Spectrum Scale has optimizations here)
- Relying on (IOR's) AIOR interface (thanks to Nathan for porting `mdtest`)
- *We are still the process to support more storage APIs*

Challenges & Approach

Scalable, i.e., run on large-scale computers and relevant storage systems

- IOR and mdtest are MPI parallelized
- Supply a parallel find version

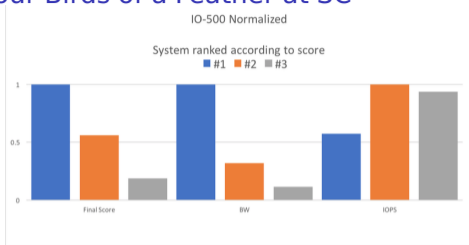
Lightweight: easy to setup and cheap to run

- 5 minute write/creation phases to limit runtime
- Extended IOR/mdtest for phase-out stonewalling options

Trustworthy: prevent (unintended) cheating

- Reveal all tunings made (also shares best practice)
- Sufficiently large working set

Visit our Birds of a Feather at SC



IO500

Getting Stared with IO500

- git clone <https://github.com/VI4IO/io-500-dev>
- cd io-500-dev
- ./utilities/prepare.sh
- ./io500.sh
- # Tune and rerun until good
- # email results to submit@io500.org

Contact us

<http://www.io500.org>

Slack: vi4io.slack.com

Twitter: [IO500benchmark](https://twitter.com/IO500benchmark)

**Come see the full IO-500 results at SC17 BOF
Wednesday, 15 November, 17:15, room 201-203**

**"Results from ThinkParQ BeeGFS, Cray DataWarp, DDN IME,
IBM Spectrum Scale, and Lustre!"**

